The effect of stress on the explore-exploit dilemma

By

Thomas Donald Grant Ferguson

Bachelor of Science, University of Victoria, 2014

Master of Science, University of Victoria, 2016

A Dissertation submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSPHY

in the Department of Psychology

University of Victoria

Supervisory Committee

The effect of stress on the explore-exploit dilemma

By

Thomas Donald Grant Ferguson

Bachelor of Science, University of Victoria, 2014

Master of Science, University of Victoria, 2016

Supervisory Committee

Dr. Olave E. Krigolson (Department of Psychology)

Co-Supervisor

Dr. Michael E. J. Masson (Department of Psychology)

Co-Supervisor

Dr. Todd C. Handy (Department of Psychology, University of British Columbia)

Outside Member

Abstract

When humans are faced with multiple options, they must decide whether to choose a novel or less certain option (explore) or stick with what they know (exploit). Exploration is a fundamental cognitive process. Importantly, when humans attempt to solve the **explore-exploit dilemma**, they must effectively incorporate both feedback and uncertainty to guide their actions. While prior work has shown that both acute (short-term) and chronic (long-term) stress can disrupt how humans solve the explore-exploit dilemma, the mechanisms of how this occurs are unclear. For example, does stress disrupt how people integrate feedback to guide their decisions to explore or exploit, or does stress disrupt computations of uncertainty regarding their choices? Importantly, the use of electroencephalography as a tool can help reveal the impact of stress on explore-exploit decision making by measuring neural signals sensitive to feedback learning and uncertainty. In the present dissertation, I provide evidence from a series of experiments where I examined the impact of both acute and chronic stress on the explore-exploit dilemma while electroencephalographic data was collected. In **experiment 1**, I exposed participants to an acute stressor and then examined their decisions to switch or stay – as a proxy for explore and exploit decisions – in a multi-arm bandit paradigm. I found tentative evidence that the acute stress response disrupted both the feedback learning signal (the reward positivity) and the uncertainty signal (the switch P300). In **experiment 2** I adopted a computational neuroscience approach and directly classified participants decisions as explorations or exploitations using reinforcement learning models. There was only an effect of the acute stress response on feedback signals, in this case, the feedback P300. In experiments 1 and 2, I used contextual bandit tasks where the reward probabilities of the options shifted throughout, and there was no behavioural effect of acute stress on task performance or exploration rate. However, in **experiment 3**, I examined a learnable bandit where one option was preferred. Again, using computational modelling and electroencephalography, I found tentative evidence that the acute stress response disrupted the feedback learning signals (the feedback P300) and stronger evidence that acute stress disrupted the uncertainty signal (the exploration P300). As well, I observed that the acute stress response reduced task performance and increased exploration rate. Lastly, in **experiment 4**, I examined the impact of chronic stress exposure on explore-exploit decision making and electrophysiology – while I found no effects of chronic stress, I believe future research is necessary. Taken

together, these findings provide novel evidence for the neural mechanisms of how the acute stress response impacts the explore-exploit dilemma through disruptions to feedback learning and assessments of uncertainty. These findings also highlight how theories of the P300 signal may not be properly capturing the varied role of the P300 in cognition.

Table of Contents

List of Tables

List of Figures

Abbreviations

Neurophysiology and Stress

EEG	Electroencephalography
ERP	Event Related Potential
HPA	Hypothalamic Pituitary Adrenal
SAM	Sympathetic Adrenal Medullary
TSST	Trier Social Stress Test

Questionnaires

PANAS	Positive and Negative Affect Schedule
STAI – S	State Trait Anxiety Inventory – State scale
STAI – T	State Trait Anxiety Inventory – Trait scale
TICS	Trier Inventory for the Assessment of Chronic Stress

Modeling and Statistics

AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
95% CI	95% Confidence Interval
η_p^2	Partial Eta Squared
MSE	Mean Squared Error

Acknowledgements

There are several individuals whose help I would like to acknowledge for this dissertation. Big thanks to Dr. Cameron Hassall and Chad Williams for helpful discussions throughout my time in the lab – and for being willing to take the occasional afternoon off at the Grad House when breaks were needed. I'd like to acknowledge Rob Trska and Wande Abimbola for their work as lab managers and for putting up with my attempts at humour. As well, thanks to all the other Theoretical and Applied Neuroscience Lab members (too many to name!) and any research assistants who helped with the collection of data. I'd also like to thank Dr. Ron Skelton and Dr. Dustin van Gerven as their mentorship during my Master's degree helped lay the groundwork for the skills I needed to do a dissertation.

I'd like to acknowledge Dr. Anna Weinberg for graciously acting as the external and providing some great comments and questions about the dissertation. Big thank you to all my supervisory committee members. Thanks to Dr. Todd Handy for providing his thoughts and making me think critically about the work. Thanks to Dr. Michael Masson for acting as my co-supervisor. Mike was always willing to take the time to answer my questions and through his mentorship I learned a huge amount about being a careful cognitive scientist. Finally, a huge thanks to my other co-supervisor Dr. Olav Krigolson. I wouldn't be where I am without your mentorship and support.

Dedication

To my parents for their love and support.

Chapter 1 – General Introduction

Perhaps one of the more interesting questions about both cognition and behaviour is the question of how the systems that guide an organism's behaviour function when they are being heavily taxed or broken. For example, our understanding of cognitive control has been enhanced through the study of how pain impacts cognitive control behaviour (e.g., Shackman, Salomons, et al., 2011). A famous example of how dysfunction can be used to understand cognition arose out of research on the patient H.M (Scoville & Milner, 1957). H.M. was a neurological case study whereby a patient had their temporal lobes severed to treat epilepsy which led to significant detriments of the patient in certain aspects of memory. Specifically, H.M. suffered from anterograde amnesia and was unable to create and remember new memories, although his procedural memory was entirely maintained. Although a sad case, the severing of H.M.'s temporal lobe provided insight into how memory works and generated a variety of theoretical approaches to memory research (Squire, 2009). While not as extreme as surgical severing of brain regions, the stressors which organisms face and overcome regularly can provide a manipulation for understanding cognition and the brain-behaviour relationship. In fact, an important aspect of human and animal behaviour is our ability to overcome stressors. There is a large body of work showing the impact of stressors and the stress response on cognition and the brain (for reviews see Lupien et al., 2007; McEwen, 2007; Porcelli & Delgado, 2017; Shields et al., 2016; Starcke & Brand, 2016). The primary goal of my dissertation is to provide evidence from a series of experiments examining decision making and learning under stress to better understand cognition.

Stress itself can be a complicated construct to define, but the impact of stress on human and animal behaviour, and in particular the decisions organisms make, has been investigated

across a variety of disciplines. For example, the role of stress in decision making has been investigated in such diverse fields as education (L. Morris, 2005), economics (Delaney et al., 2021), healthcare (Groombridge et al., 2019), and, perhaps most importantly for this dissertation, cognition (Starcke et al., 2017). While there is a wealth of knowledge of how stress impacts the decisions that animals and humans make, my goal in this dissertation is to contribute to this literature through a novel approach combining methods from psychology, artificial intelligence, and neuroscience. Specifically, I used event-related potentials (ERPs) and computational modelling to understand how humans solve what is known as the explore-exploit dilemma under stress.

The explore-exploit dilemma is a fundamental decision process whereby people must decide whether to stick with what they know (exploit) or choose a novel or less-known option (explore). Solving the explore-exploit dilemma typically involves cognitive processes such as reward valuation, feedback learning, motivation, and assessments of uncertainty/risk (J. D. Cohen et al., 2007). In this dissertation the primary focus will be on the impact of stress on how people are able to manage the trade-off between decisions to explore and decisions to exploit to effectively guide behaviour. Importantly, the explore-exploit dilemma lends itself well to event-related potential designs (Hassall et al., 2013), has clear and direct links to the neuromodulators that play a role in the stress response (dopamine and norepinephrine; Aston-Jones & Cohen, 2005), and involves key aspects of decision making such as feedback-learning and computations of uncertainty.

In this dissertation, a general introduction to the important concepts will first be provided. Specifically, I will provide a background on the neurobiology of the stress response followed by a discussion of how stress is distinguished and studied in the literature. Following this, I will

discuss the current research on stress and decision making – with an emphasis on explore-exploit decision making – and the relationship between two common ERP components (the reward positivity and the P300) and stress. The general introduction will conclude with a description of the four experiments of this dissertation.

Neurobiology of stress response

When we are faced with a threat in our environment, we need to mobilize both bodily and cognitive resources to deal with that threat. In cases where the threat is particularly pernicious and needs to be dealt with quickly, the stress response begins. During the stress response, cognitive and bodily resources are marshalled to respond to the threat but after a certain amount of time the brain and body begin to suffer from the overexertion of the response (McEwen & Wingfield, 2003). For example, imagine an animal threatened by a predator. The animal would immediately summon all its energy to escape the predator, but after a time the resources needed would begin to diminish and the animal would become exhausted. In humans, stress comes in many forms, not just the (rather unlikely) threat of a predator. That is, stress in humans can also be due to external events such as deadlines at work, family problems, or interpersonal disagreements. During the stress response, when a threat has been identified, specific parts of the brain are mobilized to respond to this stressor. Unsurprisingly, work has emphasized that the brain is the central aspect of the stress response and is key to the adaptation needed to overcome stress (McEwen, 2007). The following section will outline the nature of the stress response. The initial discussion will begin with the differences between the Sympathetic-Adrenal-Medullary (SAM) axis and the Hypothalamic-Pituitary-Adrenal (HPA) axis. Then, I will discuss the role of

different catecholamines in the stress response, with an emphasis on norepinephrine and dopamine.

During the stress response, there are two primarily active pathways, the SAM axis (Cannon, 1914) and the HPA axis (Selye, 1950). These two axes can be distinguished not just in their neural targets but also in their timeframe, as the SAM axis is fast – on the order of seconds – while the HPA axis is slow – on the order of minutes to hours (Hermans et al., 2014). On a hormonal level, the two responses begin with the release of corticotrophin-releasing hormone, which causes an increase in the activity of both the SAM axis and HPA axis (De Kloet et al., 2005). For the stress response to begin, a stressor must be perceived as having been occurred which differs for psychological and physical stressors. In the case of a psychological stressor this perception will involve appraisal from the pre-frontal cortex or amygdala while in the case of a physical stressor this perception will involve signals such as pain. Once a stressor is perceived as having occurred then the paraventricular nucleus will activate the SAM axis and, in some situations, the HPA axis (Ulrich-Lai & Herman, 2009). The paraventricular nucleus, located in the hypothalamus, is where descending autonomic fibers originate, highlighting its prominent role in the SAM axis stress response (Pacak et al., 1995). As noted by Pacak and colleagues, the paraventricular nucleus also has connections to the anterior pituitary – which plays a role in the HPA (slow) stress response, as the anterior pituitary releases adrenocorticotropin releasing hormone leading to the release of cortisol, the main stress hormone. However, following this initial activation by the paraventricular nucleus, the SAM and HPA responses begin to diverge.

For the SAM axis, the stress response can begin anytime a threat is detected, and the detection of a threat can occur in two manners. The first is a bottom-up approach involving pre-ganglionic autonomic neurons in the brainstem that detect a homeostatic/allostatic imbalance,

and the second is a top-down approach involving signals from the limbic forebrain and pre-frontal areas (Ulrich-Lai & Herman, 2009). An example of the bottom-up detection could be a threat such as high levels of pain or discomfort, while an example of a top-down detection could be a psychological stressor such as having to give a speech in front of an audience. In the SAM axis, corticotrophin-releasing hormone leads to an increase in tonic norepinephrine activity throughout the brain by directly stimulating the locus-coeruleus norepinephrine system (McCall et al., 2015), which in turn causes general activation of the sympathetic nervous system and the flight-or-fight response. This increase in activity of the sympathetic nervous system induces a state of vigilance, anxiety, and heightened attention to threat (Hermans et al., 2014; Sun et al., 2015). SAM axis activity leads to increases in norepinephrine in the cortex (particularly the pre-frontal cortex), the hypothalamus, and sub-cortical limbic areas such as the amygdala, striatum and hippocampus (Foote & Morrison, 1987; Goldstein et al., 1996). A frequent measure of SAM axis activity is salivary-alpha amylase (Nater et al., 2006), and work has shown that alpha-amylase levels peak immediately after a stressor and remain elevated for up to 25 minutes post-stress (Het et al., 2009) – suggesting that SAM axis activity remains elevated for these 25 minutes.

With the HPA axis, the stress response is more complicated. Unlike the SAM axis, initially a stressor must be perceived as stressful (whether physical or psychological) to activate the HPA axis, and the initial appraisal occurs in regions implicated in the limbic system, including the amygdala and the pre-frontal cortex (Ulrich-Lai & Herman, 2009). The activity of the HPA axis is much slower – on the order of minutes to hours – and there is evidence that long-term genomic changes occur to compensate for the rapid effects of the SAM axis (Hermans et al., 2014). In the HPA axis, corticotrophin-releasing hormone and vasopressin (acting in concert)

are released by the paraventricular nucleus in the hypothalamus, and then corticotropin-releasing hormone activates the release of corticotropin in the anterior pituitary which finally ends in the release of glucocorticoids such as cortisol in humans or corticosterone in animals (Groeneweg et al., 2011).

Glucocorticoids are hormones that affect two types of receptors that are located throughout the brain: mineralocorticoid and glucocorticoid receptors (Pearce & Yamamoto, 1993). During periods of normal (i.e., non-stressed) brain functioning, mineralocorticoid receptors are mostly saturated while glucocorticoid receptors are barely saturated (Lupien et al., 2007). When behaving normally, the hippocampus – which has the largest amount of corticoid receptors – functions as a negative feedback loop, inhibiting the activity of glucocorticoids, preventing the cascade of the stress response (Sapolsky et al., 1986). However, following the activation of the HPA axis stress response, mineralocorticoid receptors become fully saturated almost immediately and glucocorticoid receptors become fully saturated following a delay (Lupien et al., 2007). Saturation of the glucocorticoid receptors induces the cognitive effects of the HPA axis stress response. While there are glucocorticoid receptors located throughout the brain, the three regions of the brain that have the highest concentration are the hippocampus, the amygdala, and the pre-frontal cortex (Herman et al., 2005). Given the ease of collection and analysis, salivary cortisol is the most common measure of HPA axis activity (Kirschbaum et al., 1992), although some work has used serum cortisol (Vining et al., 1983) or adrenocorticotrophic hormone (Bornstein, Engeland, Ehrhart-Bornstein, & Herman, 2008). Much like during SAM axis activity, catecholaminergic activity is occurring in tandem with glucocorticoid activity (Godoy et al., 2018).

During both the SAM and HPA axis stress responses, changes to norepinephrine and dopamine levels occur. Norepinephrine strongly increases in the pre-frontal cortex during stress exposure, and moderately in the pre-frontal cortex after the cessation of the stressor (Cabib & Puglisi-Allegra, 2012).Work on the role of norepinephrine has found that stress-related norepinephrine increases activity in brain regions such as the amygdala, striate terminus, the medial pre-frontal cortex, and the lateral septum causing concomitant changes in behaviour (Morilak et al., 2005). It should be noted that these authors found that norepinephrine returns to baseline 30 minutes post-stressor, suggesting a quicker time frame for the effects of norepinephrine, a fact that makes sense when we consider how much of its activity is tied to the fast-acting SAM axis stress response. Other work suggests a longer timeframe for these norepinephrine effects as Finlay, Zigmond, & Abercrombie (1995) observed that norepinephrine in the medial prefrontal cortex of rats was increased relative to baseline for up to 60 minutes post-stress – although norepinephrine activity peaked immediately post-stress. A recent review concluded that it is actually tonic norepinephrine activity which increases during stress exposure and the author suggested this may be related to the decrease in pre-frontal cortex function during a stressor due to modulation in the signal-to-noise ratio of neurons (Arnsten, 2015). Typically in humans, salivary alpha-amylase levels peak immediately following a stressor and then return to baseline around 20 minutes post-stressor onset (Nater et al., 2006). However, that is not to suggest that SAM axis activity has entirely returned to baseline at that point, as serum norepinephrine levels remain elevated up to 40 minutes post-stressor (Thoma et al., 2012). Overall, the largest increases in norepinephrine activity occur in the 10 minutes post-stressor, although longer time-frame effects may be present for up to an hour.

During stress, dopamine levels throughout the brain are also modulated. Immediately following a stressor tonic dopamine levels in the nucleus accumbens spike (Cabib & Puglisi-Allegra, 2012). Results from a PET study in humans have shown that dopaminergic activity in the ventral striatum increases during stress exposure and is tied to the activity of cortisol (Pruessner et al., 2004). Following a stressor being dealt with, glucocorticoid activity causes large increases in dopamine in the pre-frontal cortex (Cabib & Puglisi-Allegra, 2012). Correspondingly, acute stress exposure in rats leads to an increase in extracellular dopamine in the medial pre-frontal cortex up to 90 minutes post-stress (Butts et al., 2011; Finlay et al., 1995). There is also evidence of increases in dopamine in the ventromedial pre-frontal cortex for up to 60 minutes post-stress exposure (Lataster et al., 2011). Most importantly, research has shown that stress may impair prefrontal cortex function due to the overactivity of dopaminergic neurons, as stress seems to mimic the behavioural effects of prefrontal cortex lesions (Arnsten & Goldman-Rakic, 1998). Relatedly, following acute stress exposure, prefrontal cortex dysfunction that is tied to dopaminergic activity typically leads to an adoption of habitual behaviours to the detriment of cognitive flexibility (Arnsten, 2009).

Definitions of Stress

While different definitions of stress have been provided by researchers, a uniting theme across many of these definitions is the concept of homeostasis. One of the first uses of the term "stress" in a biological context came from Hans Selye's work, where he defined stress as the state of an organism as it adapts and defends itself from a threat to its life (Selye, 1950). As has been pointed out elsewhere (e.g., Lupien, Maheu, Tu, Fiocco, & Schramek, 2007) the concept of stress was borrowed from physics and engineering where stress is defined as the internal force of

an object that is generated by the external physical strain from a different object. Taken together, this definition of stress suggests that stress and the stress response are the result of some internal or external force acting on an organism that the organism has an adaptive physiological response to. Furthermore, the idea of homeostasis – that is, the bodily functions that maintain the steady-state of an organism (Cannon, 1929) – was also integrated by Selye, further solidifying the concept that the physiological stability of organisms was a primary goal of organisms. Recent work has continued to use this definition of stress as the threat to homeostasis (Ulrich-Lai & Herman, 2009).

However, still other work has added the concept of allostasis and allostatic load to defining the construct of stress (see McEwen & Wingfield, 2003). The concept of allostasis adapts the idea of the steady state of homeostasis (which is specific to physiological systems) and suggests that allostatic systems are the physiological and behavioural processes that are used to achieve stability through adaptation. This concept of allostasis contributes the idea of an allostatic state, and more importantly for stress research, the concept of allostatic load and overload. In other words, while our allostatic systems attempt to keep us in a generalized state of adaption and comfort, at times the current state (the allostatic state) will be subject to environmental threats or demands that require energy mobilization (allostatic load). However, in some situations these demands can be too much which leads to allostatic overload, which causes the stress response and the release of stress hormones (glucocorticoids) and modulations of catecholamines in the brain (McEwen, 2007). Importantly, these definitions of allostasis incorporate the activity of SAM Axis and the HPA axis working in concert to respond to keep allostasis balanced (De Kloet et al., 2005; McEwen & Wingfield, 2003).

When considering the stress response and how it impacts behaviour and cognition, one factor that must be considered is the time-course of the stress response. That is, we can differentiate between what is termed acute stress (i.e., more immediate) in comparison to chronic stress (i.e., more long-term). Generally acute stress refers to a stressful event that occurs for a short duration (less than an hour) and usually cognitive assessments are made following the acute stressor (e.g., Duncko, Cornwell, Cui, Merikangas, & Grillon, 2007; Elzinga & Roelofs, 2005; Lighthall et al., 2012). Some examples of acute stressors include giving a public speech in front of an audience (Kirschbaum et al., 1993), dunking one's hand in cold water for three minutes (Hines & Brown, 1932), or even a combination of both approaches (Schwabe et al., 2008; Schwabe & Schächinger, 2018). In contrast, typically chronic stress is assessed in humans through the number of stressful or anxiety producing events that have occurred in either the last month (S. Cohen et al., 1983) or the past three months (Schulz & Scholtz, 1999). Chronic stress in humans must be measured using a survey or through an interview. Thus, for the present work we can differentiate stress that occurs once on the day of assessment (acute stress) and stress that occurs repeatedly for longer than one month in the lead up to the assessment (chronic stress). In addition, we can differentiate cognitive and behavioural effects due the stress response (e.g., caused by the activity of HPA or SAM axis) and effects caused by stress exposure (which elicits HPA and SAM axis responses). Generally, stress exposure causes a stress response and hormonal measures such as cortisol or salivary-alpha amylase can provide measurements of the stress response. Thus, throughout the dissertation I will define when I am examining the stress response as compared to stress exposure.

When considering acute stress specifically, we must operationally define what is meant by a "threat" to an organism in the case of a stressor to provide a working definition of the acute

stress response. Here then, it is useful to differentiate between a physical stressor and a psychological stressor. The definition of threat for a physical stressor is straight forward, as it is any external event that threatens the integrity of the organism on a physical level (such as pain or inflammation). For example, it could be exposure to ice cold water (which is used in the Cold Pressor Task – Hines & Brown, 1932), or it could be the exposure to electrical shocks (e.g., Shackman, Maxwell, McMenamin, Greischar, & Davidson, 2011). Importantly, conscious perception of the stressor is not needed to begin a physical stress response (Morilak et al., 2005). In contrast, the definition of a threat in the case of a psychological stressor is less clear, given that events that psychologically threaten an individual are varied and necessarily contain a subjective aspect. However, seminal work by Dickerson & Kemeny (2004) provides an influential theory of what causes the stress response under psychological stress. Their work defined a stressor as an event that contains motivated performance, uncontrollability, and socio-evaluative threat (in other words, threat to self-esteem). The authors conducted a meta-analysis and found that, in fact, stressors that could be classified as containing the three above aspects reliably elicited the largest stress response in terms of cortisol released across all studies (Dickerson & Kemeny, 2004). Recent work has also attempted to combine both physical and psychological stressors through the creation of the socially evaluated cold-pressor task in which participants complete a physical stressor (the Cold Pressor Task) while being assessed by a panel to induce socio-evaluative threat (Schwabe et al., 2008).

In sum, we can now synthesize a working definition of what stress means in the present work. Stress is defined as a threat to allostasis (i.e., an imbalance of homeostasis), reaching allostatic overload which then requires the mobilization of bodily and cognitive resources. Specifically, we can differentiate between stress exposure and the stress response – as stress

exposure is the threat or event which in turn causes the stress response. As well, for our definition we can distinguish between the two parts of this stress response: (1) the fast-acting SAM axis response that raises sympathetic nervous system activity and (2) the slower-acting HPA axis response that raises cortisol levels. We can define a physical stressor as any stressor that incorporates physical discomfort or the threat of injury or discomfort and a psychological stressor as any tasks that contain uncontrollability and threats to the social-self (socio-evaluative threat). Finally, we can distinguish between chronic stress (long term stress, occurring over a period of weeks to months) in comparison to acute stress (the immediate changes brought about by a single stress exposure). Generally, in this work I examined the acute stress response caused by acute stress exposure. More specifically, I was interested in the cognition and behaviour changes brought about by the acute stress response (experiments 1 through 3). In the case of chronic stress (examined in experiment 4), I considered chronic stress as repeated stress exposures and the impact of those repeated exposures on cognition and behaviour.

How to Induce the Acute Stress Response

There are a number of different approaches used in the literature to induce the stress response – including both physical stressors such as the Cold-Pressor Task (Duncko et al., 2007) and psychological stressors such as the Paced Auditory Serial Addition Task (van Gerven et al., 2016). However, the Trier Social Stress Test is the most common stressor used in the literature and can be considered as the gold-standard stressor. The Trier Social Stress Test involves participants completing both a mock job interview and difficult cognitive task in front of an audience. In fact, the Trier Social Stress Test is ideal for both eliciting an HPA axis (and thus cortisol) response (Kirschbaum et al., 1993; Kudielka et al., 2007), and a SAM axis response

(Goodman et al., 2017). The Trier Social Stress Test involves having to prepare and give a speech in front of a panel in the context of a job interview, and then conducting a difficult cognitive task involving subtracting numbers.

The theory behind the Trier Social Stress Test is that it can combine several situations that are known to cause stress: a cognitively demanding task, public speaking in front of an audience, the anticipation of having to prepare for a speech, and uncontrollability. Thus, as per Dickerson & Kemeny's (2004) theoretical model of the psychological stress response, the Trier Social Stress Test introduces both socio-evaluative threat and uncontrollability as the panel is instructed to not give positive feedback and act unemotional towards the participant. A meta-analysis by Liu et al., (2017) showed that the Trier Social Stress Test consistently elicits sex differences in cortisol reactivity – finding the typical observations of increased response in females in the luteal phase of the menstrual cycle and decreased responses by females taking oral contraceptives. Furthermore, a review by Allen et al., (2014) found that the Trier can effectively change a variety of HPA axis hormones such as vasopressin, dehydroepiandrosterone, & adrenocorticotropic hormone (thus it is not causing a cortisol specific response) and consistently elicits SAM axis responses as well. Thus, evidence across multiple large-scale reviews suggests the Trier Social Stress Test is effective at eliciting both the fast, SAM axis response and the slow, HPA response, and can successfully replicate the expected sex effects, making it an ideal stressor. Note that for the rest of the general introduction section, my focus is on the impact of acute stress and the acute stress response rather than chronic stress. Chronic stress and the impact it has on cognition and behaviour – the focus of experiment 4 – will instead be summarized in the introduction section of experiment 4.

Stress, Feedback, and Uncertainty

The acute stress response has also been shown to impair specific aspects of feedback learning. Broadly, feedback learning involves being able to use positive and negative feedback to guide future actions (Berridge, 2000). Acute stress has been shown to reduce both reward sensitivity and response bias towards reward (Bogdan & Pizzagalli, 2006). As well, acute stress has been shown to reduce feedback learning performance (Paul et al., 2020; Preston et al., 2007). Relatedly, following acute stress individuals have reductions in learning from positive prediction errors (Carvalheiro et al., 2021). Prior work has shown that acute stress impairs how participants use negative feedback (i.e., avoid punishments) following learning compared to a control group (Lighthall et al., 2013; Petzold et al., 2010), although Lighthall and colleagues also found that acute stress improved participants' performance when using positive feedback compared to control participants. Complicating matters slightly, other work has shown that cortisol responders have decreased accuracy compared to a control group when using reward to learn but no differences in learning compared to the control group when using punishments (Berghorst et al., 2013). In sum, it seems that the effects of acute stress on feedback learning are varied, and while generally feedback learning is impaired following acute stress, it is unclear whether participants exposed to an acute stressor are better at learning to use positive feedback compared to control participants.

Feedback learning following acute stress has also been investigated using functional neuroimaging. In humans, fMRI work has shown that reward-related medial pre-frontal cortex functioning is impaired by acute stress (Ossewaarde et al., 2011). Further investigations have observed that acute stress decreases reward-related (but not punishment related) activity in the dorsal striatum and the orbito-frontal cortex which could explain impairments following acute

stress exposure in participants' ability to process and use rewarding information to guide future choices (Porcelli et al., 2012). Relatedly, Robinson, Overstreet, Charney, Vytal, & Grillon, (2013) found that stress increases striatal activity in response to punishing stimuli but not rewarding stimuli. Other work has even provided evidence of sex differences following acute stress for reward processing as males have greater activity in the insula and striatum following an acute stressor in comparison to a non-stressful task, while females have less activity in the insula and striatum – corresponding to gender differences in performance (Lighthall et al., 2012). Lastly, work has found deactivation of both the orbito-frontal cortex and midcingulate cortex during acute stress (Pruessner et al., 2008). To summarize, it seems then that the circuits underlying learning from reward and feedback are impacted by acute stress with deactivation in reward-related regions being typically observed.

When making decisions and during learning, there are times when information is missing or when there is a chance for an undesired outcome. Specifically, an uncertain decision is one where the knowledge about the relationship between choice and outcome is imperfect, in other words, when there is ambiguity about the relationships between choice and outcome (Platt & Huettel, 2008). An example of an uncertain decision might be when someone with a knee injury must decide whether to either pursue a surgery or go through protracted physical therapy. In this case, neither choice is guaranteed to fix the knee problem and while perhaps the surgery has lower success it is a more immediate solution. Relatedly, a risky decision is one in which there is some chance of a consequence or punishment from the action, but all outcomes are known (Platt & Huettel, 2008). In keeping with the surgery theme, someone might elect to have a surgery to stop a degenerative disease, but the surgery itself has a chance of complications. The relationship between uncertain and risky decisions can be unclear in the literature. On one hand, some have

considered decisions under uncertainty and risk to represent a continuum spanning from a complete lack of knowledge of an outcome of a decision, to when outcomes are known but the probabilities of the outcomes aren't (uncertainty), to cases when the outcome probabilities are known but there are multiple outcomes (risk; Starcke & Brand, 2012). On the other hand, some researchers subsume risky decisions as being a type of uncertain decision making (Hsu et al., 2005). Thus, I consider risk and uncertainty as distinct but related constructs.

Importantly for the present work, acute stress has been shown to modulate decisions under uncertainty. A recent review concluded that acute stress impacts decision making by enhancing automatic processes and reducing emphasis on control-related processes when making decisions under uncertain conditions (Starcke & Brand, 2012). For example, when making decisions under uncertain conditions, participants show worse decision-making abilities in terms of points acquired when told they would have to deliver a speech in front of an audience at some point after the task (Starcke et al., 2008). A follow-up meta-analysis by these same authors found that decisions made either following or under acute stress are only impaired when risk-taking and uncertainty are disadvantageous (Starcke & Brand, 2016). Following acute stress exposure, participants exhibit less strategic decisions under uncertainty than control participants, and this difference was correlated with higher cortisol levels (Leder et al., 2013). When instead making decisions under risk, participants in the acute stress condition show a higher preference for risky options under loss than under gain – suggestive of greater automatic than rational processing (Porcelli & Delgado, 2009). Still other work has shown that acute stress improves decision making under uncertainty (Byrne et al., 2019, 2020), which the authors suggested could be because they focused on rewarding outcomes (no losses were possible) or because the task assessed participants' ability to strategize for long-term reward over short-term gains. In sum,

these findings suggest that acute stress interacts with both uncertainty and risk when people make decisions.

Explore-Exploit Decision Making

Given the wealth of research on the effect of acute stress on both feedback and uncertainty in decision making, it would be ideal then to investigate a form of decision making that incorporates both these measures. In fact, explore-exploit decision making provides an ideal candidate as it involves both the incorporation of feedback (to determine when to exploit) and uncertainty (to determine when to explore). As well, explore-exploit decision making is an ideal candidate for the application of ERPs as a method of investigation (e.g., (Hassall et al., 2013, 2019; Hassall & Krigolson, 2020).

Before discussing the effect of stress on explore-exploit decision making, it is useful to define what is meant by the terms “explore” and “exploit”. When animals or humans are faced with multiple options, they must choose between sticking with what they know (an exploitative decision) or choosing another option (an exploratory decision) – which is known as the explore-exploit dilemma. These two decisions come with a cost-benefit analysis of whether a person or animal should explore or exploit. In many tasks assessing explore-exploit behaviour, reward probabilities vary across the task and in order to effectively manage the trade-off participants not only have to identify the preferred option but explore to keep track of values of the other options (e.g., Daw et al., 2006), that is, participants must balance the known reward of a current option (while the reward could be diminishing) with the uncertain reward of alternative options. Recent work has demonstrated that when humans explore, they typically do so to reduce the uncertainty of their choices (Gershman, 2019). Perhaps more convincingly, people explore to reduce

uncertainty outside the lab when making real world decisions, such when choosing to order food for delivery (E. Schulz et al., 2019).

When investigating the explore-exploit dilemma, typically, tasks such as a multi-armed bandit (e.g., Daw et al., 2006) or the patch foraging task (e.g., Constantino & Daw, 2015) are used to assess explore-exploit behaviour. In the multi-armed bandit, participants select one bandit (also known as a slot machine or arm) from *n* number of bandits on each trial. Following their selection, the participants are given feedback for their selection. This feedback could be in the form of points (e.g., Daw et al., 2006) or a win or a loss (Littman, 2015). The goal of the participant in this task is to maximize reward, and as such they need to determine which bandit is preferred by accurately sampling from the environment and then remaining consistent with their selection once they've found the bandit that provides the highest points or most wins. However, remaining stable comes with the implicit assumption that the bandit win-probabilities remain fixed (a stationary bandit). In fact, there are several different types of bandit environments– including "contextual bandits" in which the underlying reward of each bandit shifts across the task and "learnable bandits" where the reward probabilities remained fixed (R. S. Sutton & Barto, 2018). In the patch-foraging task, on each trial participants are faced with the choice to harvest from a patch or move on to another patch. As the participants harvest from each patch, some amount of reward is removed from the patch, meaning that participants need to take care to not overharvest each patch or they end up sacrificing reward. However, there is also an opportunity cost as travelling between patches reduces the total amount of time the participant has available to harvest, as such participants do not want to leave patches too quickly. Importantly, the reward value and harvest time for each patch is randomly determined, and as

such participants must effectively balance the explore-exploit dilemma to ensure they are not leaving patches too early (under-harvesting) or perseverating in patches (over-harvesting).

In terms of neuroanatomy, exploration trials are associated with greater activity in the frontal-polar cortex and intraparietal sulcus (Daw et al., 2006), with the suggestion that the frontal-polar cortex disengages control processes leading to exploration (Mansouri et al., 2015). More recent research has even identified different types of exploratory behaviour – known as directed and random exploration (Wilson et al., 2014). Wilson and colleagues suggest that directed exploration involves the incorporation of information of other choices such as expected feedback while random exploration involves computations of decision noise (i.e., exploring options purely at random). Further work on directed and random exploration has shown that the two exploration strategies likely rely on dissociable neural substrates (Zajkowski et al., 2017), and that different computations of uncertainty underlie these two strategies (Gershman, 2019). Specifically, directed exploration is sensitive to relative uncertainty (as it is directed to reduce uncertainty regarding specific choices) while random exploration is sensitive to total uncertainty (as it is sensitive to randomness).

Given that stress is typically suggested to lead to decreases in pre-frontal cortex function (Arnsten, 2009), it seems likely that stress might impact how people manage the explore-exploit dilemma. Moreover, two key neuromodulators involved in the explore-exploit dilemma are dopamine and norepinephrine (Aston-Jones & Cohen, 2005; J. D. Cohen, McClure, & Yu, 2007; Costa, Tran, Turchi, & Averbeck, 2014), and coincidently both these neuromodulators are affected by the stress response (van Oort et al., 2017). Specifically, decisions to explore involve norepinephrine (which reflects unexpected uncertainty) while decisions to exploit involve dopamine (which reflects feedback learning and strengthening of stimulus-action pairs). The

suggestion of a role of norepinephrine in signaling unexpected uncertainty is that norepinephrine is used as a signal to tell the brain when to disengage from a current choice or option (that is to explore) when unexpected uncertainty is detected due to changes in utility or value (Dayan & Yu, 2006; A. J. Yu & Dayan, 2005). That is, norepinephrine acts as a neural interrupt signal tied to exploration.

Recent work in humans has highlighted sub-optimal behaviour when solving the explore-exploit dilemma following an acute stressor (Lenow et al., 2017). Lenow and colleagues used a patch-foraging task (Constantino & Daw, 2015) to study decisions to explore or exploit following an acute stressor. The authors found that when the acute stress response is induced using a Cold-Pressor Task, participants tend to overexploit due to the participants appraising the value of their environment as being lower. Specifically, Lenow et al. (2017) showed that participants with a higher cortisol response exploit more often when they should instead be exploring. Thus, according to marginal value theorem (Charnov, 1976), the acute stress response caused participants to appraise the overall environment as being less rewarding and as such they choose to stick with what they know because their expected value of the entire environment was lowered.

While promising, work understanding the explore-exploit dilemma under stress requires more research. For example, specific modulations to feedback learning under stress during the explore-exploit dilemma have not been investigated. Given modulations to feedback learning following acute stress (e.g., Berghorst et al., 2013; Lighthall et al., 2013; Petzold et al., 2010) and modulations to dopamine following stress (Cabib & Puglisi-Allegra, 2012), it seems possible that there could be interactions between exploration-rate and feedback learning. In addition, Lenow and colleagues did not directly examine the impact of uncertainty in the effect of stress

on explore-exploit behaviour. Given modulations to decision uncertainty following stress (e.g., Starcke & Brand, 2012) and research highlighting a role of uncertainty in guiding exploratory behaviour (Gershman, 2019), I believe that an investigation of how acute stress and the acute stress response modulate uncertainty during exploration is an important question to resolve. Importantly, and as will be discussed shortly, there are two event-related potential components sensitive to feedback learning (the reward positivity; Holroyd & Coles, 2002) and uncertainty (the P300; Kopp et al., 2016). Relatedly, the use of a task such as the multi-arm bandit paradigm (R. S. Sutton & Barto, 2018) which allows for individual decisions to be distinctly categorized as exploratory or exploitative decisions seems like a natural extension of the work by Lenow and colleagues for understanding how stress impacts the explore-exploit dilemma.

Stress and EEG

The reward positivity, also known as the feedback related negativity (Proudfit, 2015), is a neural component tied to feedback learning which occurs following feedback. More specifically, the reward positivity is a frontal-central ERP component that occurs 250 to 350 ms post-feedback and overlaps in timeframe with the N200. However, the two components are functionally distinct; the reward positivity provides an index of reward processing and feedback learning while the N200 provides an index of conflict (Baker & Holroyd, 2011). The reward positivity is found by comparing the neural response to positive outcomes to the neural response to negative outcomes (e.g., Miltner et al., 1997). The reward positivity has been shown to provide a neural measure of a prediction error and has been tied to reinforcement learning in the brain, reflecting the activity of dopaminergic neurons following feedback (Holroyd & Coles, 2002; c.f. Schultz, Dayan, & Montague, 1997). This same line of research has tied the reward

positivity to activity in the midcingulate cortex (Miltner et al., 1997; Nieuwenhuis et al., 2004). A meta-analysis on the reward positivity found that the component is sensitive to both reward magnitude and likelihood – further supporting the claim that this component is sensitive to feedback learning and behavioural updating (Sambrook & Goslin, 2015). In terms of functional significance, the reward positivity is thought to reflect feedback learning, and tends to propagate back in time to predictive cues (Baker & Holroyd, 2009; Holroyd, Krigolson, & Lee, 2011; Ferguson et al., 2019) much like reinforcement learning theories would predict (R. S. Sutton & Barto, 2018). Thus, the reward positivity is thought to reflect dopaminergic activity used by the midcingulate cortex to guide behaviour in a reinforcement learning framework.

Generally, the reward positivity is reduced following acute stress. That is, while the vast majority of studies have found that acute stress reduces the reward positivity (Banis et al., 2014; Banis & Lorist, 2012; Burani et al., 2021; Paul et al., 2020; X. Zhang et al., 2020), some work has found stress increases the reward positivity (Glienke et al., 2015). In addition, some work has shown that the acute stress response reduces reward processing more generally, as both the neural response to win and neural response to losses were reduced following acute stress (Ethridge et al., 2020). I should note that while Glienke et al., (2015) did observe that the reward positivity was enhanced, they examined the reward positivity following learning, suggesting their work might be tied to feedback salience rather than an impact of stress on feedback learning per se. These results suggest that despite increases in dopaminergic activity during stress (Cabib & Puglisi-Allegra, 2012), acute stress negatively impacts the reward positivity and reward processing more generally. As will be discussed below, there are several important methodological differences between these studies examining the impact of acute stress on the

reward positivity, including the choice of stressor used, the task design, and the use of within-subject comparisons.

The P300 is a neural component tied to stimulus salience, uncertainty, and working memory and is one of the most well studied event-related potential components. The P300 is a positive spike in voltage that occurs 300 to 500 ms following stimulus or feedback presentation (S. Sutton et al., 1965). Perhaps the most well-known task used to measure of the P300 is the Oddball task (Donchin, 1981; Donchin & Coles, 1988). In the Oddball task, participants are exposed to two types of stimuli – a frequent stimulus that occurs more often and an oddball stimulus that occurs less-often. The P300 is much larger to the infrequent than to the frequent stimulus, and P300 amplitude is inversely related to the frequency of the infrequent stimulus (Donchin & Coles, 1988). The most prominent theory of the P300 is that the P300 is a signal which reflects context updating – in other words, participants perceive a stimulus which is then compared to past experience using working memory and if it is novel or uncommon then attentional resources are marshalled and updating of the stimulus representation occurs – generating the larger P300 voltages on trials where infrequent stimuli are presented (Donchin, 1981; Polich, 2007). Note that in the contextual updating theory, the frontal P300 is distinguished as the P3a and suggested to be more sensitive to novelty, while the posterior P300 is sensitive to context and is known as the P3b. Subsequent physiological examinations have tied the P300 to the locus-coeruleus norepinephrine system and have argued for a role of the P300 in decision making (the LC-P3 theory; Nieuwenhuis et al., 2005). Work on the P300 and learning has directly tied P300 amplitude to learning rate and prediction errors (Jepma et al., 2016), with larger P300 amplitudes tied to greater learning (Fischer & Ullsperger, 2013). In addition, the P300 has been shown to scale with evidence accumulation when making decisions (Kelly &

O'Connell, 2013) and is sensitive to trial-by-trial adjustments related to decision surprise (Mars et al., 2008).

The role of the P300 in decision making has been further elucidated, showing a distinct role in the P300 in learning and decision-making processes, including being a signal tied to uncertainty. Importantly, the P300 has been shown to scale to uncertainty in classification tasks (Kopp et al., 2016). Relatedly, there is a compelling line of evidence demonstrating that the P300 is larger to decisions to explore than decisions to exploit (Hassall et al., 2013, 2019; Hassall & Krigolson, 2020), including both prior to decisions to explore (Hassall et al., 2013) and to the feedback from a decision to explore (Hassall & Krigolson, 2020). Importantly, when exploration or exploitation trials are preceded by a larger number of identical trial types, there is a greater reduction in the amplitude of the P300 (Hassall, 2019). Hassall interpreted his results in the context of the P300 reflecting a neural interrupt signal tied to the activity of phasic norepinephrine (e.g., Nieuwenhuis et al., 2005). That is, his work is building upon the influential theory of norepinephrine suggesting that norepinephrine reflects detections of unexpected uncertainty tied to neural interruption (Dayan & Yu, 2006). Finally, given that decisions to explore are suggested to be related to the desire to reduce uncertainty (Gershman, 2019), I believe that these multiple lines of evidence demonstrate that the P300 is a signal sensitive to uncertainty in the explore-exploit dilemma.

There have only been a handful of investigations on the effect of acute stress on the P300 and they have typically found decreases in P300 amplitude. In a Go No-Go task, the acute stress response was found to reduce P300 amplitude in high-cortisol responders – which the authors took to mean stress modulated late-stage information processing (Dierolf et al., 2017). Other work has found decreases in the P300 amplitude following acute stress tied to the evaluation of

stimulus information (Shackman, Maxwell, et al., 2011) and attentional processing (Sänger et al., 2014). Work on the impact of acute stress on feedback learning has also examined the effect of acute stress on the P300, with the P300 being found to be reduced following acute stress (Banis & Lorist, 2012; Paul et al., 2020). However, none of the papers specifically examined the P300 as an uncertainty signal, nor did they consider the role of the P300 in exploration, suggesting a novel avenue of investigation. Of course, this might cause one to question whether the P300 to feedback and the P300 to exploration reflect the same cognitive process. No current overarching theory of the P300 exists which can explain all the experimental findings (Polich, 2020), which is a point I will return to in the general discussion.

Thus, both the reward positivity and P300 are impacted by acute stress. If one is interested in the neural mechanisms for how stress impacts the explore-exploit dilemma then the reward positivity and P300 are ideal targets to help explain how stress modulates the explore-exploit dilemma. Given the previously suggested role of the reward positivity in feedback learning, and a role of the P300 in uncertainty and exploration, I believe these are two important neural components to investigate.

The Present Research

Broadly, this dissertation will investigate how stress impacts the explore-exploit dilemma. Specifically, I am interested in whether stress modulates the neural signals involved in feedback learning, the neural signal involved in uncertainty, and the corresponding impact of stress on decisions to explore or exploit. Given the fact explore-exploit decision making involves both the mid-brain dopamine and the locus-coeruleus norepinephrine systems (J. D. Cohen et al., 2007), I believe that explore-exploit decision making is an ideal candidate for both revealing the

mechanisms of how stress impacts decision making and allowing for EEG to be effectively applied to the research questions this dissertation intends to investigate. In addition, relatively few research studies have considered the effect of stress on explore-exploit decision making, meaning it is a field ripe for investigation.

In my dissertation, I conducted four experiments examining the effect of stress on the explore-exploit dilemma and two common neurophysiological components tied to feedback learning (reward positivity) and uncertainty (P300). Experiments 1 through 3 examined the impact of acute stress and the acute stress response. In these first three experiments, acute stress was induced using the Trier Social Stress Test (Kirschbaum et al., 1993). One important manipulation was that in experiment 1 and 2, a contextual (i.e., non-learnable) four-armed bandit task was used with win-loss feedback (experiment 1) and point feedback (experiment 2). In experiment 1 I examined switch-stay behaviour while in experiment 2 I used a series of reinforcement learning models to classify participant choices as exploratory or exploitative. In experiment 3, a learnable two-armed bandit task was used. That is, one of the bandits was better than the other and participants had to learn which bandit it was. Again, I used a reinforcement learning model to classify trials as exploratory or exploitative. In the last project, I switched my emphasis from acute stress induced via the Trier Social Stress Test to an examination of chronic stress and anxiety. In experiment 4, I again used a learnable two-armed bandit task to examine how participants' history of chronic stress exposures and anxiety impact decisions to explore and exploit, the reward positivity, and the P300.

Chapter 2 - Experiment 1

Introduction

When someone is faced with a particularly important event – for example, giving an oral presentation – a common physiological and emotional response to that presentation is the experience of stress. Intuitively it would seem likely that conducting that presentation would change the behaviour and cognitive abilities of that person as they deal with the stressor and its aftereffects (the acute stress response). As one possibility, it could be that the person exposed to the acute stressor would be less able to integrate the feedback they get from their surroundings to guide their actions due to the acute stress response. For example, they might be unable to use negative feedback from their audience to stop talking about evidence which undermines their argument. As another possibility, it could be that the acute stress response changes how the person make decisions when faced with uncertainty. For example, the person might be willing to make a risky suggestion to end the presentation early to get some libations. Luckily for our hypothetical example above, there is evidence that stress, and specifically the acute stress response, modulates cognition by changing both how feedback information is learned and used (e.g., Petzold et al., 2010) and how decisions are made under uncertainty (e.g., Starcke & Brand, 2016).

Previous investigations have observed that acute stress modulates feedback learning. Specifically, some work has shown that the acute stress response impairs feedback learning performance (Paul et al., 2020; Preston et al., 2007). Following acute stress, individuals have been shown to be impaired at using negative feedback following learning (Petzold et al., 2010) but may be better able to learn to choose rewarding outcomes than to avoid punishments

(Lighthall et al., 2013). Moreover, the acute stress response has been shown to reduce the rate of learning of positive prediction errors (Carvalheiro et al., 2021). Neuroimaging data has corroborated these findings that the acute stress response modulates feedback learning, by showing a reduction in reward sensitivity in the brain (Porcelli et al., 2012) and a reduction in activation in reward circuits following feedback (Lincoln et al., 2019). These modulations to feedback learning following acute stress have been tied to changes caused by cortisol to the midbrain dopamine system and the midcingulate cortex under acute stress (Paul, 2019). I should note that Lighthall et al., (2013) found that the acute stress response actually enhanced using positive feedback following learning compared to a control group and Glienke et al., (2015) found no effect of the acute stress response on feedback learning performance, suggesting that acute stress might not always impair feedback learning.

Prior work has also shown that acute stress impacts the type of decisions people make under uncertainty, though no work has specifically examined how acute stress impacts uncertainty signals in the brain. When decisions are uncertain, acute stress enhances automatic decisions while impairing control-related decisions (see Starcke & Brand, 2012 for a review). As well, the acute stress response causes participants to focus on larger feedback leading to less strategic decisions under uncertainty (Leder et al., 2013), and the acute stress response reduces decision-making performance in terms of points acquired (Starcke et al., 2008). However, there are discrepancies in the literature, as some work has shown that the acute stress response might improve decision making under uncertainty – following an acute stressor participants were more likely to select the optimal option in a two-choice task (Byrne et al., 2019, 2020). Neuroimaging research has shown that the encoding of uncertainty is tied to brain regions such as the midcingulate cortex and the dorsolateral pre-frontal cortex (Critchley et al., 2001; Nassar,

McGuire, et al., 2019). Moreover, specific investigations of unexpected uncertainty (the type of uncertainty which I've argued is tied to exploration) have shown that posterior parietal regions of the brain, such as the posterior cingulate cortex, are more active for unexpected uncertainty at the time of outcome feedback (Payzan-LeNestour et al., 2013). Importantly, there is evidence that the acute stress response modulates neural activity in both the midcingulate and the dorsolateral pre-frontal cortices (Dedovic, D'Aguiar, et al., 2009) and the posterior-cingulate cortex (Qin et al., 2009). To my knowledge no research has examined how neural signals tied to uncertainty are specifically impacted by the acute stress response. Thus, given there are discrepancies in the behavioural findings of the effect of acute stress on decisions made under uncertainty, one possible solution is by examining how neural signals tied to uncertainty are impacted by acute stress.

Returning now to the question of how the acute stress response might impact both feedback learning and uncertainty, one question which arises is whether acute stress related modulations to these two constructs can concurrently change behaviour? An ideal means of investigating this question is using tasks which examine the explore-exploit dilemma. Decisions to either explore or exploit will occur when an agent is in an environment with multiple options. In this case, the agent (human, animal, or artificial) must determine whether to stick with its current choice (an exploitation) or search its environment for a less known, but potentially better choice (an exploration). Effective exploration requires the incorporation of feedback, as the agent must decide to exploit an option which gives some known amount of reward, and uncertainty, as the agent must compare its current choice to unselected options where the reward distribution is unknown but could offer more reward. That is, exploration requires knowledge of not only the feedback provided by the current option, but computations of uncertainty of the outcome of other

possible outcomes. Unsurprisingly, both feedback learning and uncertainty have been hypothesized to be key components of effectively managing the explore-exploit dilemma (J. D. Cohen et al., 2007).

In humans, there has been one prior investigation of how the acute stress response impacts the explore-exploit dilemma (Lenow et al., 2017). The authors found that acute stress modified the explore-exploit dilemma by causing participants to engage in over-exploitative strategies to the detriment of task performance. Lenow and colleagues hypothesized that this reduction in exploration was due to the participants' assessments of environmental quality being lower under stress and the expected value of the environment was lower. My suggestion is that this assessment decrease could be driven by concurrent modulations to how participants used feedback and experience uncertainty. Thus, I believe acute stress may modulate the explore-exploit dilemma by changing how participants incorporate feedback and uncertainty to guide their actions. Importantly, electroencephalography (EEG) is an ideal method to investigate this problem, as EEG has been used to show that both feedback learning and uncertainty are associated with specific neural signals in the brain. These two signals are known as the reward positivity (feedback learning; Holroyd & Coles, 2002) and the P300 (uncertainty; Kopp et al., 2016).

The reward positivity is a reinforcement learning signal which has been localized to the midcingulate cortex (Miltner et al., 1997; Nieuwenhuis et al., 2004). The reward positivity is generally taken to be the difference between positive and negative or neutral feedback around 300 ms post-feedback. The name itself comes from the fact that the presentation of positive or better than expected feedback, produces a higher positive voltage between 200 and 400 ms than negative or neutral feedback (Proudfit, 2015). Likely the most prominent theory of the reward

positivity is that it is a neural signal of reinforcement learning which is elicited in the front of the scalp when feedback or events are better than expected (Holroyd & Coles, 2002; Proudfit, 2015). Prior investigations using EEG to investigate how acute stress impacts the reward positivity have provided mixed findings. Specifically, the acute stress response has typically been shown to reduce the reward positivity in learnable tasks (Banis et al., 2014; Banis & Lorist, 2012; Paul et al., 2020) but increase the reward positivity when learning is not present (Glienke et al., 2015). Thus, it seems that the acute stress response impacts neural signals of reward by reducing feedback learning in most cases but may cause feedback to be more salient after learning has concluded.

In contrast, the P300 is a positive going spike in voltage between 300 and 500 milleseconds following stimulus or feedback presentation (S. Sutton et al., 1965). While there are a number of competing theories of the functional significance of the P300, one of the most prominent is the context-updating hypothesis, which states the P300 is elicited in events that are contextually relevent or surprising (Polich, 2007). Importantly however, a compelling series of experiments have argued that the P300 might represent a neural interupt signal when tied to decisions to explore (Hassall, 2019), in turn related to unexpected uncertainty and the activity of phasic norepinephrine (c.f. Aston-Jones & Cohen, 2005; Dayan & Yu, 2006; Nieuwenhuis et al., 2005). As mentioned in the general introduction, other research has directly argued for a role for the P300 as sensitive to uncertainty (Kopp et al., 2016) and highlighted how changes to the P300 reflect modulations to decision related surprise and uncertainty (Nassar, Bruckner, et al., 2019). Prior investigations of the P300 under acute stress have typically found that acute stress reduces the P300 in paradigms investigating attention (Dierolf et al., 2017; Sänger et al., 2014) and reduces the P300 in paradigms investigating feedback learning (Paul et al., 2020). Thus, in the

present work, I will examine how acutes stress modulates the P300 during decisions to explore or exploit.

For the present project I induced acute stress to determine how the acute stress response modulates both feedback learning and uncertainty signals. I induced. the acute stress response through the use of the Trier Social Stress Test (TSST; Kirschbaum et al., 1993). After this, participants completed a task that required effectively managing the explore-exploit dilemma – the four-armed contextual bandit (Daw et al., 2006). The four-armed bandit is an ideal task to understand decision making under uncertainty as the win percentage of the bandits change throughout the task. However, in my experiment I made two changes to the four-armed bandit task itself and the analysis protocol of Daw and colleagues. One is that I changed the points feedback (usually 1 to 100) to Win/Loss feedback. The reason for this is that I expected win-loss feedback to provide a clear reward positivity when compared to point feedback. As well, rather than use a conventional reinforcement learning model to extract explore-exploit behaviour (as Daw and colleagues did), I instead examined stay trials (i.e., a trial where the participant selected the exact same bandit) and switch trials (i.e., a trial where a participant switched to a new bandit). My rationale for this choice was that trials on which someone stays with their current option should be equivalent to an exploit trial, while a trial on which someone switches to another option should be equivalent to an explore trial.

Thus, in this experiment I investigated how the acute stress response impacted switch and stay behaviour in a four-armed bandit task. Specifically, my goal was to determine how the acute stress response modified both behaviour and the neurophysiology of participants by examining neural signals tied to feedback learning and uncertainty. My expectation was that the acute stress response would modify behaviour by causing participants to win less often, which would be

related to participants switching less often, akin to the finding by Lenow et al., (2017) that participants over-exploit to the detriment of performance. As well, I expected that the acute stress response would reduce the neural signal associated with feedback learning (the reward positivity) given prior work showing decreases in the reward positivity following an acute stressor (Banis et al., 2014; Banis & Lorist, 2012; Burani et al., 2021; Paul et al., 2020; X. Zhang et al., 2020). In a novel analysis, my expectation was also that acute stress would reduce the neural signal associated with uncertainty (the switch P300). To verify that switch and stay trials are similar behaviourally to explore and exploit trials respectively, I would expect that switch trials should be slower than stay trials and should produce less wins much like how explore trials are generally slower and produce less wins (e.g., Hassall & Krigolson, 2020). In addition, I expected the P300 to be larger for switch trials than stay trials akin to how explore trials elicit a larger P300 than exploit trials (Hassall et al., 2013, 2019; Hassall & Krigolson, 2020).

Method

Participants

54 healthy participants were recruited from the University of Victoria. Due to data quality issues, I had to exclude five participants. For three of the participants, independent component analysis could not be completed due to data quality issues. For the remaining two excluded participants, they had greater than 50% of trials as invalid (i.e., no response was made before a two second response window ended). For the final sample of 49 participants (25 females, age range 18 to 35, mean age = 20.92, 95% CI [19.94, 21.90]), I had 27 participants in the control condition and 22 participants in the acute stress condition.

As per the suggestion of Shields (2020) for conducting research on acute stress, I did not include any participants who had neuropsychological or health issues, did not test any females who were on hormonal birth control[1], I attempted to balance the ratio of males to females across all conditions, I tested no participants who smoked regularly, and I ensured that no participants had either eaten a large meal, exercised, or had smoked for at least two hours prior to the experiment. All testing occurred in the afternoon, between 12 pm and 7 pm to avoid fluctuations in cortisol caused by circadian rhythms. All participants provided written informed consent prior to the completion of the experimental session. Participants received course credit for a Psychology course for their participation. The Human Research Ethics Board at the University of Victoria approved all experimental procedures.

Power. To determine the appropriate sample size for the present work, I recruited a similar group size to previous work examining feedback-learning and acute stress (e.g., Glienke et al., 2015; Paul et al., 2020). Glienke and colleagues had 20 participants in both the acute stress and the control conditions, while Paul and colleagues had 17 participants in the acute stress condition and 20 participants in the control condition. In Paul and colleagues, the authors found a large effect of the acute stress response on the reward positivity (η^2 = .14). Thus, to detect an effect of acute stress on EEG with 80% power and assuming α = .05, I would need 26

[1] It is important to note that while I excluded females on hormonal birth control due to physiological differences in their cortisol response (Dedovic, Wadiwalla, et al., 2009) and possible interaction with cognitive effects (Nielsen et al., 2014). However, given the lack of collection of cortisol in experiment 1 and the high proportion of college females on hormonal birth control, this sample may not be entirely representative of the University population from which it was sampled.

participants in the control condition and 26 participants in the stress condition, similar to the group sizes used here.

Apparatus and Materials

All testing occurred in a soundproof room, with subjects seated 60 cm in front of a 59-cm monitor (1680 by 1050 pixels resolution). Stimulus presentation and data acquisition were controlled using MATLAB (Version 8.2, Mathworks, Natick, USA) through the Psychophysics Toolbox extension (Brainard, 1997; Pelli, 1997), on a Linux-based operating system (Ubuntu version 16.04).

Questionnaires and Heart Rate. To assess the effect of acute stress exposure, I measured participants' heart rate and their emotional response to two questionnaires (State-Trait Anxiety Inventory, and the Positive and Negative Affect Schedule). Heart rate was recorded continuously throughout the task with the use of a Polar H3 heart rate monitor attached using a chest strap (Polar Inc.) while data was streamed via Bluetooth to a mobile device. I measured participants' heart rate in two blocks, by taking the average heart rate across the entire baseline period, and their average heart rate during the stressor or placebo task. Note that four participants had heart-rate connection issues (the data stopped streaming at different points in the experiment), and thus I ended up with a total of 46 participants with proper heart-rate data. As well, I had participants complete the State-Trait Anxiety Inventory – State scale (STAI-S) and the Positive and Negative Affect Schedule (PANAS). I had participants complete these questionnaires four times throughout the experiment: (1) upon entry to the lab, during the time when the EEG cap was being put on participants (baseline period), (2) immediately following completion of the stressor or placebo task, (3) immediately before the four-armed bandit task, and (4) immediately after the four-armed bandit task.

Stress or Placebo Task. In order to induce the acute stress response in participants, I had participants either complete the Trier Social Stress Test (stress condition; Kirschbaum et al., 1993) or the placebo version (control condition; Het et al., 2009). Briefly, the Trier Social Stress Test involves participants being brough to a new room and being instructed that they would have to give a talk in front of a panel to justify why they are an ideal candidate for a job while the interview was videotaped. Participants were then given three minutes to prepare their talk after which they had five minutes to give their talk. The panel members were research assistants who the participants have never met before, were instructed to act coolly and business-like towards the participant and were instructed to provide the participant with their attention throughout the task by maintaining eye contact and writing occasional notes about the participant's behaviour. After the eight-minute interview period, the participants completed a five-minute cognitive task where they had to subtract the number 17 consecutively from a series of numbers, starting from the number 2023. Anytime the participants made a mistake, the panel members would stop the participants and have them begin again at the number 2023.

In contrast, the Placebo version of the Trier Social Stress Test (Het et al., 2009) was designed to be very similar to the stressor version but removed any feelings of socio-evaluative threat. Thus, the placebo version of the Trier Social Stress Test was like the stress version but had a few key differences. The major difference was that the participant was not instructed to prepare for a job interview – instead, they were instructed that they had three minutes to think about something like a movie or novel to talk about. They then had five minutes to talk aloud about whatever they chose. Crucially, participants were informed that no one is present in the room for either their preparation for the talk or the talk itself. After the talk period, participants

were given a cognitive task where they added the number 15 to itself consecutively. Again, no one other than the participant was in the room itself.

Multi-Armed bandit. Participants completed a four-armed bandit task (Daw et al., 2006). Unlike the Daw et al., 2006 bandit task in which participants received 1 to 100 points per bandit selection, participants received win or loss feedback. Before beginning the task, participants were given instructions of the task demands and shown an example of the bandits. Participants were told that their goal for the task was to win as much as possible by selecting the bandit that gave them the most wins. As well, participants were informed that throughout the task the winning percentage of the bandits would be slowly changing. I used the parameters from Daw et al., 2006 to randomly determine each of the four bandits' winning percentages across the task – for each participant these values were randomized (figure 1). Specifically, four noisy means (from one to 100) were chosen for each bandit drawn from a Gaussian distribution with a standard deviation (σ) of four. For each trial, the means diffused in a Gaussian random walk with the following formula:

$$\mu_{i,t+1} = \lambda\mu_{i,t} + (1-\lambda)\theta + v \qquad (2.0)$$

For equation 1.0, the decay parameter (λ) was .9836, the decay center (θ) was 50, and the diffusion noise (v) was a zero-mean Gaussian (σ = 2.8). This equation (2.0) computed the win percentage of each bandit for each trial. That is, if the bandit value was 77% then participants had a 77% chance of receiving win feedback and conversely a 23% chance of receiving loss feedback.

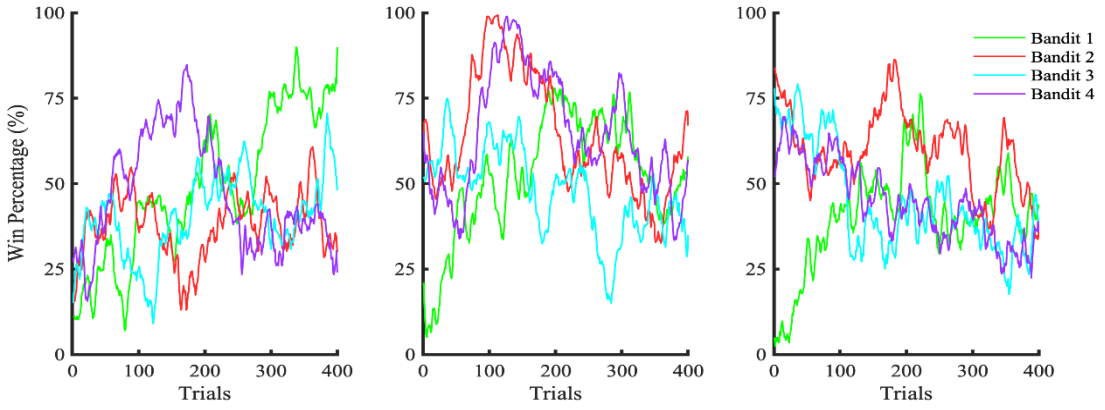

Figure 1. Experiment 1 - Sample bandit win probabilities. Win Percentages were taken from the percentages three randomly chosen participants saw.

Participants completed 400 trials, with a rest-break provided after the completion of a block of 100 trials. On each trial, participants had 2000 ms to select a bandit (figure 2). If they did not select in time, then they received a message indicating “too slow” for 1000 ms which was considered an invalid trial and no feedback was provided. On valid trials in which participants made their selection in time, their bandit choice (the coloured square) was highlighted in yellow for 400 to 600 ms. After this, the participants were provided with “Win” or “Loss” feedback for 1000 ms. The bandits were represented by different colour squares – for each participant a randomized set of four different colours were chosen – presented against a black background. To make a selection, participants used the mouse to click on a square.

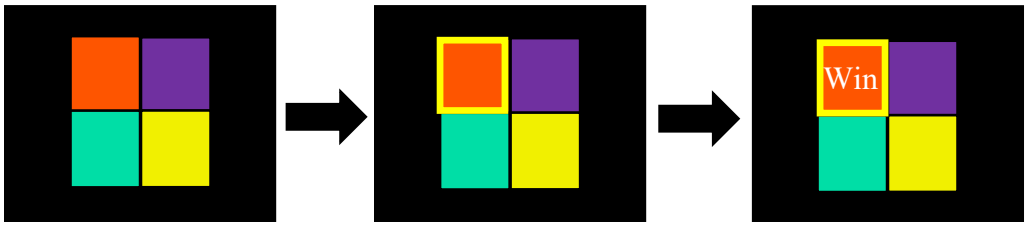

Figure 2. Experiment 1 - Four-armed bandit task. On each trial participants selected a coloured square. Their choice was highlighted in yellow, and the participants received win or loss feedback for their choice.

Procedure

The general procedure of the task is provided below (figure 3). Specifically, stress and control participants experienced an almost identical procedure whereby after they came to the lab, they were capped and filled out questionnaires for at least 30 minutes and up to an hour, which ensured an adequate baseline period (Shields, 2020). After this, the stress and control participants were either exposed to the Trier Social Stress Test, or the placebo version of the Trier Social Stress Test which took 13 to 15 minutes total to complete. Participants then completed a two-armed bandit task and an Oddball task, the order of which was counterbalanced across participants[2]. Following the two-armed bandit task and Oddball task participants completed the four-armed bandit – which occurred on average thirty minutes post-stressor onset, when the participants' HPA axis should be maximally active (Dickerson & Kemeny, 2004).

[2] The data from the Oddball and Two-armed bandit were not analyzed here but will be presented in experiment 3.

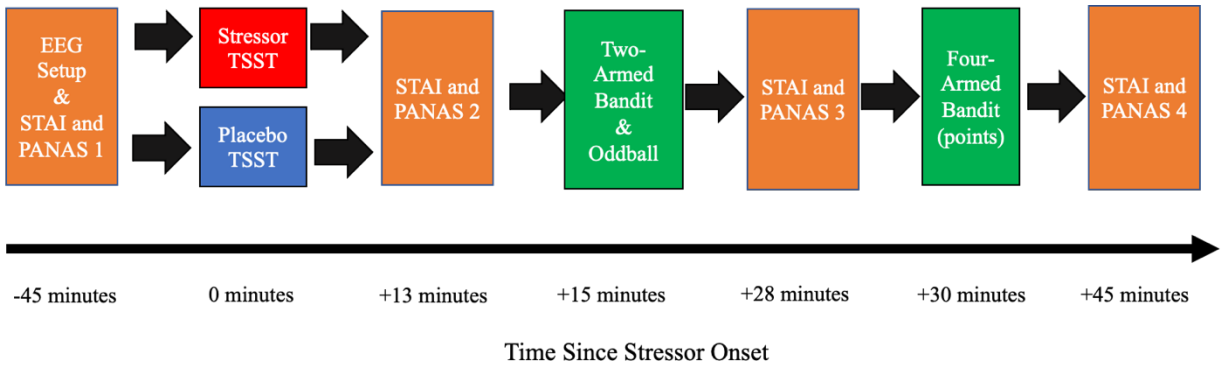

Figure 3. Experiment 1 - Procedures. The only difference between the control and stress condition was the completion of either the stressful or placebo Trier social Stress Test (TSST). STAI = State-Trait Anxiety Inventory - State scale, PANAS = Positive and Negative Affect Schedule.

Data Processing

EEG data were recorded from either 32 or 64 channels mounted in a fitted cap using a standard 10-20 layout (ActiCAP, Brain Products GmbH, Munich, Germany) and using Brain Vision Recorder software (Version 1.10, Brain Products GmbH, Munich, Germany). All electrodes were referenced to a common ground (channel AFz) and during recording electrode impedances were kept below 20 kΩ. EEG data were sampled at 500 Hz, amplified (ActiCHamp, Revision 2, Brain Products GmbH, Munich Germany), and filtered through a low-pass anti-aliasing filter of 245 Hz. Stimuli and EEG markers were temporally synched using a DataPixx stimulus synchronization unit (VPixx, Vision Science Solutions, Quebec, Canada).

All EEG data were processed in MATLAB using custom code (code can be found at www.github.com/Neuro-Tools) that depends on EEGLAB (Delorme & Makeig, 2004). EEG data

were processed using the standard methods of our laboratory (http://www.krigolsonlab.com/data-analysis.html). Data were processed twice – once to remove noisy or broken channels, and a second time to fully process the data. On the first pass, the EEG data were re-referenced to an average of the two mastoids (TP9, TP10), and channel AFz was topographically interpolated using the method of spherical splines. Then data were filtered using a dual-pass Butterworth filter with a band-pass of 0.1 to 30 Hz, and a notch-filter at 60 Hz. Following this, I used EEGLAB's independent component analysis (ICA) algorithm to identify blinks, the corresponding ICA components were plotted using topographic maps and component loadings, and components deemed to correspond to eye blinks were removed manually. Data were then reconstructed using the remaining ICA components. Data were then segmented from -200 to 800 ms related to feedback onset, baseline corrected from -200 to 0 ms, and ran through an artifact rejection algorithm with a gradient criteria of 10 μV/ms, and an absolute within-segment difference criteria of 100 μV. Any electrodes where more than 40% of trials were rejected by the algorithm were then subject to removal on the second pass. On the second pass, all steps were identical to above except before re-referencing to the Mastoids the previously identified noisy channels were removed, and after inverse ICA all previously removed channels were topographically interpolated using the method of spherical splines. The final data had an artifact rejection rate of 5.16%, 95% CI = [3.40, 6.92] of trials.

Data Analysis

Manipulation Checks. First, to verify that my stressor (the Trier Social Stress Test) led to an acute stress response, I compared the stress and the control condition across three measures. The three measures were: (1) heart rate, (2) scores on the State Trait Anxiety Inventory – State scale, and (3) scores on the Positive and Negative Affect – Negative Affect

scale. To compare conditions, I computed the average percent change from baseline across the three measures by taking the participants' heart rate or questionnaire response either during (heart rate) or post-stress (the questionnaires) and subtracting their score from the baseline period, then dividing that difference by their baseline scores and multiplying by 100. For heart rate I computed their baseline heart rate by averaging their heart rate across the entire baseline period while they were being capped but not including the stressor and computed their stressor score by taking their average heart rate during the stressor. For the two questionnaires, the baseline score was measured by having participants complete the State Trait Anxiety Inventory – State scale and the Positive and Negative Affect Schedule – Negative Affect scale while they were being capped, while the post-stressor score was measured by having them fill out the same questionnaires immediately post-stress. I compared participants in the stress and control conditions using independent samples *t*-tests, measured effect size with Cohen's *d*, and measured variability with 95% between-subject confidence intervals.

Behaviour. In order to determine how the acute stress response impacted participants in the four-armed bandit, I measured participants' time to choose a square (reaction time), the total number of wins they received divided by their total number of valid trials (win percentage), the amount of times they selected the arm with the highest win percentage divided by the total number of valid trials (optimal arm selection), how often they stayed on their current arm selection following a win divided by all win trials (win-stay percentage), and how often they switched to another arm following a loss divided by all loss trials (lose-shift percentage). I compared participants in the stress and control conditions using independent samples *t*-tests, measured effect size with Cohen's *d*, and measured variability with 95% between-subject confidence intervals.

As an exploratory analysis, I also examined optimal arm choice within three different trial blocks: trials 1 to 50, trials 51 to 100, and trials 351 to 400. In this case, I used a 3 x 2 mixed factorial ANOVA where factor 1 was time (within: trials 1 to 50, 51 to 100, 351 to 400) and factor 2 was condition (between: control, stress). I measured effect size using η_p^2, and measured variability with the mean squared error (*MSE*) term. Follow-up *t*-tests were conducted using a false-discovery rate correction (Benjamini & Hochberg, 1995) to account for multiple comparisons.

In addition, I compared participants' switch and stay behaviour as a proxy for exploration and exploitation. I compared participants' behaviour on switch and stay trials, and the effect of acute stress on these trials, by examining their reaction time on both switch and stay trials, and their win percentage – the number of win trials divided by the total number of valid trials – on both switch and stay trials. To confirm that participants' switch and stay trial behaviour differed and to determine whether acute stress influenced switch and stay trials behaviour, I compared participants on these measures using a 2 x 2 mixed ANOVA. The first factor (trial type), which was within subjects, was whether a trial was stay or switch. The second factor (condition) was between-subjects and was whether the participant was part of the control or stress condition. I measured effect size using η_p^2, and measured variability with 95% between-subject confidence intervals.

EEG. To assess the neural effects of the acute stress response, I compared both the reward positivity and the P300 in the four-armed bandit task. For both components, data were segmented from -200 to 800 ms. ERPs were then constructed on each conditional waveform (Reward Positivity: win, loss; P300: switch, stay) within each participant.

For the reward positivity, following the computation of the conditional ERPs for each participant, a difference ERP was created by subtracting the neural response to losses from the neural response to wins while collapsing across switch and stay trial type. Grand average ERPs were then computed by taking the average of the conditional waveforms (win, loss) and the difference wave (win minus loss) across all participants at electrode FCz (Krigolson, 2017; Williams et al., 2021). To avoid biasing my window choice to conditional differences (Kappenman & Luck, 2016), I examined the grand-grand average while collapsing across condition and trial type. Then I chose a window based on the peak of the grand-grand average (peak = 270 ms) which included the entire component within the reward positivity timeframe. This window ended up being a 100-millisecond window between 220 and 320 ms. I then computed each participant's reward positivity using the mean value of each participant's difference wave (win minus loss) within that window. As an exploratory analysis, I also examined the P300 to feedback (feedback P300). I used the same approach as the reward positivity again at electrode FCz. However, in this case I used a 150 ms window selected from the peak of grand-grand average (peak = 396 ms) between 320 and 470 ms, avoiding overlap with the reward positivity window. I examined the effect of acute stress and exploration on the reward positivity and the feedback P300 using a 2 x 2 mixed-factorial ANOVA where factor one was trial type (within: win, loss) while the second factor was condition (between: control, stress). I measured effect size using η_p^2, and measured variability with 95% between-subject confidence intervals. As an additional check, I then conducted a one-sample *t*-test on the mean reward positivity and mean feedback P300 for both the control condition and the stress condition, and computed Cohen's *d* and 95% between-subject confidence intervals for the difference wave for each component.

For the switch P300, I again examined the neural response to feedback presentation but rather than comparing wins and losses – I compared switch and stay trials while collapsing across feedback type. I computed the conditional ERPs for each person by averaging across all the stay trials and switch trials. Following the computation of the conditional ERPs for each participant, I computed a difference ERP by subtracting the neural response to stay trials from the neural response to switch trials. Grand average ERPs were then computed by taking the average of the conditional waveforms (stay, switch) and the difference wave across all participants. Given the paucity of research on the exploration P300, for selecting the electrode of choice I instead found the electrode where the grand-grand average value was maximum between 300 to 600 ms (electrode P4). Then I took the peak of the grand-grand average (385 ms) and computed each participant's P300 using a +/- 100 ms time window to calculate the mean of the difference wave. I examined the effect of acute stress on the switch P300 using a 2 x 2 mixed-factorial ANOVA where factor one was trial type (within: switch, stay) while the second factor was condition (between: control, stress). I measured effect size using η_p^2, and measured variability with 95% between-subject confidence intervals. Again, as an additional check I then conducted a one-sample *t*-test on the mean P300 for both the control condition and the stress condition, and computed Cohen's *d* and 95% between-subject confidence intervals for the difference wave in each condition.

For all components I calculated split-half reliability to measure internal consistency. For the reward positivity, wins had a split-half reliability of .91 while losses had a split-half reliability of .77. For the feedback P300, wins had a split-half reliability of .82 while losses had a split-half reliability of .79. For the switch P300, stays had a split-half reliability of .86 while switches had a split-half reliability of .87.

For all significance testing, an α value of .05 was used. All comparisons were checked for homogeneity of variance (α = .05) using Levene's test. In the case of a two-sample *t*-test where homogeneity was violated, Welch's *t*-test was used instead. For any mixed ANOVA where the within subjects' factor had more than two levels, Mauchly's test of sphericity was used. Greenhouse-Geisser corrections were applied when Mauchly's test was violated. All statistics were completed in R (version 4.0.3; R Core Team, 2021).

Results

Manipulation Checks

I found that acute stress exposure increased the scores of the stress condition across all three manipulation check measures when compared to the control condition for the change from baseline (figure 4), and thus did induce an acute stress response. That is, acute stress exposure ($\bar{X}_{str}$= 11.47%, 95% CI [5.54, 17.39]) increased the heart-rate change from baseline relative to controls ($\bar{X}_{con}$= -1.39%, 95% CI [-5.73, 2.95]; $t(43) = 3.73$, $p < .001$, d = 1.13). In addition, acute stress exposure increased the change from baseline on the Positive and Negative Affect Schedule ($\bar{X}_{str}$= 36.91%, 95% CI [15.31, 58.51]) compared to the control condition ($\bar{X}_{con}$= -5.47, 95% CI [-10.78, -0.15], $t_{welch}(22.45) = 3.97$, $p < .001$, d =1.25). Finally, acute stress exposure increased the change from baseline scores on the State-Trait Anxiety Inventory ($\bar{X}_{str}$= 42.11%, 95% CI [26.41, 57.80]) compared to controls ($\bar{X}_{con}$= 9.10%, 95% CI [-2.56, 15.65]; $t_{welch}(32.88) = 4.08$, $p < .001$, d =1.25).

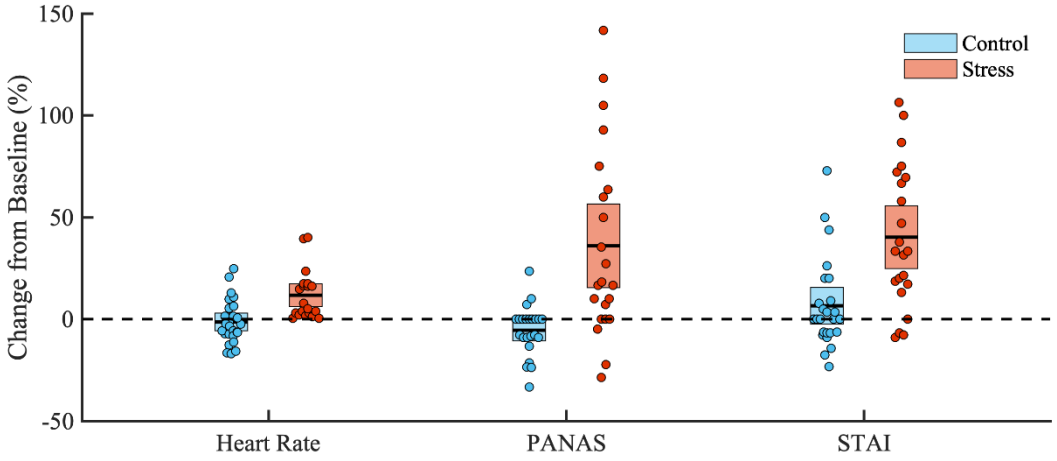

Figure 4. Experiment 1 - Manipulation checks. Effect of acute stress on heart rate, the Positive and Negative Affect Schedule – Negative Affect scale (PANAS), and State-Trait Anxiety Inventory – State scale (STAI). Error bars are 95% between-subject confidence intervals.

Behaviour

There were no effects of the acute stress response on behaviour in the four-armed bandit (Table 1). That is, there was no difference between the stress and control condition in terms of reaction-time ($t(47) = 0.65$, $p = .52$, $d = 0.19$), win percentage ($t(47) = 0.56$, $p = .57$, $d = 0.16$), optimal arm percentage ($t(47) = 0.19$, $p = .85$, $d = 0.05$), win-stay percentage: ($t(47) = 0.20$, $p = .84$, $d = 0.06$), lose-shift percentage ($t(47) = 0.20$, $p = .84$, $d = 0.06$), and switch percentage ($t(47) = 0.40$, $p = .69$, $d = 0.12$).

To determine how quickly participants' strategy stabilized in the task, I compared optimal arm choice in three blocks: trials 0 to 50, 51 to 100, and 350 to 400. However, while there was no effect of condition ($F(1, 47) = 0.35$, $p = .57$, $\eta_p^2 = .01$, $MSE = 550.93$) and no interaction ($F(1.98, 93.09) = 2.49$, $p = .09$, $\eta_p^2 = .05$, $MSE = 412.77$), there was an effect of time

($F(1.98, 93.09) = 3.61$, $p < .05$, $\eta_p^2 = .07$, $MSE = 412.77$). Both the effect of time and the interaction violated the assumption of sphericity, and as such, the Greenhouse-Geiser corrected degrees of freedom and *p*-values are reported here. A follow-up paired-samples *t*-test of the first 50 trials (0 to 50) and the last 50 trials (350 to 400), revealed a difference ($t(48) = 2.09$, $p < .05$, $d = -0.40$). A follow-up paired-samples *t*-test of the second 50 trials (51 to 100) and the last 50 trials (350 to 400), showed no difference ($t(48) = 0.23$, $p = .82$, $d = -0.05$).

Table 1

Experiment 1 – Control and Stress condition performance

	Control		Stress	
Measure	Mean	95% CI	Mean	95% CI
RT (ms)	370.14	[321. 06, 419.22]	346.65	[288.97, 404.32]
Win Percentage	57.12	[54.70, 59.52]	56.21	[54.07, 58.36]
Optimal Arm Percentage	41.36	[37.02, 45.71]	40.80	[36.44, 45.14]
Win-Stay Percentage	66.75	[58.75, 74.75]	67.95	[58.12, 77.77]
Lose-Shift Percentage	40.21	[31.97, 48.45]	38.94	[28.90, 48.99]
Switch Percentage	28.56	[18.53, 38.58]	31.30	[21.47, 41.11]

When instead comparing the stress and control conditions on switch and stay trial behaviour, there was an effect of trial type but not condition (see Table 2). For reaction time, there was an effect of trial type ($F(1, 47) = 26.67$, $p < .001$, $\eta_p^2 = .362$, $MSE = 6835.62$), but not condition ($F(1, 47) = 0.30$, $p = .59$, $\eta_p^2 = .01$, $MSE = 24734.46$), and no trial-type by condition

interaction ($F(1, 47) = 0.43$, $p = .52$, $\eta_p^2 = .01$, $MSE = 6835.62$). The same pattern was true for win percentage, as there was an effect of trial type ($F(1, 47) = 39.93$, $p < .001$, $\eta_p^2 = .46$, $MSE = 31.62$), but not condition ($F(1, 47) = 0.44$, $p = .51$, $\eta_p^2 = .01$, $MSE = 67.32$), and no trial-type by condition interaction ($F(1, 47) = 0.13$, $p = .72$, $\eta_p^2 = .00$, $MSE = 31.62$).

Table 2

Experiment 1 - Switch and stay performance

Measure		Switch		Stay	
		Mean	95% CI	Mean	95% CI
Reaction Time (ms)	Control	421.745	[370.08, 473.40]	346.02	[294.00, 398.02]
	Stress	415.17	[360.80, 469.53]	317.48	[266.72, 368.24]
Win Percentage (%)	Control	52.05	[48.55, 55.55]	58.86	[56.21, 61.50]
	Stress	50.52	[47.76, 53.29]	58.16	[55.71, 60.61]

EEG

When examining the neural response to feedback, there is tentative evidence of the acute stress response modulating the reward positivity but not the feedback P300 (figure 5). For the reward positivity, there was an effect of trial type ($F(1, 47) = 10.46$, $p < .001$, $\eta_p^2 = .18$, $MSE = 1.16$), but not condition ($F(1, 47) = 1.09$, $p = .30$, $\eta_p^2 = .02$, $MSE = 20.91$), and no trial-type by condition interaction ($F(1, 47) = 1.51$, $p = .22$, $\eta_p^2 = .03$, $MSE = 1.16$). However, when examining the difference wave of the reward positivity (win trials minus loss trials), there was a reward positivity in the control condition ($\bar{X}_{con} = 1.12\ \mu V$, 95% CI [0.44, 1.80]) which differed from zero ($t(26) = 3.39$, $p < .005$, $d = 0.65$). However, in the stress condition, there was no

reward positivity ($\bar{X}_{str}$= 0.50 μV, 95% CI [-0.29, 1.29]), as the difference wave did not differ from zero ($t(21) = 1.33$, $p = .20$, $d = 0.28$).

For the feedback P300, there was an effect of trial type ($F(1, 47) = 17.93$, $p < .001$, $\eta_p^2 = .28$, $MSE = 3.16$), but not condition ($F(1, 47) = 0.09$, $p = .77$, $\eta_p^2 = .00$, $MSE = 18.21$), and no trial-type by condition interaction ($F(1, 47) = 0.50$, $p = .49$, $\eta_p^2 = .01$, $MSE = 3.16$). When instead comparing the difference wave of the feedback P300, there was a feedback P300 in both the stress and control conditions. That is, in the control condition wins elicited a smaller P300 than losses ($\bar{X}_{con}$= -1.79 μV, 95% CI [-2.72, -0.85]; $t(26) = 3.91$, $p < .001$, $d = 0.75$). In the stress condition, the same pattern emerged, that win trials elicited a smaller P300 than losses ($\bar{X}_{str}$= -1.28 μV, 95% CI [-2.47, -0.08]; $t(21) = 2.22$, $p < .05$, $d = 0.47$).

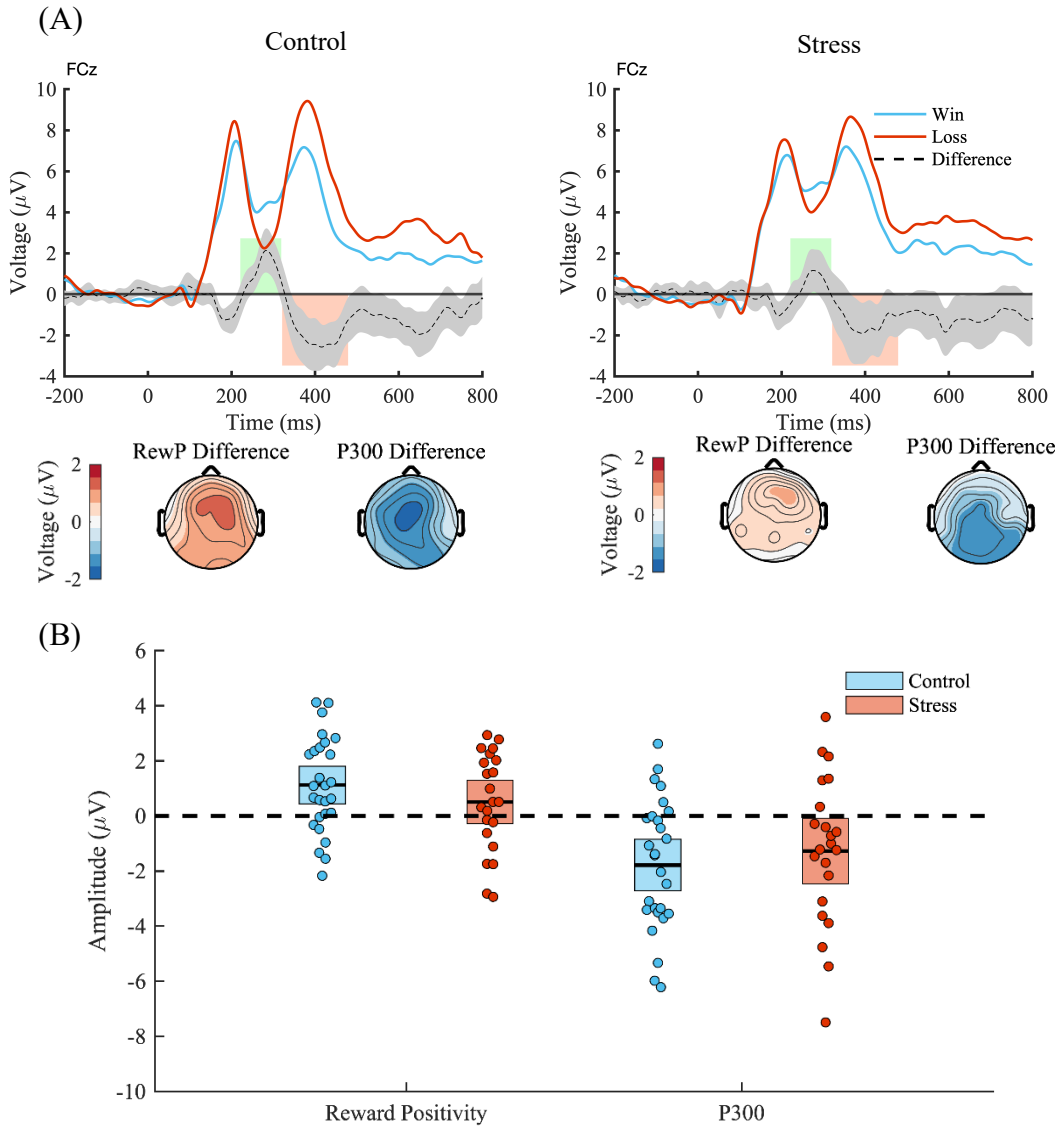

Figure 5. Experiment 1 - Feedback ERPs. (A) Waveforms for the Control condition and Stress condition. The shaded region indicates 95% between-subject confidence intervals. The green box indicates the region of analysis for the reward positivity while the pink box indicates the region of analysis for the P300. The scalp topographies indicate the difference (win minus loss) within the region of analyses. (B) Data show the average of the difference wave (win trials minus loss trials) for the region of analysis. Boxes indicate 95% between-subject confidence intervals.

When instead collapsing across feedback type and examining the neural response to decisions to stay or switch (figure 6), there is tentative evidence that the acute stress response disrupted the neural signal associated with switching. For the switch P300, there was no effect of trial type ($F(1, 47) = 3.32$, $p = .07$, $\eta_p^2 = .07$, $MSE = 4.53$), no effect of condition ($F(1, 47) = 0.17$, $p = .68$, $\eta_p^2 = .01$, $MSE = 18.21$), and no trial-type by condition interaction ($F(1, 47) = 1.73$, $p = .19$, $\eta_p^2 = .04$, $MSE = 4.53$). However, the additional check of the one sample *t*-tests revealed that in the control condition, there was a larger switch P300 in comparison to the stay P300 ($\bar{X}_{con} = 1.36\ \mu V$, 95% CI [0.19, 2.58]; $t(26) = 2.38$, $p < .05$, $d = 0.45$). However, in the stress condition, the P300 on switch and stay trials did not differ ($\bar{X}_{str} = 0.22\ \mu V$, 95% CI [-1.14, 1.58]; $t(21) = 0.33$, $p = .74$, $d = 0.07$).

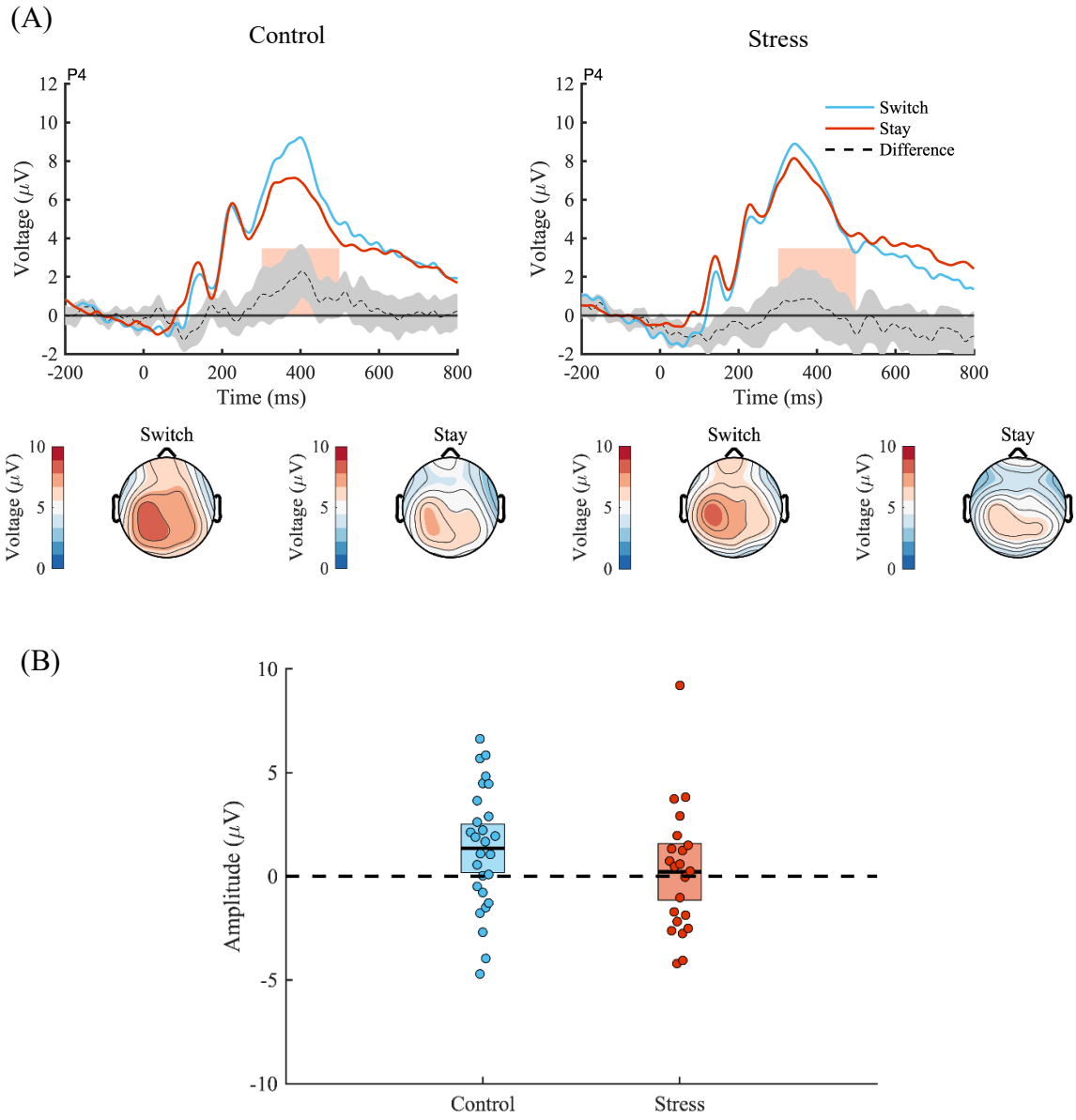

Figure 6. Experiment 1 – Switch ERPs. (A) Waveforms for the control condition and stress condition. The gray shaded region indicates 95% between-subject confidence intervals. The pink box indicates the region of analysis for the exploration P300. The scalp topographies indicate switch trials and stay trials within the pink box. *(B)* Data show the average of the difference wave (switch trials minus stay trials) for the region of analysis. Note that the boxes indicate 95% between-subject confidence intervals.

Discussion

In the present work, I found tentative evidence that the acute stress response modulated the underlying neurophysiology of participants, but no evidence for behavioural effects or changes to switch and stay behaviour. That is, I observed no differences between the stress and control condition on any of the measures of behaviour that I examined. Moreover, despite observing different behaviour on switch and stay trials (higher reaction time, less wins on switch trials) as I expected, I found no differences between the stress and control condition on any of the measures of switch and stay behaviour (and thus, I infer that there was no effect of acute stress on exploration and exploitation). I obtained suggestive evidence that acute stress disrupted the reward positivity – participants in the stress condition had no neural difference to win and loss trials while participants in the control condition did. A similar pattern was observed for the P300 to switch and stay trials. In the control condition, I found that switch trials elicited a larger P300 than stay trials (mirroring P300 effects to exploration and exploitation trials – Hassall et al., 2013, 2019; Hassall & Krigolson, 2020). In contrast, in the stress condition the switch P300 difference was absent – both switch and stay trials were identical. Thus, although I found no behavioural effects of the acute stress response on feedback learning or switch and stay behaviour (my proxy for explore and exploit trials), I did observe tentative evidence of an acute stress modulation to feedback and uncertainty signals.

Despite the lack of an impact of the acute stress response on behaviour, the suggestive evidence of a disruption of the reward positivity and the switch P300 in the stress condition is promising. The reward positivity finding suggests that the acute stress response did partly impact feedback related processing in my paradigm (although it may not have impacted how that feedback was used by participants to guide actions). Importantly, this replicates previous work

showing that the reward positivity typically decreases following an acute stressor (Banis et al., 2014; Banis & Lorist, 2012; Paul et al., 2020) – and follows from work showing that the acute stress response causes the midbrain dopamine system to malfunction (Arnsten, 2009). In addition, this extends the previous findings of acute stress reducing the reward positivity under typical probabilistic learning paradigms and extends them to decision making under uncertainty – as in the present task participants are never certain which bandit is the ideal choice due to the shifting reward probabilities. In turn, I believe that the disruption of the reward positivity by the acute stress response could contribute to sub-optimal behaviour in explore-exploit tasks under acute stress (e.g., Lenow et al., 2017). That is, participants are unable to effectively integrate feedback to guide their actions due to the acute stress response – in the present work it seems that participants may not view win-loss feedback as being better or worse than expected. I should note that there was only weak evidence (i.e., a small effect) of an effect of both acute stress and the interaction between acute stress and trial-type on the reward positivity. The small effect would suggest that the experiment was underpowered to detect an effect of the acute stress response or to detect the interaction. I relied on the use of a one-sample *t*-test to infer a disruption by the acute stress response on the reward positivity – a point I will return to later in the limitations section of the general discussion.

In addition, that the acute stress response disrupts the switch P300 mirrors neurophysiological work showing the acute stress response reduces P300 amplitude in attentional tasks (Dierolf et al., 2017; Sänger et al., 2014) and feedback learning tasks (Banis & Lorist, 2012; Paul et al., 2020). However, my argument is the switch P300 in this paradigm represents a signal sensitive to uncertainty (Kopp et al., 2016) tied to decisions to explore or exploit (Hassall et al., 2013). If this assumption is correct, the present work provides suggestive

evidence that following acute stress the uncertainty signal associated with exploration and exploitation is disrupted. Given that exploration is related to a goal of reducing uncertainty in one's environment (J. D. Cohen et al., 2007; Gershman, 2019), the present findings could suggest that the acute stress response causes participants exposed to either not correctly encoding uncertainty (as occurs during substance abuse – J. C. Yu et al., 2020) or that the acute stress response modulates their experience of uncertainty (as occurs during schizophrenia – Kreis et al., 2021). I believe that this disruption of the uncertainty signal under acute stress may be tied to the physiological activity of norepinephrine (e.g., Thoma et al., 2012), as phasic norepinephrine bursts have been suggested to be closely tied to the timing of the P300 (Nieuwenhuis et al., 2005). Moreover, some work has tied explorations to pre-frontal brain regions such as the frontal-polar cortex (Daw et al., 2006), and there is evidence that acute stress increases norepinephrine levels leading to the deactivation of the pre-frontal cortex and a switch to habit-based sub-cortical areas (Arnsten, 2015). Alternatively, unexpected-uncertainty has been tied to the posterior cingulate cortex (amongst other regions – Payzan-LeNestour et al., 2013), and there is evidence that the acute stress response and cortisol decrease activity in the posterior cingulate cortex (Qin et al., 2009). Like the reward positivity findings, it is important to note that there was only a small effect of the acute stress response and a small interaction between stress condition and trial type on the switch P300 per the ANOVA, and thus these findings should be considered as tentative.

Thus, when considering the tentative evidence showing a disruption of the feedback signal (the reward positivity) and uncertainty signal (the switch P300) due to the acute stress response, I believe that these two EEG findings could in fact be an indication of neural evidence that might support why the explore-exploit dilemma is impacted by acute stress (Lenow et al.,

2017). It is worth asking, however, how might the disruption of these two signals impact the explore-exploit dilemma? When an agent is deciding to explore, they must incorporate the current feedback they are getting and their certainty about the reward available from other options (J. D. Cohen et al., 2007). My suggestion is that the acute stress response might disrupt the agent's ability to integrate the feedback they are getting from their current choice leading to them either over-valuing or under-valuing the chosen option due to reduced reward salience (e.g., Porcelli et al., 2012) reflected in the disruption of the reward positivity. In turn, the acute stress response disrupts their computations of uncertainty when they decide to explore, leading to them under-exploring, as per Lenow and colleagues (2017).

An alternative account is that by modulating the feedback and uncertainty signals in the brain, the acute stress response causes a shift into a more habit-based mode of thought where the full history of the reward and certainty of other options are not properly updated (Schwabe & Wolf, 2011). However, it remains to be seen how to integrate work showing that the acute stress response can cause participants to make better decisions under uncertainty and, in fact, that participants may be thinking more long-term (Byrne et al., 2019). Byrne and colleagues provided one argument that their decision to only include gain feedback could possibly explain their findings. Given the use of loss feedback in the present work however, it seems that future work in this area should consider how to disentangle these discrepancies. For example, this could occur by varying the type of feedback learning available to participants (e.g., include a win-loss condition, a loss condition, and a win condition).

The lack of behavioural difference in this experiment are concerning and require further investigation. One major issue with the present paradigm is the fact that I switched point feedback (as was originally used in Daw et al., 2006) to win/loss feedback. I did this to elicit a

clear reward positivity – something which is more difficult when using points. However, because this task is a contextual bandit where reward probabilities change throughout, I believe that the win-loss feedback was not informative enough for participants to guide their behaviour. If I expect participants to compute an underlying reward value for each bandit (i.e., how much reward do you expect to receive from each selection), then it seems possible that only receiving a win or a loss might not be informative enough. In contrast, if one were to receive points that would allow a person to compare whether they've received less or more points relative to previous trials and make more informed comparisons relative to their expectations. I should note that I am not the only researcher who has modulated Daw's four-armed bandit task to use win/loss feedback (Aylward et al., 2019). However, much like in the present work, in this previous investigation by Aylward and colleagues on the impact of anxiety – no behavioural differences were found, which I hypothesize could be due to the lack of fine-grained information conveyed by win-loss feedback.

A separate possibility is that cognitive strategies in the contextual bandit stabilize relatively quickly – as evidenced by the fact that when averaging across trials 51 to 100 and comparing to trials 351 to 400, there was no difference in optimal arm choice. Thus, a competing possibility is that the contextual bandit may not be an adequate type of task to investigate behavioural differences due to a ceiling effect. In addition, a power calculation using the largest behavioural effect (reaction time, $d = .19$) revealed that to detect a behavioural effect of acute stress would require 394 people in each condition (assuming a goal of 80% power and $\alpha = .05$), suggesting the experiment may be underpowered to detect an effect. This lack of effect being yoked to the contextual nature of the bandit task will be returned to in experiment 3 where a learnable bandit task was used.

In sum, I found tentative evidence that the acute stress response may modulate feedback and uncertainty signals in the brain by disrupting both signals. Given the importance of feedback and uncertainty in exploratory behaviour (Gershman, 2019), I believe this may provide an explanation for why explore-exploit behaviour is impacted by the acute stress response (Lenow et al., 2017). In experiment 2, I hope to extend these neural findings to a task where explore-exploit behaviour is more clearly elicited, and feedback can be used more effectively to guide behaviour. Specifically, rather than use win-loss feedback, I instead used point feedback (0 to 100 – akin to the original Daw et al., 2006 investigation). As well, I used a series of reinforcement learning models to classify participants' trial-by-trial decisions as either exploratory or exploitative (as in Daw et al., 2006) rather than use the proxy measure of switch and stay behaviour.

Chapter 3: Experiment 2

Introduction

In experiment 1, I sought to investigate how the acute stress response modulated exploration and exploitation by examining how acute stress disrupted neural signals associated with feedback and uncertainty. Specifically, I hoped to build upon prior work by Lenow and colleagues who found that the acute stress response reduced exploratory behaviour by using electroencephalography to investigate two common cognitive processes associated with exploration: feedback learning and uncertainty assessments. In line with previous work (Banis & Lorist, 2012; Burani et al., 2021; Paul et al., 2020; X. Zhang et al., 2020), I found suggestive evidence that the acute stress response disrupted the reward positivity. In addition, I also found suggestive evidence that the acute stress response modulated the uncertainty signal by disrupting the P300 to switch and stay trials. In experiment 1, rather than examine exploration and exploitation trials through a computational neuroscience approach (the more common method – see Daw et al., 2006), I used switch and stay trials as proxy for exploration and exploitation. I replicated the expected effects of exploration on behaviour in switch trials, as switch trials were both slower and were associated with lower wins in comparison to stay trials (e.g., Hassall & Krigolson, 2020). However, I did not find that the acute stress response had any effect on behaviour, including in switch percentage, and thus I was unable to replicate Lenow and colleagues (2017). Thus, my goal in experiment 2 is to extend these findings from experiment 1 related to feedback and uncertainty using a slightly modified experimental design to directly elicit exploratory behaviour, rather than rely on switch behaviour as a proxy measure.

While acute stress does modulate feedback learning, different studies have found different results. Briefly, the acute stress response has been shown to impair feedback learning (e.g., Paul et al., 2020; Preston et al., 2007), and reduce reward responsiveness and reward discriminability (Bogdan & Pizzagalli, 2006). However, these findings are not always consistent. For example, Lighthall et al., (2013) showed an enhancement to positive feedback learning following acute stress compared to a control group, but an impairment for negative feedback learning. In addition, Glienke et al., (2015) showed no effect of acute stress on feedback learning. Thus, while we can say the acute stress response does impact feedback learning, it seems that different studies observe either impairments, enhancements, or no effect.

These changes to feedback learning following an acute stressor have been tied to a neural signal associated with learning, the reward positivity (Paul et al., 2020). The reward positivity is a brain signal which reflects activity of the mid-brain dopaminergic system (Holroyd & Coles, 2002) and the midcingulate cortex specifically (Miltner et al., 1997). The acute stress response has been shown to reduce the amplitude reward positivity in learnable tasks (Banis & Lorist, 2012; Burani et al., 2021; Paul et al., 2020; X. Zhang et al., 2020) but enhance the amplitude of the reward positivity when feedback is presented following learning (Glienke et al., 2015).

I should note however, that there are other components involved in the processing of feedback in addition to the reward positivity. Specifically, there is also a feedback P300 signal. While initially it was argued that the feedback P300 was sensitive to outcome magnitude rather than outcome valence (Yeung & Sanfey, 2004), more recent work has shown that feedback P300 amplitude is sensitive to valence (Hajcak et al., 2007) and expectancy (Hajcak et al., 2005). Building upon these findings, the feedback P300 has been argued to reflect more top-down feedback evaluation processes (Wu & Zhou, 2009) rather than the simple better or worse than

expected (i.e., a reinforcement learning signal) conveyed by the reward positivity. The findings by Wu & Zhou (2009) are supported by work demonstrating that the feedback P300 carries addition information related to feedback evaluation including magnitude, valence, and expectancy, which has been suggested to be related to motivation (San Martín, 2012; Schuermann et al., 2012). That is, the feedback P300 is enhanced when motivation is higher to obtain a desired outcome – such as when making decisions under risk (Schuermann et al., 2012) – and reflects top-down modulation of feedback signals.

The acute stress response has also been shown to modulate decisions made under uncertainty. When decisions are uncertain, acute stress has been shown to enhance automatic decisions while impairing control-related decisions (Starcke & Brand, 2012). Following an acute stressor, participants also tend to focus on larger feedback values, in turn, causing impaired decision making under uncertainty (Leder et al., 2013). In addition, there is evidence that acute stress causes reductions in decision-making performance under uncertainty (Starcke et al., 2008). A meta-analysis by Starke and Brande (2016) concluded that under conditions of decision uncertainty, acute stress only impairs decision making when there is a penalty for reward seeking or for risk taking. There is some literature which runs counter to this as under uncertainty the acute stress response causes participants to be more likely to select the optimal option in a two-choice task (Byrne et al., 2019, 2020). The authors hypothesized that their findings could be either because they only examined the gain-context only or because they examined a task which required maximizing future reward over short-term gains.

However, despite these findings, it was previously unclear how the acute stress response might impact neural signals tied to uncertainty. I argued previously that the exploration P300 is a neural component sensitive to uncertainty (e.g., Kopp et al., 2016) related to decisions to explore

(Hassall & Krigolson, 2020). That is, prior work has shown that a parietal (non-feedback) P300 is related to uncertainty, as the P300 was larger in amplitude to stimuli in uncertain conditions rather than certain conditions (Kopp et al., 2016). However, another theory by Cameron Hassall is that the P300 might be tied to decisions to explore as a neural interrupt signal (2019) which could be related to unexpected uncertainty (Dayan & Yu, 2006). While no prior research had investigated the exploration P300 to decisions to explore following an acute stressor, in my previous investigation (experiment 1) I found tentative evidence that the acute stress response disrupted the P300 during decisions to switch, which I believe was tied to impairments in how participants experience or encode uncertainty.

To categorize when a participant makes an exploratory decision, I have adopted an approach from computational modelling (Daw et al., 2006). Daw and colleagues categorized participants' trials as either exploratory or exploitative by modelling each participant's behaviour with a reinforcement learning model. Then by comparing the participants' behaviour to the model itself, they were able to categorize decisions when the participant chose the highest value option (as per the model) as an exploitative decision, while decisions when the participant did not choose the highest value as an exploratory decision. In this experiment, I used a similar approach as I categorized participants' behaviour on a trial-by-trial basis using multiple reinforcement learning models (R. S. Sutton & Barto, 2018). However, this does raise the question of why one might use a computational model to classify trials as explorations rather than rely on switches (as was done in experiment 1)? The main reason is that the decision to explore is not just a decision to switch but relies on the computation of value from the environment, as exploitations are decisions where the highest value option is chosen (Daw et al., 2006). That is, when a person is faced with multiple options which they have some knowledge

about, an exploration is specifically tied to decisions which are not related to value. By using a model to simulate the value of each option for each participant, one can determine when a participant chooses the highest value option (even if it isn't their current option due to changes in feedback and expectations) or when they choose the non-highest value option. This approach allows for trials where the participants do not select the highest value option to be defined as explorations. When only examining switches, value computations are not specifically accounted for. I should note however that I expect that switch trials and exploration trials would have a high degree of overlap.

Experiment 2 is a further investigation of the neural findings I observed in experiment 1 (and hence the effects of the acute stress response) but with a modification to the experimental paradigm. As described above, the win-loss feedback that was originally provided in the four-armed bandit task was instead replaced by point feedback (0 to 100). I suggested in experiment 1 that it is possible that I did not find any behavioural effects in experiment 1 because the win-loss feedback did not provide enough information to participants to successfully elicit behavioural differences. My hope is that the use of point feedback rather than win-loss feedback might provide more information to participants to elicit different behavioural effects. I hoped that this experimental change would allow me to extend my suggestive findings that acute stress disrupts both the exploration P300 and reward positivity, and to also provide a direct link between explore-exploit behaviour and feedback learning to the neural changes I found previously.

Thus, my goal with experiment 2 was to determine whether I could replicate the behavioural effects of Lenow and colleagues, and whether I could extend the neural effects of experiment 1 to direct examinations of exploration in a paradigm that used a different type of feedback (point feedback rather than win-loss feedback). In terms of computational modelling,

my primary goal was to develop and compare three reinforcement learning models to determine if they could adequately model behavioural effects and determine which trials participants explored and which trials they exploited. I expected that the models would be able to simulate participants' performance both in terms of average performance and in terms of trial-by-trial performance curves. Behaviourally, my goal was to determine if the acute stress response would modify the exploration rate of participants and the number of points participants acquired in a task where they made decisions under uncertainty (the contextual bandit). Based on prior work, I expected that the acute stress response would both reduce the number of points acquired in the four-armed bandit as in other feedback learning tasks (e.g., Paul et al., 2020), and that it would reduce exploratory behaviour (Lenow et al., 2017). In addition, a second goal was to determine whether the acute stress response would disrupt the frontal P300 to feedback (when collapsing across exploration and exploitation) and the parietal P300 to decisions to explore (when collapsing across feedback type). Based on experiment 1, my expectation was that the acute stress response would reduce both these neural components, in turn providing neural correlates for the modification of behaviour.

Method

There was a large overlap in the method used in the present work with the method used in experiment 1. As such only areas where the method differed will be mentioned.

Participants

Fifty-seven healthy participants (32 females, age range 17 to 41, mean age = 21.77, 95% CI [20.37, 23.17]; see Table 3 for a breakdown) recruited from the University of Victoria. No participants had to be removed due to EEG data quality or task performance issues. Participants

either received course credit for a Psychology course for their participation or were compensated at a rate of 15.00$ Canadian per hour. The Human Research Ethics Board at the University of Victoria approved all experimental procedures.

Power. To determine the appropriate sample size for experiment 2, I recruited a similar group size to previous work (e.g., Glienke et al., 2015; Paul et al., 2020) and experiment 1. However, it is important to note that if instead relying on the effect size of the interaction observed in experiment 1 for the reward positivity (η_p^2 = .03), then a sample size of 123 participants in each condition (246 total) would be needed to detect an effect with 80% power and where α = .05. In terms of a behavioural effect of acute stress, I adopted a liberal approach and used the largest effect size between stress and control conditions in the previous experiment (reaction time, d = .19). Thus, assuming a small effect size, I found I would need 393 people in each condition for the two groups to detect an effect with 80% power and where α = .05. In turn, these two power calculations suggest that the experiment is under-powered to detect small or medium effects (a point I will return to later in the general discussion).

Table 3

Experiment 2 – Condition breakdown

	Condition	
Sex	Stress	Control
Male	10	15
Female	17	15

Apparatus and Materials

Questionnaires and Heart Rate. Two participants had heart rate monitor connection issues as the data stopped streaming during the experiment, and thus I ended up with a total of 55 participants with usable heart-rate data. In terms of the State Trait Anxiety Inventory - State scale and Positive and Negative Affect Schedule – Negative Affect scale, one participant had incomplete data for both surveys, and thus I ended up with a total of 56 participants with usable questionnaire data.

Task

The only difference in the four-armed bandit task used in experiment 2 was that point feedback was provided rather than the win-loss feedback provided in experiment 1. Participants completed 400 trials of the four-armed bandit with point feedback (Daw et al., 2006; figure 7).

Figure 7. Experiment 2 – Four-armed bandit. On each trial participants selected from one of four options, their choice was highlighted, and then participants received point feedback for their selection.

Protocol

The task protocol for experiment 2 was identical to experiment 1, except for two major differences. The first major difference was that the two-armed bandit task and visual Oddball task were completed after the four-arm bandit (figure 8). The second major difference was that two cortisol samples were taken[3]. The first cortisol sample was taken at the conclusion of the baseline period, immediately before the acute stressor or control task. The second cortisol sample was taken immediately before the four-armed bandit task, thirty minutes post-stressor onset when cortisol receptors should be fully saturated (Dickerson & Kemeny, 2004).

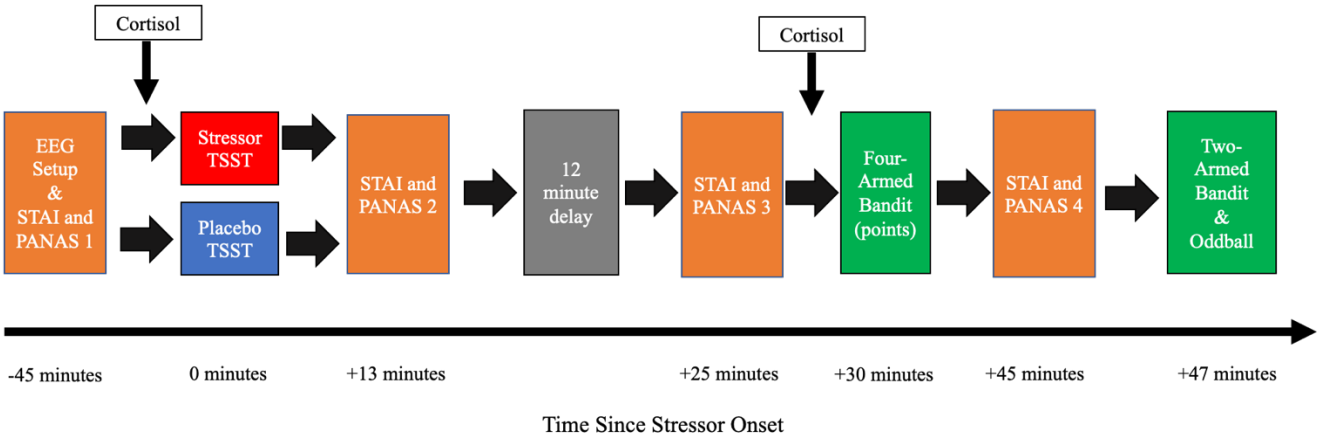

Figure 8. Experiment 2 – Procedures. The only major difference between the two conditions was that participants in the stress condition underwent a stressor TSST, while participants in the control condition underwent a placebo TSST. Note: STAI = State-Trait Anxiety Inventory State scale, PANAS = Positive and Negative Affect Schedule, and TSST = Trier Social Stress Test.

[3] There were delays in the analysis of cortisol (which was contracted outside of the laboratory) due to the COVID-19 pandemic. As such the cortisol data are not analyzed here but will be reported in any published material

Computational Modeling

To determine when participants decided to explore or exploit, I modelled participants' behaviour using four different models. The four models I used were the chance model (to provide a baseline model to compare the other models to; Wilson and Collins, 2019), an ε-greedy model, a Softmax model, and a Win-Stay Lose-Shift model. The bandit values of each model were updated using a State-Action-Reward-State-Action (SARSA) algorithm – that is, all 4 bandits have associated values but only the chosen bandit has its value updated on each trial. I used this approach for all models except the Win-Stay Lose-Shift model – which does not require value updating. In my approach all values (v_t) were initialized to 0.5 and the values for each stimulus (s) was updated using the following formula at time *t*:

$$v_{t+1}(s) = v_t(s) + \alpha * \delta_t \qquad (3.0)$$

With α being the learning rate and δ_t being my prediction error with the following formula:

$$\delta_t = r_t - v_t(s) \qquad (3.1)$$

In this case, r_t is the reward value for the stimulus and was equal to the point value of the selected arm, divided by 100 (60 points would thus be .60). Finally, as an additional constraint, v_t was never allowed to go lower than -1 or larger than 1. All modelling was completed in MATLAB (Version 9.7, Mathworks, Natick, USA).

Action Selection.

Chance Model. For the chance model, action selection was simply based on chance. That is, on each trial (*t*) the probability of making an action (a) to select one of the four stimuli (*i*) was 1/4th.

$$P_t(a_i) = \frac{1}{4} \qquad (3.2)$$

The value of the stimulus chosen was then updated as per the SARSA algorithm Above.

ε-greedy Model. For the ε-greedy model, the model utilized a near-greedy approach to select the action. In this case, the action values for each bandit (v_t) are updated on every trial. To select an action (a_i), the ε-greedy model will typically choose the bandit with the highest value but on a sub-set of trials the model will choose one of the lower values as per the exploration rate parameter (ε). To put it another way, on each trial (*t*) the probability of selecting a stimulus *i* was given by the following formula:

$$P_t(a_i) = \begin{cases} 1 - \varepsilon \ \ if \ argmax \ v_t(i) \\ \frac{\varepsilon}{3} \ otherwise \end{cases} \qquad (3.3)$$

Thus, a larger value for the exploration rate mean that the model explored more, while a smaller value means that the model explored less and choose the highest valued stimulus more often. Much like the Chance model, the value of the stimulus chosen was then updated as per the SARSA algorithm.

Softmax Model. For the Softmax model, action selection was determined using a Softmax action selection. Unlike the ε-greedy model, the Softmax model typically chooses the highest value stimulus although it explores the other stimuli as per the temperature parameter (β; Daw et al., 2006). Thus, on each trial (t) the probability of selecting stimulus i is divided by the sum of all possible actions (j) and is given by the Softmax formula:

$$P_t(a_i) = \frac{exp*(\beta*v(t_i))}{\sum exp*(\beta*v(t_{ij}))} \qquad (3.4)$$

Thus, β controls the amount of exploration and exploitation. Much like the exploration rate parameter of the ε-greedy model, a larger value of β means that the model explored more while a

smaller value means that the model explored less. The value of the stimulus chosen was then updated as per the SARSA algorithm.

Win-Stay Lose Shift Model. The last model I used was the Win-Stay Lose-Shift model. In this case, the model only depends on the previous trials feedback – the long-run values of each of the stimuli are not considered (in contrast to the previous models). The selection of a stimulus uses the following simple rules: (1) if the reward (r_t) given by the stimulus on the trial was a win then the same action is selected with the probability $P(stay \mid win)$ & (2) if the reward given by the stimulus on the trial was a loss then the action was avoided with the probability $P(shift \mid loss)$. For the Win-Stay Lose-Shift model, two parameters were computed – the probability of a win-stay ($P(stay \mid win)$) and the probability of a lose-shift ($P(shift \mid loss)$). This is because the probabilities of the other actions (win-shift and lose-stay) are simply the opposite probabilities of win-stay and lose-shift respectively. A win trial was defined as a trial where participants had more points than the previous selection while a loss was defined as a trial where participants received less points than the previous selection.

Parameter Optimization. All parameters were optimized using the optimization toolbox provided by MATLAB (Version 9.7, Mathworks, Natick, USA). Specifically, all model parameters were optimized by the function *fmincon,* which is a non-linear optimization algorithm that uses minimization and allows for boundaries to be placed on the parameters. This is particularly important for the learning rates of the ε-greedy and Softmax models and the exploration parameter of the ε-greedy model as these three parameters need to be bound between 0 and 1. All models had their parameters optimized on a participant-by-participant basis – that is, all participants had their parameters individually optimized based on their behaviour in the four-armed bandit task.

In order to avoid biases due to the choice of starting parameters (Daw, 2011; Lewandowsky & Farrell, 2018), optimization occurred through the use of a randomized parameter space to determine the starting parameters. Specifically, I drew 100 random samples from a distribution – dependent on the parameter chosen, and then I ran *fmincon* on each of these 100 different starting parameters (Table 4). For the learning rates across the three models which used them (Chance, ε-greedy, Softmax), I used a Beta distribution with an α_{beta} shape parameter of 1.1 and a β_{beta} shape parameter of 1.1 (Palminteri et al., 2015). For the exploration parameter of the ε-greedy model, the Win-Stay parameter, and the Lose-Shift parameters, I used a Beta distribution with the same starting parameters – *Beta* (α_{beta} = 1.1, β_{beta} = 1.1). For the temperature parameter of the Softmax model, I used a Gamma distribution with a shape parameter (k_{gam}) of 1.2, and a scale parameter (θ_{gam}) of 5 (Palminteri et al., 2015).

Table 4

Experiment 2 - Starting distributions for parameter optimization

Model	Mode Parameter	Distribution	Distribution Parameters
Chance	Learning Rate (α)	Beta	α_{beta} = 1.1, β_{beta} = 1.1
ε-greedy	Learning Rate (α)	Beta	α_{beta} = 1.1, β_{beta} = 1.1
	Epsilon (ε)	Beta	α_{beta} = 1.1, β_{beta} = 1.1
Softmax	Learning Rate (α)	Beta	α_{beta} = 1.1, β_{beta} = 1.1
	Temperature (β)	Gamma	k_{gam} =1.2, θ_{gam} = 5
Win-Stay	Win-Stay Probability ($P(stay \mid win)$)	Beta	α_{beta} = 1.1, β_{beta} = 1.1
Lose-Shift	Lose-Shift Probability ($P(shift \mid loss)$)	Beta	α_{beta} = 1.1, β_{beta} = 1.1

For all models, I used *a posteriori* estimation and optimized the values based on the minimization of the negative log-likelihood (Daw, 2011). However, depending on the model the calculation of the negative log-likelihood differed. For the first three models (Chance model, ε-greedy model, Softmax model), the negative log likelihood (L) was calculated by using the bandit probability (P_t) following action selection (a_i) on each trial using the following formula:

$$L = -log\ (P_t(a_i)) \tag{3.5}$$

Following the completion of each run for each of the models, the sum of the negative log likelihood was taken.

However, in the case of the Win-Stay Lose-Shift model, the log-likelihood was calculated in a slightly different manner. That is, the probabilities of the log-likelihood for each of the four trial types (Win-Stay, Win-Shift, Lose-Shift, Lose-Stay) was calculated using the following formula:

$$L_{win-stay} = log(P\ (stay \mid win))$$

$$L_{win-shift} = log\ (\frac{1-(P\ (shift\ |win)}{3})$$

$$L_{lose-shift} = log(P\ (shift \mid loss))$$

$$L_{lose-stay} = log\ (\frac{1-(P\ (stay \mid loss)}{3}) \tag{3.6}$$

Parameter Recovery. To ensure that the models' parameters are adequately identifiable and can be used to describe behaviour in an ideal circumstance (Wilson & Collins, 2019), I conducted parameter recovery across each of the parameters of the models. To do this, I first randomly simulated parameter values for a participant and for each individual model using the distributions specified in table 2. Then I had the simulated participant "complete" the experiment

using the task parameters specified above and extracted the simulated participant's performance – that is, the point values obtained and the bandits chosen. Following this, I optimized each model on a participant-by-participant basis by using the parameter optimization method laid out above which produced a series of fitted parameter values for each simulated individual within each model. These fitted parameter values were then correlated with the corresponding simulated values which were used to generate the simulated behaviour. I conducted this simulation across each of the models using 60 simulated participants as 60 was close to the final sample size of the experiment.

Data Processing

EEG data were processed identically to experiment 1. The final data had an artifact rejection rate of 11.89%, 95% CI = [10.06, 13.72] of trials.

Data Analysis

Modelling.

Model Comparison. All models were compared using Bayesian Information Criterion (BIC) to determine if model complexity and data points contribute to the model fit, as BIC more heavily penalizes models based on the complexity and number of data points (Lewandowsky & Farrell, 2018). I used the following formula for BIC:

$$BIC = -2L + KlnN \qquad (3.7)$$

Where L was the negative log-likelihood of the model, K was the number of parameters of the model, and N was the number of datapoints that the likelihood was calculated from. BIC values

were computed across all participants and models[4]. I also compared each model's BIC relative to the baseline model. To do this, I took the BIC for each participant from the chance (baseline) model and subtracted the BIC for each participant from each of the other three models (ε-greedy; Softmax; Win-Stay Lose-Shift). Larger BIC difference values (chance minus model) indicate the model is a better fit.

Exploration-Exploitation. For this experiment, I used three of the four models to classify trials as either an exploration or exploitation. As the chance model was considered a baseline model, I did not examine its ability to classify exploration and exploitation trials. For the ε-greedy model, an exploitation trial was any trial in which the participant selects the greedy (highest value) choice while an explore trial was one in which they did not select the greedy choice. For the Softmax model, I used the same logic as the ε-greedy model – that is, trials on which the participant selected the highest value choice (from the Softmax probability) are classified as exploitations – all other trials are explorations. For the Win-Stay Lose-Shift model, trials were classified as an exploitation if the participant completed a Win-Stay (selecting the same bandit after gaining more points than the previous trial) or a Lose-Shift (selecting a different bandit after gaining less points than the previous trial). Trials were classified as explorations if either of the opposing strategies are observed (Win-Shift and Lose-Stay). Lastly, to ensure model agreement, I measured the proportion of trials on which each model classified a

[4] I also extracted Akaike Information Criterion (Akaike, 1973; Wagenmakers & Farrell, 2004) values for each participant and model. The AIC values demonstrated the exact same trend as the BIC values, and as such, are not discussed further nor are they reported here.

trial as an exploration or an exploitation and then compared whether the three models' overlap in their classification of each trial across all participants.

Model's Ability to Fit Behaviour. As a method of determining the effectiveness of each model**,** I compared the ability of each model to correctly replicate the behavioural patterns of the participants through two measures of performance – points acquired and win-stay/lose-shift probabilities. To do this, I took the optimized parameters of the models for each participant and ran the model on the win probabilities of the four bandits that the participants saw. In other words, I ran the model through the actual bandit win probabilities that the participants experienced when they were completing the task, and each model selected a bandit and received feedback using the optimized parameters of that participant. For the first measure of performance, I compared the model to behaviour by examining the points the models acquired and the points the participant acquired. For the second measure of performance, I compared the model to behaviour by computing a win-stay percentage and lose-shift percentage that updated on each trial. For win-stay percentage, I took the number of times the participants or models stayed following win – defined as anytime their selection earned more points than the previous trial – divided by the number of wins. For the lose-shift percentage, I calculated the number of times the participants or models shifted following a loss – defined as anytime the selection earned less points than the previous trial – divided by the number of losses.

To determine how well the models were able to fit behaviour, I ran each model through the task on the participants' optimized parameters and computed the average points acquired. I then ran a 5 x 2 mixed ANOVA on the average points for both the behavioural data and the model data. Specifically, the first factor was model (within: behavioural, chance, ε-greedy, Softmax, Win-Stay Lose-Shift) while the second factor was condition (between: stress, control).

I measured effect size using η_p^2, and measured variability with the mean squared error (*MSE*) term. Follow-up paired sample t-tests were conducted using a false-discovery rate correction (Benjamini & Hochberg, 1995) to account for multiple comparisons.

Behaviour. I used an identical data analysis approach for the behavioural data as per experiment 1 with two differences. The first difference was that I examined exploration rate rather than switch percentage. Exploration rate was calculated as the number of explore trials – as classified by the model – divided by sum of the total number of exploration and exploitation trials. As well, I compared participants' total points acquired. To do this, I compared participants in the stress and control conditions using independent samples *t*-tests, measured effect size with Cohen's *d*, and measured variability with 95% between-subject confidence intervals. The second difference was that I compared participants on the average number of points they received (points acquired) rather than win percentage by using a 2 x 2 mixed-factorial ANOVA. In the ANOVA, factor one was trial type (within: explore, exploit) while factor two was condition (between: control, stress). I measured effect size using η_p^2, and measured variability with 95% between-subject confidence intervals.

EEG. To assess the neural effects of the acute stress response, I examined the P300 to feedback in the four-armed bandit task. Note that I did not observe a reward positivity because both win and loss and waveforms were completely overlapped within the timeframe of the reward positivity. To define a win, I took any trial where the participant acquired more points than the trial previously while to define a loss, I took any trial where the participant acquired less points than the trial previously. I examined the feedback P300 difference wave (win minus loss) at electrode Fz and chose a window based on the peak of the grand-grand average (Kappenman

& Luck, 2016; peak = 402 ms) which included the entire component within the feedback P300 timeframe. This window ended up being a 200-millisecond window between 302 and 502 ms.

To compute the exploration P300, I collapsed across feedback type and instead compared explore and exploit trials. Given the dearth of research on the exploration P300, for selecting the electrode of choice I instead found the electrode where the grand-grand average value was maximum within a 400 to 600 ms time window (electrode Pz). Then I took the time of the peak of the grand-grand average at Pz (500 ms) and computed each participant's P300 using a +/- 100 ms time window to calculate the mean of the difference wave.

For all components I calculated split-half reliability to measure internal consistency. For the feedback P300 wins had a split-half reliability of .83 while losses had a split-half reliability of .84. For the exploration P300, exploration trials had a split-half reliability of .70 while exploitation trials had a split-half reliability of .82. All statistics were completed in R (version 4.0.3; R Core Team, 2021).

Results

Manipulation Checks

As with experiment 1, I found that acute stress modulated all three manipulation check measures (figure 9). That is, the acute stressor ($\bar{X}_{str}$= 19.46%, 95% CI [13.74, 25.19]) increased the heart rate change from baseline relative to controls ($\bar{X}_{con}$= 1.26%, 95% CI [-0.99, 3.52]; t_{welch} (31.54) = 6.10, p < .001, d = 1.76). In addition, the acute stressor increased scores on the Positive and Negative affect schedule in the stress condition ($\bar{X}_{str}$= 37.62%, 95% CI [25.20, 50.06]) when compared to the control condition ($\bar{X}_{con}$= 8.40%, 95% CI [-4.35, 21.15]; t(54) = 3.36, p < .005, d =0.90). Finally, the acute stressor increased scores on the State-Trait anxiety

inventory in the stress condition ($\bar{X}_{str}$= 62.68%, 95% CI [47.38, 77.99]) compared to the control condition ($\bar{X}_{con}$= -13.87, 95% CI [0.88, 26.84]; $t(54) = 5.01$, $p < .001$, $d = 1.34$).

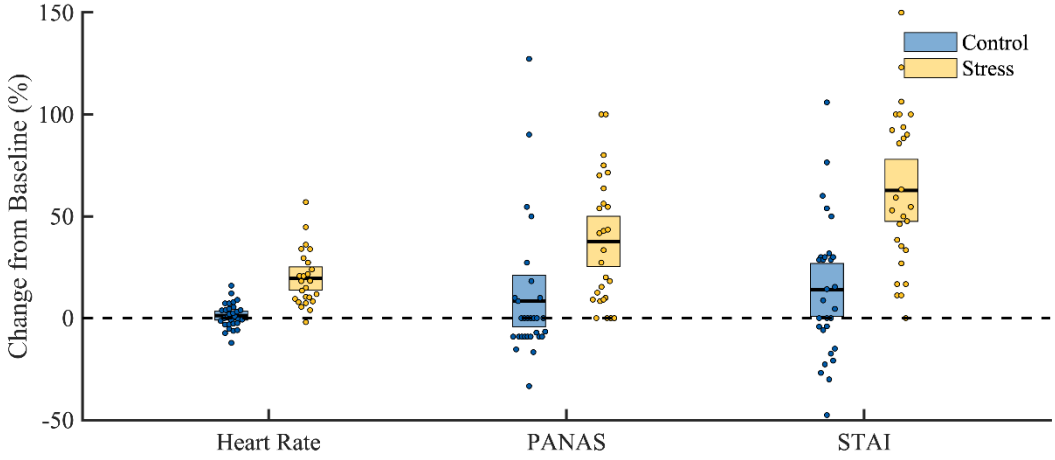

Figure 9. Experiment 2 - Manipulation checks. Effect of the acute stressor on heart rate, the Positive and Negative Affect Schedule – Negative Affect scale (PANAS), and State-Trait Anxiety Inventory State (STAI-S). Error bars are 95% between-subject confidence intervals.

Modelling

When comparing model fit to model simulation, it appears that all model parameters can be recovered effectively (all $r > .67$; figure 10). That is, when fitting each model to simulated data, there is generally a high correlation between model fit and model simulation. The only parameter where there is concern is the learning rate of the ε-greedy model ($r = .67$), as it has a much lower correlation value than the other model parameters, which all have correlation values greater than .93. Attempts to better recover the ε-greedy model learning rate, such as including positive and negative learning rates (Carvalheiro et al., 2021) and simulating the learning rate on

a truncated range of values (Wilson & Collins, 2019) reduced the correlation rather than increased it. As such, I chose to use with the single learning rate for the ε-greedy model presented here.

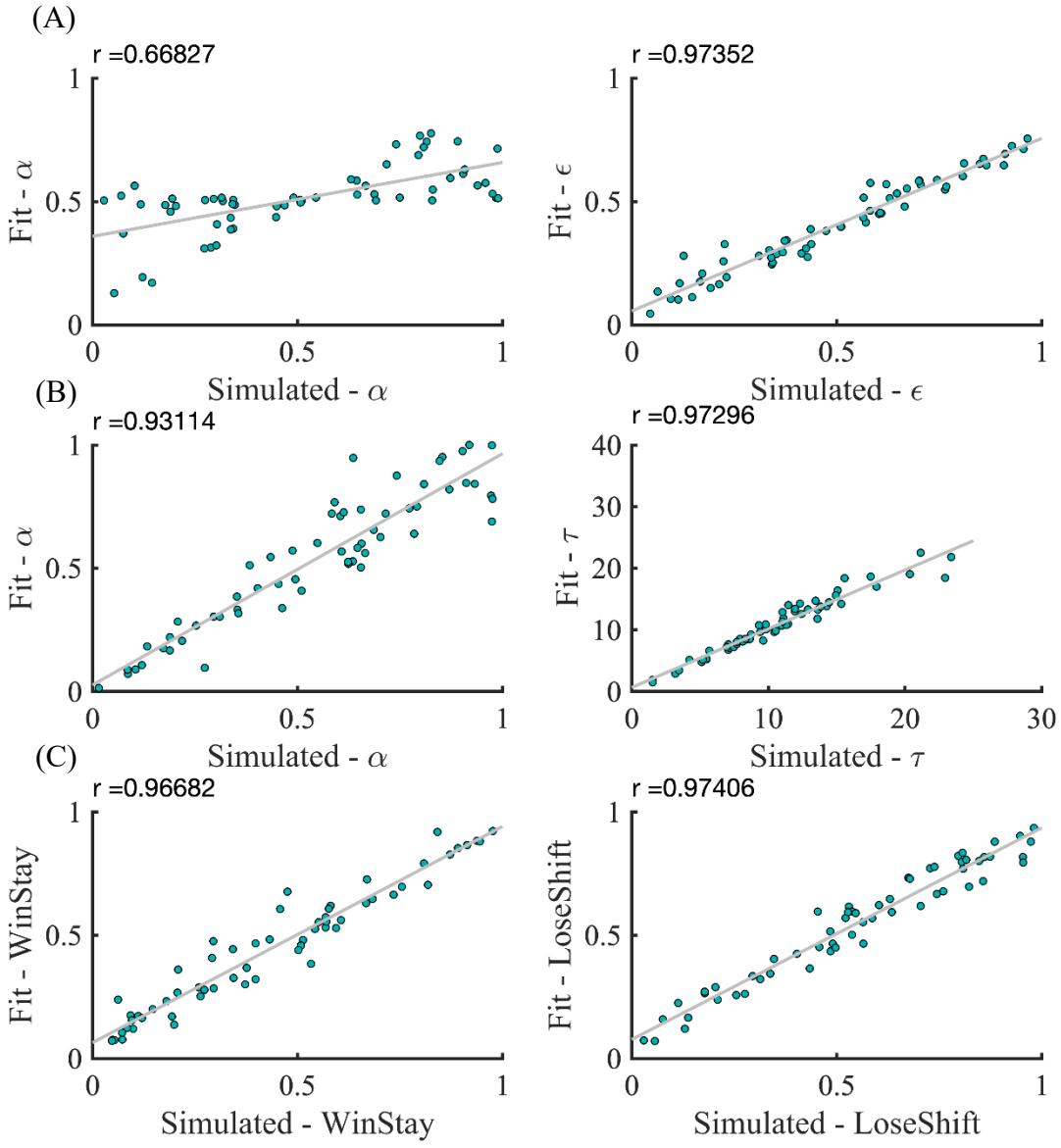

Figure 10. Experiment 2 - Parameter recovery. (A) Greedy model, (B) Softmax model, and (C) Win-stay Lose-shift. The r values indicate the correlation between simulated parameter values and fitted parameter values. α = learning rate, ε = epsilon parameter, and τ = temperature parameter.

When comparing across each model's ability to fit behaviour (figure 11), it was evident that only the chance model was unable to correctly simulate behaviour. The analysis of this difference showed an effect of model ($F(2.21, 121.54) = 0.32$, $p < .001$, $\eta_p^2 = .77$, $MSE = 12.13$) but no effect of condition ($F(1, 55) = 0.32$, $p = .57$, $\eta_p^2 = .01$, $MSE = 172.88$) and no condition by model interaction ($F(2.2, 121.54) = 0.32$, $p = .68$, $\eta_p^2 = .01$, $MSE = 12.13$). Note that Mauchly's Test for Sphericity was violated for both the model and Condition by model interaction, as such the degrees of freedom and *p*-values reported are the corrected values. Follow-up pairwise false-discovery rate corrected *t*-tests revealed that only the chance model ($t(56) = 15.71$, $p < .001$, $d = 1.80$) produced lower point values than the behavioural data when comparing the data using paired-samples *t*-tests. In contrast, the ε-greedy model ($t(56) = 2.09$, $p = .12$, $d = 0.10$), the Softmax model ($t(56) = 0.63$, $p = .91$, $d = 0.02$), and the Win-Stay Lose-Shift model ($t(56) = 0.10$, $p = .91$, $d = 0.00$) did not differ from the behavioural data.

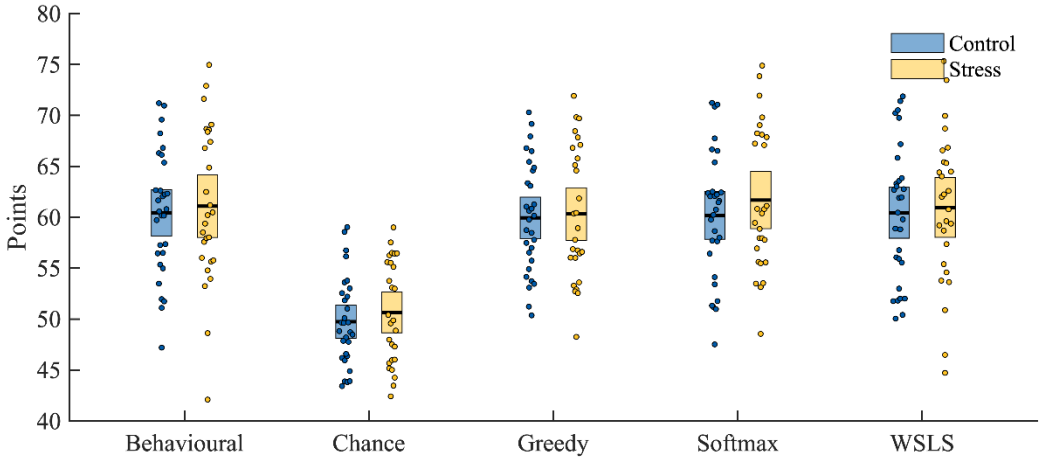

Figure 11. Experiment 2 - Model average performance simulation. The average points acquired across all four hundred trials for both behaviour and each of the four models. Boxes indicate 95% between-subject confidence intervals. Note: The key comparison is between the two left most columns (the behavioural data) and the columns demonstrating the data of each of the four models. WSLS = Win-Stay Lose-Shift.

When instead comparing each model's ability to fit trial-by-trial behavioural curves (figure 12), only the chance model was unable to produce a comparable win-stay percentage and lose-shift percentage curve. That is, for the ε-greedy, the Softmax, and the Win-Stay Lose-Shift model, we can see that the 95% confidence intervals have considerable overlap with the behavioural curves for both win-stay percentage and lose-shift percentage. The only major areas where the models do appear to produce differences from the behavioural curves is in the initial trials (typically between trials 1 to 50). This inability of the models to account for behaviour are more evident in the lose-shift percentages than the win-stay percentages. Moreover, given the lack of behavioural difference in win-stay percentage ($t(55) = 0.90$, $p = .38$, $d = -0.24$) and lose-

shift percentage ($t(55) = 0.45$, $p = .65$, $d = 0.12$) between the control and stress conditions when collapsing across trials, it is unsurprising that there are no apparent differences in the models' ability to fit behaviour when comparing the control and stress conditions across both win-stay and lose-shift percentages. In contrast, the chance model shows much lower win-stay percentage than behaviour across both the control and stress conditions, while showing a much higher lose-shift percentage than behaviour across the control and stress conditions.

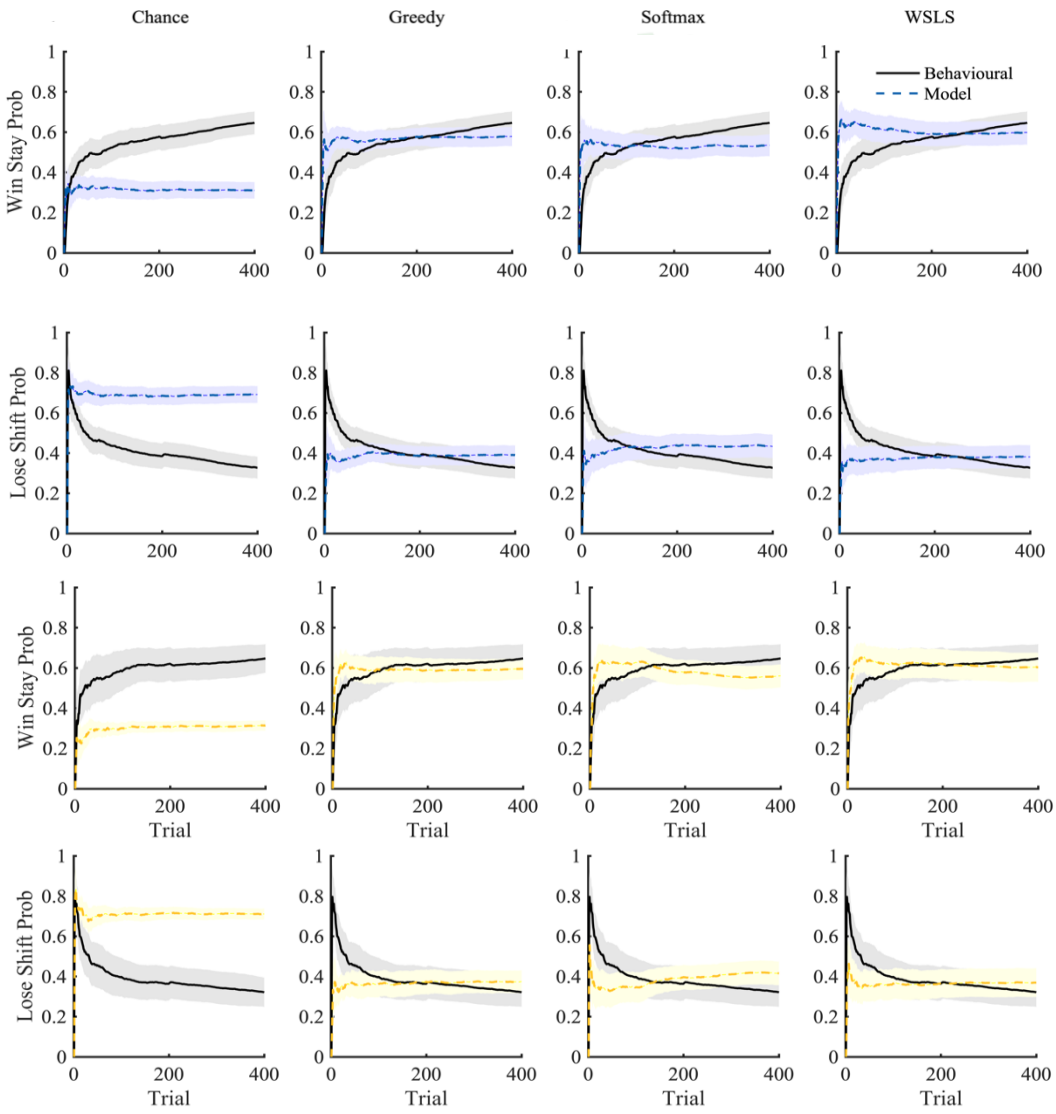

Figure 12. Experiment 2 - Trial-by-trial performance simulation. The black lines indicate the average behaviour of participants while the dotted lines indicate the model averages for both win-stay probabilities and lose-shift probabilities across all four-hundred trials. Blue lines are modelled control participants' averages and yellow lines are modelled stress participants' averages. Error bars are 95% between-subject confidence intervals. Note: the black lines demonstrate identical data within each row of figures.

To compare each of the three model's fit, I calculated a relative BIC using the chance model as a baseline measures (figure 13). That is, for each model, I subtracted the BIC of the chance model minus the BIC of the model itself, and a higher BIC indicates a better fit for the model. When comparing the relative fit across participants, the Win-Stay Lose-Shift model provided the best fit for the majority of participants (40/57, 70%), the Softmax model provided the best fit for the second largest group (16/57, 28%), and the ε-greedy model provided the best fit for the last participant (1/57, 2%). When comparing stress and control participants, the same trends emerge, as the Win-Stay Lose-Shift model provides the best fit for the majority of the control participants (22/30, 73%) and the majority of the stress participants (18/27, 67%), while the Softmax model provides the best fit for most of the other participants in the control condition (7/27, 26%) and the stress condition (9/27, 33%). The ε-greedy model provided the best fit for one participant in the control condition. To model participants behaviour in the rest of the task (i.e., to compare exploration and exploitation trials), I thus chose the best-fitting model on a participant-by-participant basis.

Lastly, I compared the overlap between each (non-baseline) model in their classification of exploration and exploration trials. Generally there was high agreement between the models in terms of the overlap of which trials were exploration and exploitation. While the SoftMax model and ε-greedy model had the highest overlaps of trials at 95%, the Win-Stay Lose-Shift model also had a large overlap with the ε-greedy model (78%), and the Softmax model (79%). Generally then, the models are in agreement of which trials are exploration and which trials are exploitation. In terms of exploration rate, the Win-Stay Lose-Shift model explored the least as 26.37% of trials were explorations, followed by the Softmax Model (29.48%) and then the ε-greedy model (30.81%)

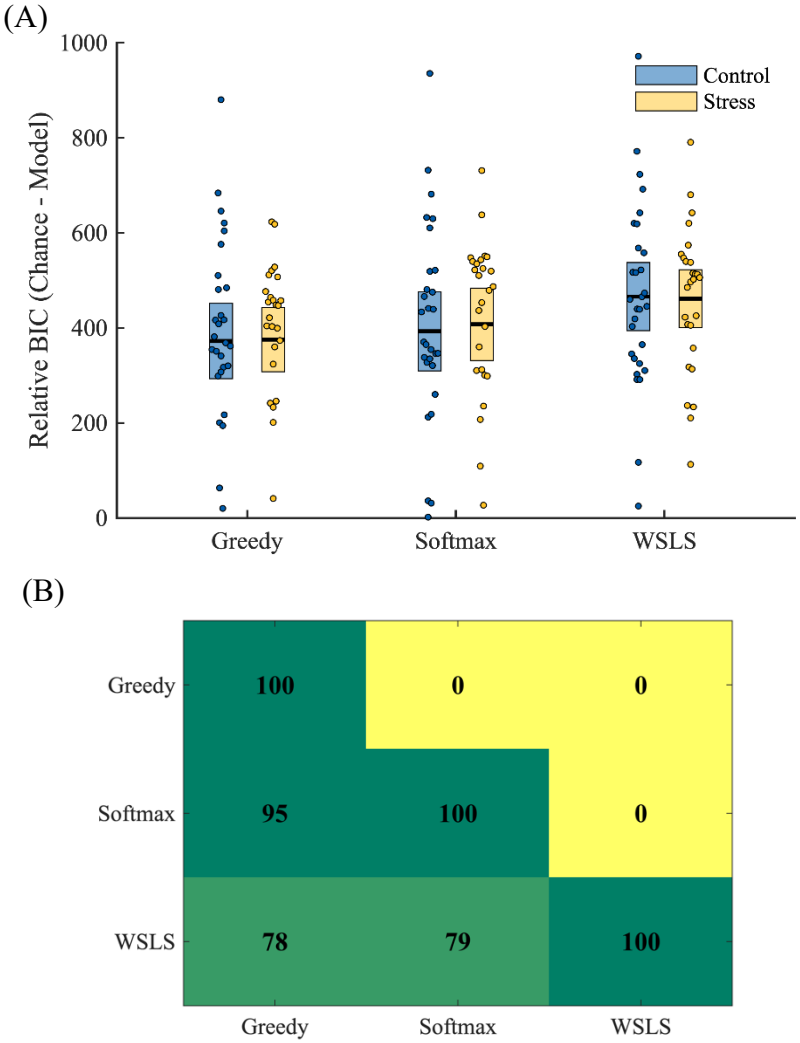

Figure 13. Experiment 2 - Model fit. (A) BIC indicates Bayesian Information Criterion from the chance model minus each of the other three models. Higher values indicate a better fit for the model. All models outperform the chance model. Boxes indicate 95% between-subject confidence intervals. (B) The percentage of explore and exploit trial classification overlap for each model.

Behaviour

In terms of behaviour, there was no effect of the acute stress response (figure 14). Stress ($\bar{X}_{str}$= 29.93%, 95% CI [23.23, 36.62]) and control ($\bar{X}_{con}$= 29.53% , 95% CI [23.92, 35.14]) participants did not differ in terms of exploration rate ($t(55) = 0.09$, $p = .92$, $d = 0.02$). In addition, a *t*-test comparing stress ($\bar{X}_{str}$= 56.93% , 95% CI [51.91, 61.94]) and control ($\bar{X}_{con}$= 57.91% , 95% CI [53.11, 62.71],) participants on optimal arm choice also revealed no effect of acute stress ($t(55) = 0.29$, $p = .77$, $d = 0.08$).

To determine how quickly participants developed a stable strategy in the task, I collapsed across control and stress conditions and examined optimal arm choices on trials 0 to 50, trials 51 to 100, and trials 350 to 400. To do this, I completed a 3 x 2 mixed ANOVA. However, there was no effect of condition ($F(1, 55) = 0.32$, $p = .58$, $\eta_p^2 = .01$, $MSE = 594.47$), no effect of time ($F(1.56, 85.92) = 2.66$, $p = .09$, $\eta_p^2 = .05$, $MSE = 294.66$), and no interaction ($F(1.56, 85.92) = 0.35$, $p = .65$, $\eta_p^2 = .01$, $MSE = 294.66$). Both the effect of time and the interaction violated sphericity, and as such, the Greenhouse-Geiser corrected degrees of freedom and *p*-values are reported here. A follow-up *t*-test of the first 50 trials (0 to 50) and the last 50 trials (350 to 400), showed no difference ($t(56) = 1.98$, $p = .06$, $d = -0.36$).

When comparing explore and exploit trials between the stress and control conditions, there was an effect of trial type but not condition (Table 5). For reaction time, there was an effect of trial type ($F(1, 55) = 14.01$, $p < .001$, $\eta_p^2 = .20$, $MSE = 1928.09$), but not condition ($F(1, 55) = 0.09$, $p = .77$, $\eta_p^2 = .00$, $MSE = 24734.46$), and there was no trial-type by condition interaction ($F(1, 55) = 1.74$, $p = .19$, $\eta_p^2 = .03$, $MSE = 1928.09$). The same pattern was true for win percentage, as there was an effect of trial type ($F(1, 55) = 403.18$, $p < .001$, $\eta_p^2 = .88$, $MSE =$

12.85), but not condition ($F(1, 55) = 0.32$, $p = .57$, $\eta_p^2 = .01$, $MSE = 57.03$), and no trial-type by condition interaction ($F(1, 55) = 0.00$, $p = .96$, $\eta_p^2 = .00$, $MSE = 12.85$).

Table 5

Experiment 2 - Explore-exploit behavioural data

Measure		Explore		Exploit	
		Mean	95% CI	Mean	95% CI
Reaction Time (ms)	Control	364.34	[323.29, 405.40]	344.38	[290.84, 397.93]
	Stress	366.51	[330.16, 402.87]	324.82	[283.56, 366.09]
Points	Control	50.40	[48.30, 52.50]	63.87	[61.82, 65.91]
	Stress	51.17	[48.90, 53.43]	64.71	[62.01, 67.41]

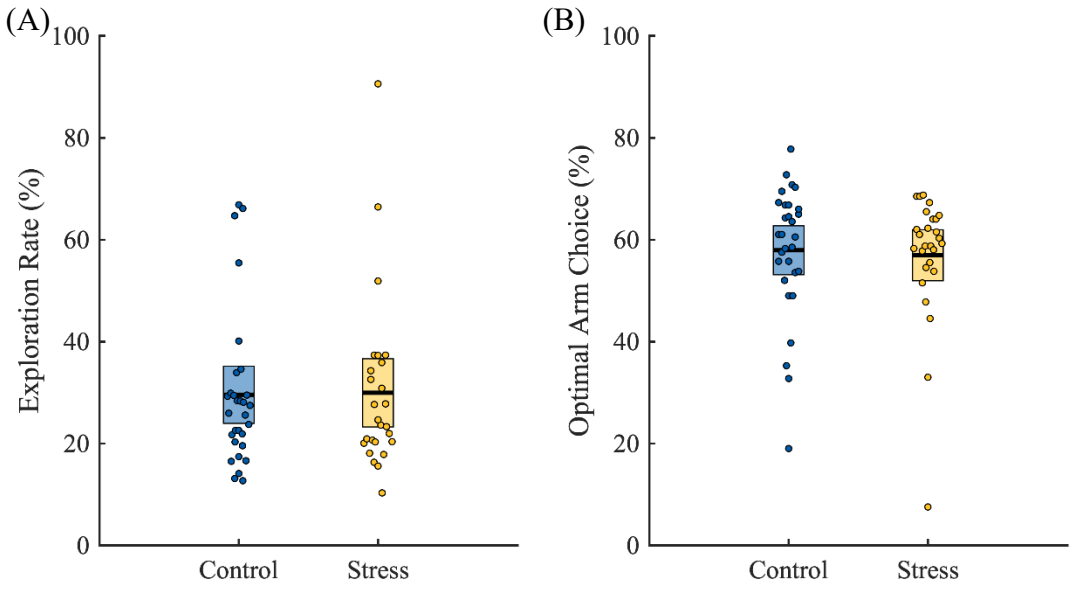

Figure 14. Experiment 2 - Exploration rate by condition. (A) Average exploration rate and (B) optimal arm choice for the control and stress conditions. Boxes indicate 95% between-subject confidence intervals

EEG

When examining the impact of the acute stress response on feedback (figure 15), it is evident acute stress modulates the feedback P300. A two-way (2 x 2) mixed ANOVA revealed only an interaction between trial type and condition (win/loss; $F(1, 55) = 8.95$, $p < .005$, $\eta_p^2 = .14$, $MSE = 2.09$). There was no effect of trial-type ($F(1, 55) = 3.06$, $p = .08$, $\eta_p^2 = .05$, $MSE = 2.09$), and no effect of condition ($F(1, 55) = 0.11$, $p = .74$, $\eta_p^2 = .00$, $MSE = 13.71$).

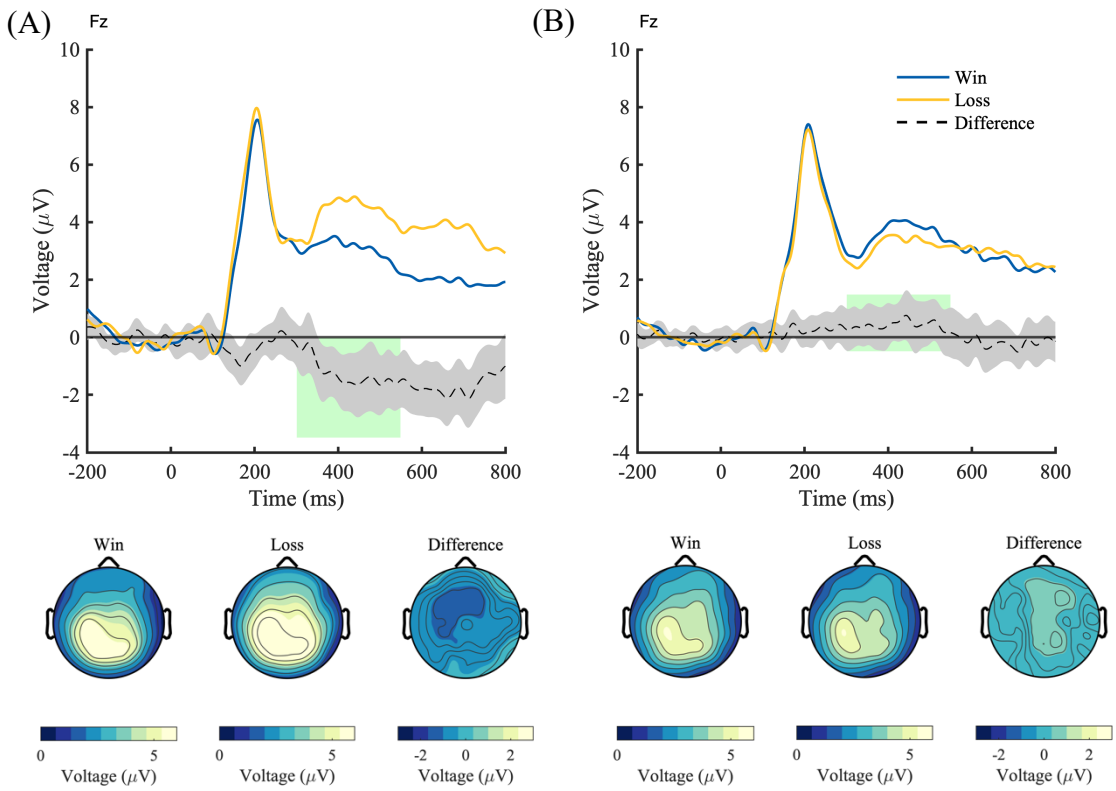

Figure 15. Experiment 2 - Feedback ERPs. (A) Control condition and (B) Stress condition. The shaded region indicates 95% between-subject confidence intervals. The green box indicates the region of analysis for the P300 to feedback. The scalp topographies indicate win trials, loss trials, and the difference (win trials minus loss trials) within the shaded region.

When examining the impact of the acute stress response on exploration signals (Figure 16), there was no effect of acute stress. A two-way (2 x 2) mixed ANOVA revealed an effect of trial type (exploration/exploitation; $F(1, 55) = 48.18$, $p < .001$, $\eta_p^2 = .47$, $MSE = 3.51$), but no effect of condition was present ($F(1, 55) = 0.69$, $p = .41$, $\eta_p^2 = .01$, $MSE = 14.20$), and no trial type by condition interaction ($F(1, 55) = 0.00$, $p = .98$, $\eta_p^2 = .00$, $MSE = 3.51$).

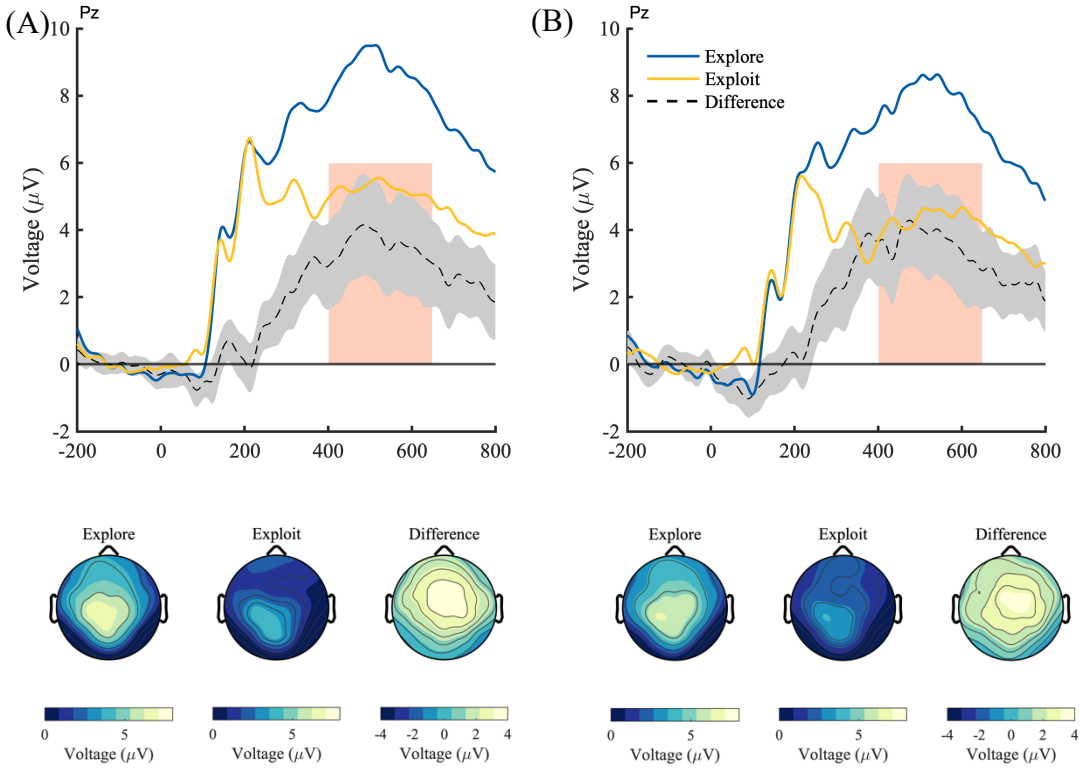

Figure 16. Experiment 2 - Exploration ERPs. (A) Control condition and (B) Stress condition. The gray shaded region indicates 95% between-subject confidence intervals. The pink box indicates the region of analysis for the P300 for exploration. The scalp topographies indicate explore trials, exploit trials, and the difference (explore trials minus exploit trials) within the shaded region

Table 6

Experiment 2 - Component values

Component		Win		Loss	
		Mean	95% CI	Mean	95% CI
Feedback P300	Control	2.89 μV	[1.97, 3.83]	4.21 μV	[3.01, 5.41]
	Stress	3.49 μV	[2.34, 4.65]	3.15 μV	[2.08, 4.22]
		Explore		Exploit	
		Mean	95% CI	Mean	95% CI
Exploration P300	Control	6.37 μV	[5.01, 7.66]	3.94 μV	[2.94, 4.93]
	Stress	5.79 μV	[4.41, 7.16]	3.34 μV	[2.54, 4.14]

Discussion

In the present work, I found that the acute stress response did affect the neurophysiology of participants but had no impact on behaviour – echoing the findings from experiment 1. The three reinforcement learning models I used to simulate behaviour showed an ability to both correctly model average task performance in terms of points acquired, and replicate trial-by-trial behaviour for both win-stay and lose-shift percentage curves. Importantly, there was a high degree of overlap between exploration and exploitation trial classification across the models. In terms of task performance, there were no differences between the acute stress and control conditions. However, there was the expected effect of exploration on both reaction time and points acquired, as exploration trials were both slower and had less points acquired on them than exploit trials (akin to experiment 1 and replicating Hassall & Krigolson, 2020). With regards to

the two components I examined, the feedback P300 and the exploration P300, I found an effect of the acute stress response on neural signals of feedback but not exploration. I found no effect of the acute stress response on the exploration P300, which does not replicate the tentative acute stress disruption of the switch P300 in experiment 1.

I found an effect of the acute stress response on feedback signals in the brain. In the present experiment the disruption of the feedback signal by the acute stress response was on the feedback P300. In contrast, in experiment 1, there was suggestive evidence that acute stress disrupted the reward positivity as there was no reward positivity under acute stress. Prior investigations of the effect of acute stress on feedback signals in the brain have shown that acute stress reduces both the reward positivity and the feedback P300 (Banis & Lorist, 2012; Paul et al., 2020). Thus, it could that the findings of experiment 2 fit into the line of evidence that acute stress simply disrupts feedback signals and processing. However, other work has argued that the feedback P300 signal represents a more top-down cognitive signal in feedback learning tied to motivation (Wu & Zhou, 2009). If interpreted within this framework by Wu & Zhou (2009) then the effect of acute stress on the feedback P300 may indicate a more specific effect of acute stress on top-down feedback processing. That is, in the present experiment the feedback P300 may in-part reflect a salience signal which is greater under losses as a signal to recruit cognitive control processes (Ridderinkhof et al., 2004). In turn, this top-down feedback signal is then modulated under acute stress. If the argument that the P300 reflects top-down, controlled feedback evaluation is correct, and given that work has shown that acute stress typically shifts participants away from controlled to habit-based modes of thought (Bogdanov et al., 2021; Schwabe & Wolf, 2011), then the present findings might be a neural indication of acute stress disrupting more top-down cognitive processes in the evaluation of feedback. I should note that although in

experiment 1 I did not observe an effect of the acute stress response on the P300, the effect of the acute stress response did reduce the effect size of the feedback P300 difference wave from a large effect in the control condition (d = 0.75) to a moderate effect in the stress condition (d = 0.47).

As noted, in experiment 2 there was no evidence of a reward positivity signal – meaning I was unable to replicate the previous experiment's reward positivity finding. Although I attributed the lack of a reward positivity to the use of point feedback, it is also possible that if participants were given monetary compensation for performance that could have produced a reward positivity. That is, prior work has shown that the reward positivity is enhanced when participants are given monetary compensation (Weinberg et al., 2014). Thus, researchers using contextual bandits with points might consider offering monetary compensation to participants in order to ensure that the reward positivity is present and detectable.

In experiment 2, I found no effect of the acute stress response on the exploration P300, an uncertainty signal. The lack of effect may stand in contrast to experiment 1 where I found tentative evidence that acute stress disrupted the uncertainty signal (the switch P300). The lack of effect of the acute stress response on the uncertainty signal in this paradigm is worth explaining. In terms of task design and experiment protocol, the only major difference between experiment 1 and experiment 2 was that the feedback provided in experiment 1 was win-loss feedback and in the present work the feedback was points. Another difference is that in experiment 1 I examined stay and switch trials, while in the present work I examined exploration and exploitation as per a computational model. Thus, perhaps my claim that switch and stay trials and explore and exploit trials may not be true. That is, perhaps switch trials are not equivalent to exploration trials. However, I compared the overlap of model classification of trials

(explore-exploit) with the more straight forward behavioural classification of trials (switch-stay) and found that the two were in general agreement with an 80% overlap in explore/switch and exploit/stay trial classification.

Another possible explanation is that there might have been differences in how the two tasks recruited attentional processes, leading to participants being more engaged in experiment 2. In experiment 1, participants only received binary information about their choice as it was either a win or loss. In experiment 2, participants received point feedback. Given that the bandit tasks were contextual, the use of point feedback would likely signal more information than a simple binary win or loss. As such, this additional information might have led participants to be more engaged in experiment 2 – preventing the exploration P300 from being disrupted by the acute stress response. Specifically, there is evidence that parietal P300 signals (which the exploration P300 is) are sensitive to attention and arousal (Polich & Kok, 1995) and thus the additional recruitment of attentional processes due to higher levels of engagement in experiment 2 could be preventing the exploration P300 from being disrupted.

Perhaps a more concerning explanation is that the lack of disruption of the uncertainty signal in experiment 2 is the result of a lack of power due to the weak effect of stress on uncertainty signals. As mentioned in the power analysis, it seems likely that this experiment is under powered to detect a small or medium effect which it could be the effect of acute stress on feedback processing is. Moreover, in experiment 1, I compared the difference waves against zero to tentatively infer an effect of acute stress. Thus, a more parsimonious explanation is that that the inability of experiment 2 to replicate experiment 1 with regards to how acute stress impacted the uncertainty signal comes down the tenuous effect observed in experiment 1 – that is the

effect of stress on the uncertainty signals may only small or medium and the experiment was underpowered to detect such an effect.

As with experiment 1, there was no effect of the acute stress response on either task performance (points acquired, reaction time) nor on exploration rate. While I hypothesized in experiment 1 that this could be due to the information conveyed by the win/loss feedback – it was not informative enough to provide a nuanced behavioural effect – this explanation does not hold given the lack of behavioural effects in experiment 2 where point feedback was given. My other suggestion for the lack of behavioural effects was that the contextual bandit may not be an ideal task for eliciting behavioural differences between conditions. One piece of evidence is that participants seem to learn to select an optimal arm in a relatively stable manner early in the task – that is, there is not a difference in optimal arm choice percentage between early trials and late trials. In fact, studies using contextual bandits have shown no performance differences between conditions, such as when comparing high anxiety to low anxiety participants (e.g., Aylward et al., 2019). As such, for experiment 3, I will instead use a stationary (i.e., learnable) bandit task. In this case, one bandit will produce more wins consistently across a block. At the beginning of each block, participants will have to re-learn which bandit is the ideal target. My hope is that this task will induce performance differences between stress and control conditions as it introduces a learnable component.

In sum, I found that the acute stress response only modulated feedback signals – but not uncertainty signals – in the present experiment. In addition, there was no effect of the acute stress response on behaviour. Thus, the present experiment some-what replicates the tentative effect of the acute stress response eon feedback signals from experiment 1 but on the feedback P300 and not the reward positivity. However, the fact that the contextual bandit task used in experiment 2

did not produce a clear reward positivity in participants, it's unclear if this disconfirms or confirms the findings from experiment 1, other than suggesting that the acute stress response modulates feedback learning more generally. The lack of effect of the acute stress response on exploration P300 could be explained by the fact that the effect of the acute stress response on the uncertainty signal observed in experiment 1 was only a tenuous effect. In experiment 3, I examined explore-exploit behaviour in learnable bandit task to investigate whether the contextual bandit tasks used previously were unsuited to detecting a behavioural effect. In addition, I investigated whether the assessment of explore-exploit behaviour depends on the time post-stressor onset. There is evidence that the impact of acute stress on cognitive abilities can depend on the delay between stressor and assessment (Shields et al., 2016). Thus, to determine whether this is the case in the explore-exploit dilemma I examined participants' explore-exploit behaviour following acute stress both immediately after an acute stressor and following a delay.

Chapter 4: Experiment 3

Introduction

In experiment 1, I observed a disruption in the reward positivity and the switch P300 following an acute stressor which I argued represented tentative modulations to participants' feedback learning and uncertainty assessments respectively. However, in experiment 2, I only observed an effect of the acute stress response on the feedback P300 but not the exploration P300. While these findings suggest tentative evidence that both feedback learning and uncertainty signals are modulated following acute stress, one issue I hope to resolve with experiment 3 is the lack of behavioural effect. Specifically, in neither experiment 1 nor experiment 2 did I observe any effects of the acute stress response on exploratory behaviour nor task performance. I believe that this was due to the task design as participants tended to remain relatively stable in optimal arm selection in both experiments when examining trial-by-trial behaviour. In experiment 1, while there was a difference in the trials 1 to 50 compared to trials 351 to 400, there was no difference in trials 51 to 100 compared to 351 to 400, suggesting that improvement in optimal arm selection had leveled off at this point. In experiment 2, a similar trend emerges, as there was no difference between trials 1 to 50 and trials 351 to 400 in terms of optimal arm selection. My suggestion was that the lack of behavioural effects is because the contextual bandit is not an ideal task design to elicit behavioural effects. Specifically, the evidence gleaned from optimal arm selection suggest that participants remain stable in their behaviour throughout the task.

Thus, the primary goal of experiment 3 is to determine whether the use of a block-based learning design might induce effects of the acute stress response on explore-exploit behaviour

and performance. Specifically, rather than use a contextual bandit task where reward probabilities shift throughout (as was done in experiment 1 and 2; Daw et al., 2006), I instead used a learnable two-armed bandit task where one of the options offers a consistently higher proportion of win feedback than the other option. In this case, participants completed multiple blocks of trials, where at the beginning of each block the winning option is randomly assigned to one of the two bandits. As such, participants had to re-learn which bandit is the better option every block. As prior work has shown that acute stress impacts performance in feedback learning tasks (e.g., Paul et al., 2020), I believe the use of a learnable bandit might elicit behavioural differences between stress and control conditions.

The second goal of this experiment is to determine whether the effects of the acute stress response on the explore-exploit dilemma differs depending on when participants complete cognitive tasks relative to the acute stressor. To test this goal, I examined both immediate and delayed effects of the acute stress response on the Oddball task and the two-armed bandit. To do this, I examined participants from experiment 1 (who completed the Oddball and two-armed bandit task immediately post-stress) and the participants in experiment 2 (who completed the Oddball and two-armed bandit task after a delay of around 30 minutes). Prior work has shown that the SAM axis is activates immediately during stress while the HPA axis reaches maximal activity after a delay (Joëls et al., 2011). Importantly, SAM axis activity is closely linked to an increase the activity of norepinephrine (Morilak et al., 2005). Phasic norepinephrine bursts in turn are suggested to be tied to P300 activity (Nieuwenhuis et al., 2005) and the exploration P300 (Hassall, 2019). In contrast, activity of the HPA stress-response axis has been closely tied to modulations of dopamine (Arnsten, 2009). Relatedly, feedback components such as the reward positivity are hypothesized to reflect dopaminergic activity, specifically a dopaminergic burst

when an action is better than expected (Proudfit, 2015; Schultz et al., 1997). Thus, it seems possible that by assessing explore-exploit behaviour immediately post-stress and following a delay could reveal an interaction in the neural effects of the acute stress response (both SAM and HPA axis effects) on feedback and uncertainty. Moreover, if an interaction between the neural effects and the behavioural findings occurs, it could provide an indication of how feedback and uncertainty signals interact under exploratory behaviour.

Relatedly, work has shown that the time of assessment relative to a stressor can impact the effect of the acute stress response. As one example, in a meta-analysis Shields and colleagues found that increasing the delay post-stressor increased the effect of acute stress on working memory (2016). As well, prior work on the effect of acute stress on memory has shown that the delayed effect of cortisol is necessary to induce behavioural effects (Joëls et al., 2011). Thus, even if the SAM axis is not directly tied to timing effects, it could be that the time of assessment is related to the activity of cortisol, which peaks 20 to 40 minutes post-stressor onset (Dickerson & Kemeny, 2004). As such, by varying the time of assessment post-stressor, I may be able to provide insight into this problem during the explore-exploit dilemma.

In addition, in the present experiment, I collected data from a conventional visual Oddball paradigm. In the Oddball paradigm, participants are exposed to a series of stimuli which vary in their frequency and the Oddball paradigm is the most common method to elicit the P300 (Donchin, 1981; Donchin & Coles, 1988). My rationale for including this task is to determine whether the acute stress response is modulating the exploration P300 specifically or whether acute stress is modulating the P300 signal more generally. As noted, the P300 is elicited in a wide variety of circumstances, is sensitive to a variety of modulations, and there is currently no well accepted theory on the P300's functional significance (Polich, 2020). Work has shown that

the posterior P300 is sensitive to task uncertainty (Kopp et al., 2016) and is elevated when participants decide to explore (Hassall et al., 2013, 2019; Hassall & Krigolson, 2020). Thus, my belief is that the exploration P300 in the bandit tasks used previously reflected uncertainty rather than the claim that the P300 is a context updating signal tied to stimulus processing and working memory (Polich, 2007). If in the present work I do not observe an effect of the acute stress response on the P300 in the Oddball paradigm, it would contribute evidence to the claim that the acute stress response is impacting the uncertainty signal specifically.

As such, in experiment 3, I used data collected from participants from both experiment 1 and 2. More specifically, participants in experiment 1, completed the Oddball task and two-armed bandit task immediately following acute stress – when I expect that the effects of the SAM axis are greater than the HPA axis. In contrast, participants in experiment 2 completed the Oddball and two-armed bandit after a considerable delay – when the effects of cortisol and the HPA axis are maximal. My expectations for goal 1 were that generally the acute stress response would modulate the neural signals involved (uncertainty, feedback) as I observed previously in experiments 1 and 2. Moreover, my hope is the addition of a learnable bandit requiring learning in each block would correspondingly reveal effects of the acute stress response on bandit performance and exploration rate. As with experiment 2, I validated and used a series of reinforcement learning models to classify trials as explorations and exploitations. For my second goal, I expected that there would be a dissociation between the time of assessment and the behavioural and neural effects of acute stress. Specifically, I believe that the exploration P300 will be impacted both immediately and following a delay. In contrast the feedback signals (the reward positivity, and the feedback P300) will only be impacted after the delay. My hope is that if there were effects of acute stress on exploration behaviour and two-armed bandit performance,

then these effects would either only be present at the delay or would be larger in comparison to the groups who were assessed immediately. A tertiary goal is to determine whether effects of the acute stress response are present in the oddball P300 in a visual Oddball task and the exploration P300 in the two-armed bandit task. I expect that the effect of the acute stress response will only be present on the exploration P300 rather than the oddball P300, as the exploration P300 provides a more direct index of uncertainty rather than contextual updating.

Method

Experiment 3 had an almost identical method to experiment one and two except where noted.

Participants

A total of 111 participants were recruited from the University of Victoria. However, two participants were removed due to data quality issues (independent component analysis could not be completed), two participants were removed due to task performance issues in the four-armed bandit (see experiment 1), and one participant chose to end the experiment before completing the visual Oddball and two-armed bandit task. Thus, the final sample had a sample size of 106 participants (58 females, age range 17 to 41, mean age = 21.39, 95% CI [20.56, 22.23]; see Table 7 for a breakdown).

Note that participants are taken from experiment 1 and 2. However, there are some slight differences. One participant in the stress early condition's data was unable to be analyzed in experiment 1 due to ICA not being completed. However, their data was usable here, as such they were included. As well, one participant in the control and late condition chose to end the

experiment after the four-armed bandit and as such did not complete the two tasks discussed in this experiment.

Table 7

Experiment 3 - Participant Counts

	Early		Late	
	Male	Female	Male	Female
Control	10	17	15	14
Stress	14	9	10	17

Power. In experiment 2, I found a large effect of stress on the feedback P300 (η_p^2 = .14; identical to Paul et al., 2020). Using G*Power (Faul et al., 2009) to detect a neural effect of acute stress with 80% power and assuming α = .05 using a two-way ANOVA, I would need a total of 111 participants across all the four conditions (control-early, control-late, stress-early, stress-late; 27 participants in each group), which was close to the final sample size. In terms of a behavioural effect of acute stress, I used the effect size from Paul and colleagues (2020) who found a large effect of stress on feedback learning (η^2 = .13). I found I would need a total of 111 participants across the four conditions.

Apparatus and Materials

Questionnaires and Heart Rate. Six participants had heart-rate connection issues as the data stopped streaming during the experiment), and thus I ended up with a total of 100 participants with usable heart-rate data. For the Positive and Negative Affect Schedule – Negative Affect scale, four participants had incomplete data and I ended up 102 participants with

usable data. In terms of the State Trait Anxiety Inventory – State scale, five participants had incomplete data, as they either missed or chose not to answer a question, and I ended up with a total of 100 participants with usable State Trait Anxiety Inventory – State scale data.

Two-Armed bandit. The participants completed a conventional two-arm bandit (figure 17; also known as a gambling task). Before beginning the task, participants were instructed that their goal in the task was to select an option on each trial with a keypress and through the feedback given from their selection determine which of two options won more than the other option. These two options were represented as two coloured squares and the underlying win probability of the two options was 60% and 10%. While participants were informed of the underlying win probability before beginning the task, they were not told which square corresponded to which win probability. The squares were also randomly assigned to either the left or right side of the screen on each trial, with each side of the screen requiring a different response (a response with the left hand for the left side and conversely a response with the right hand for the right side). To make a selection, participants either pressed the “f” key for a left selection or the “j” key for a right selection. At the conclusion of each block of trials, the colours of the squares changed, and the participants had to re-learn which coloured square won more than the other. Participants completed 6 blocks of 20 trials each.

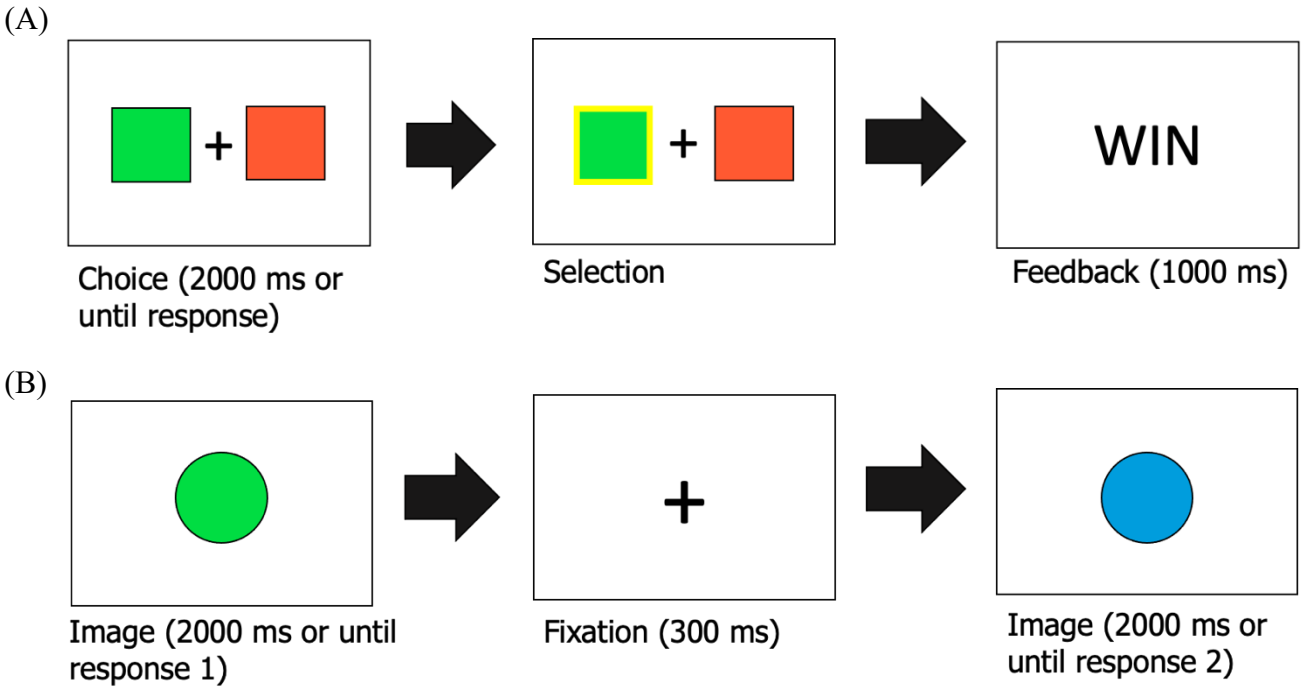

Figure 17. Experiment 3 - Tasks. (A) Two-armed bandit task. (B) Visual oddball task.

Oddball task. The other task that participants completed was a visual Oddball task (Donchin & Coles, 1988). In this task, participants were presented with a series of coloured circles (either green or blue). For participants, one circle was presented 70% of the time (the frequent circle) while the other circle was presented 30% of the time (the infrequent circle). Participants were required to make a different keypress to each coloured circle – either on the left or right side of the keyboard. Whether the keypress for the frequent circle was located on the left side or the right side was varied randomly across the sample of participants. Participants completed 4 blocks of 50 trials each.

Protocol

Participants completed two tasks: (1) a two-armed bandit, and (2) a visual Oddball task. The order of tasks was counterbalanced across participants. Participants either completed the two cognitive tasks immediately following either the Trier Social Stress Test or the placebo-Trier Social Stress Test or completed the cognitive tasks following a 40-minute delay post-stressor or placebo.

Computational Modeling

Much like experiment 2, to determine when participants decided to explore or exploit, I modelled participants' behaviour using four difference models. The four models I used were the chance model (to provide a baseline measure), an ε-greedy model, a Softmax model, and a Win-Stay Lose-Shift model. In this case, r_t is the reward value for the stimulus – either 0.5 for a win or -0.5 for a loss. All modelling was completed in MATLAB (Version 9.7, Mathworks, Natick, USA).

Action Selection.

Chance Model. For the chance model, action selection was based on a bias parameter. That is, on each trial (t) the probability of making an action (a) was fit to a parameter biased to select one square over the other:

$$P_t(a_i) = \begin{cases} Square\ 1 & if\ b > \mathcal{U}(0,1) \\ Square\ 2 & otherwise \end{cases} \qquad (4.3)$$

where $\mathcal{U}$ is a uniform distribution sampling a number between 0 and 1 and b is the bias parameter. Thus, the model attempted to account for if participants were biased to select one square over the other (for example if they preferred the color) regardless of win or loss feedback.

The ε-greedy model, Softmax model, and Win-Stay Lose-Shift model were similar to the three corresponding models used in experiment 2. The only major difference was that for the Win-Stay Lose-Shift model the selection of a stimulus used the following simple rules: (1) if the reward given by the stimulus on the trial was a win then the same action is selected with the probability $P(stay \mid win)$ & (2) if the reward given by the stimulus on the trial was a loss then the action was avoided with the probability $P(shift \mid loss)$.

Parameter Optimization. All parameters were optimized using the optimization toolbox provided by MATLAB (Version 9.7, Mathworks, Natick, USA). One important difference between the ε-greedy model fit in the present work and the ε-greedy model fit in experiment 2, is that the learning rate parameter was kept constant ($\alpha = .5$) rather than included as a parameter to be fit. The reason for this is because the learning rate parameter in the ε-greedy model was unable to be recovered during parameter recovery. That is, there was no correlation between simulated and fit model parameters for the ε-greedy learning rate. See table 8 for the starting distribution for each parameter.

Table 8

Experiment 3 - Starting distribution for parameter optimization

Model	Parameter	Distribution	Distribution Parameters
Chance	Bias (b)	Beta	α_{beta} = 1.1, β_{beta} = 1.1
ε-greedy	Epsilon (ε)	Beta	α_{beta} = 1.1, β_{beta} = 1.1
Softmax	Learning Rate (α)	Beta	α_{beta} = 1.1, β_{beta} = 1.1
	Temperature (β)	Gamma	k_{gam} =1.2, θ_{gam} = 5

Win-Stay	Win-Stay Probability ($P(stay \mid win)$)	Beta	α_{beta} = 1.1, β_{beta} = 1.1
Lose-Shift	Lose-Shift Probability ($P(shift \mid loss)$)	Beta	α_{beta} = 1.1, β_{beta} = 1.1

Parameter Recovery. I conducted parameter recovery across each of the parameters of the models (Wilson & Collins, 2019) to ensure that the model parameters are identifiable. I ran each model on 6 blocks where one bandit had a win value of .6 and the other bandit had a win value of .1. These win values were randomly assigned to one bandit or the other bandit on each block. I conducted this simulation across each of the models using 100 simulated participants.

Data Processing

Data were processed identically to experiment 1 and 2. The sole difference was the events of interest from which segments were created from. That is, the segments were created either relative to stimulus (Oddball task) or feedback (two-armed bandit task) onset. The final data had an artifact rejection rate of 5.13%, 95% CI = [3.71, 6.55] of trials in the Oddball task, and 9.71%, 95% CI = [8.03, 11.39] of trials in the two-armed bandit task.

Data Analysis

Modelling. Both model comparison and the definition of trials as either exploration or exploitation was identical to experiment 2. The only difference in experiment 3 was that to determine the model's ability to fit behaviour, I compared the ability of each model to correctly replicate the behavioural patterns of the participants through win probabilities and win-stay/lose-shift probabilities. To determine how well the models were able to fit behaviour, I ran each model through the task on the participants' optimized parameters and computed the average win percentage (rather than the points acquired as in experiment 2). I then ran a 5 x 2 x 2 mixed ANOVA on the average points for both the behavioural data and the model data. Specifically,

the first factor was model (within: behavioural, chance, ε-greedy, Softmax, Win-Stay Lose-Shift), the second factor was condition (between: stress, control) and the third factor was time (between: early, late). I examined all three main effects and the interaction terms. I measured effect size using η_p^2, and measured variability with the mean squared error (*MSE*) term. Follow-up paired samples or independent samples *t*-tests were conducted using a false-discovery rate correction (Benjamini & Hochberg, 1995).

Behaviour. I examined performance in both the Oddball task and the two-armed bandit task[5]. For the oddball task, I examined both the time participants took to hit a key (reaction time) and the number of errors made (errors). For the two-armed bandit task, I examined the total number of wins participants received divided by their total number of valid trials (win percentage), the number of times they selected the arm with the highest win percentage divided by the total number of valid trials (optimal arm selection), and how often they explored divided by the sum of how often they choose to exploit and explore (exploration rate). I compared participants on these measures using a 2 x 2 between-subjects ANOVA. The first factor was time (early, late) while the second factor was condition (stress, control). I examined both main effects and the interaction. I measured effect size using η_p^2, and measured variability with 95% between-subject confidence intervals. Follow-up paired sample t-tests were conducted using a false-discovery rate correction (Benjamini & Hochberg, 1995).

---

[5] In the interest of space, I did not report the explore and exploit trial differences for reaction time and win-percentage. However much like experiment 2, explore trials were slower and had more losses.

EEG. To assess the neural effects of the acute stress response, I examined the following components: (1) oddball P300, (2) reward positivity, (3) feedback P300, (4) exploration P300. The oddball P300 was assessed in the visual Oddball task by examining the neural response to frequent and infrequent trials. The reward positivity and feedback P300 were assessed in the two-armed bandit task by collapsing across explore-exploit decisions and examining the neural response to feedback (wins, losses). The exploration P300 was assessed in the two-armed bandit task by collapsing across feedback type and examining the neural response to exploration (explore, exploit).

To assess the effect of the acute stress response on the oddball P300, ERPs were constructed on each trial type (frequent circles, infrequent circles) to stimulus presentation at electrode Pz (Polich, 2007). I chose a window based on the peak of the grand-grand average (peak = 370 ms; Kappenman & Luck, 2016) which ended up being a 200-millisecond window between 270 and 470 ms.

To assess the neural effects of the acute stress response on feedback learning, I examined both the reward positivity and the P300 to feedback in the two-armed bandit task. First, I collapsed across explore and exploit decisions. ERPs were computed at electrode FCz (Krigolson et al., 2019; Williams et al., 2021). I choose a window based on the peak of the grand-grand average (peak = 302 ms; Kappenman & Luck, 2016) which was a 100-millisecond window between 252 and 352 milliseconds. To examine the feedback P300 I used the same approach. That is, using the grand-grand average at FCz, I identified a peak (438 ms) and selected a 125 ms window around that peak (between 375 ms and 500 ms).

For the exploration P300, I collapsed across feedback type and instead compared explore and exploit trials. I found the electrode where the grand-grand average value (Kappenman &

Luck, 2016) was maximum within a time window comparable to previous work (Hassall & Krigolson, 2020) which was electrode CP1. Then I took the time of the peak of the grand-grand average (400 ms) and computed each participant's P300 using a +/- 100 ms time window to calculate the mean of the difference wave (300 to 500 ms).

For all four ERP comparisons (oddball P300, reward positivity, feedback P300, exploration P300), I compared participants using a 2 x 2 between-subjects ANOVA where factor one was time (between: early, late) and factor two was stress condition (between: control, stress). I measured effect size using η_p^2, and measured variability with 95% between-subject confidence intervals. Follow-up planned comparisons were conducted on the difference wave for each condition (control-early, control-late, stress-early, stress-late) and were false-discovery rate corrected (Benjamini & Hochberg, 1995). I also computed Cohen's *d* and 95% between-subject confidence intervals for the difference wave in each condition.

For all components I calculated split-half reliability to measure internal consistency. For the oddball P300, frequent trials had a split-half reliability of .70 while infrequent trials had a split-half reliability of .60. For the reward positivity, wins had a split-half reliability of .73 while losses had a split-half reliability of .74. For the feedback P300 wins had a split-half reliability of .79 while losses had a split-half reliability of .80. For the exploration P300 exploration trials had a split-half reliability of .72 while exploitation trials had a split-half reliability of .88.

I also ran Pearson correlations comparing the neural findings and behaviour. I collapsed across all conditions for the correlations. The first correlation was between exploration and win percentage. The second correlation was between win percentage and the feedback P300 difference wave. The last correlation was between exploration rate and the explore P300 difference wave. All correlations were tested for significance, and all *p*-values were corrected

using a false-discovery rate correction (Benjamini & Hochberg, 1995). All statistics were completed in R (version 4.0.3; R Core Team, 2021).

Results

Manipulation Checks

As expected, the acute stressor was found to modulate the three stress manipulation checks (Table 9). Specifically, the acute stressor led to a larger increase in heart-rate relative to baseline in comparison to the control condition ($F(1, 96) = 39.276$, $p < .001$, $\eta_p^2 = .29$, $MSE = 124.60$). Interestingly, there was an effect of time, with the early condition having a lower change in baseline than the late condition ($F(1, 96) = 6.04$, $p < .05$, $\eta_p^2 = .06$, $MSE = 124.60$). There was no interaction ($F(1, 96) = 2.82$, $p = .10$, $\eta_p^2 = .03$, $MSE = 124.60$). When examining scores on the Positive and Negative Affect Schedule – Negative Affect scale, the stress condition again had a larger change in baseline than the control condition ($F(1, 98) = 23.51$, $p < .001$, $\eta_p^2 = .20$, $MSE = 1042.471$). In the case of the Positive and Negative Affect Schedule – Negative Affect scale there was no effect of time ($F(1, 98) = 1.43$, $p = .24$, $\eta_p^2 = .02$, $MSE = 1042.471$) nor any interaction ($F(1, 98) = 0.40$, $p = .53$, $\eta_p^2 = .00$, $MSE = 1042.471$).

When examining the State Trait Anxiety Inventory – State scale, the stress condition also had a larger change in baseline than the control condition ($F(1, 96) = 26.15$, $p < .001$, $\eta_p^2 = .21$, $MSE = 1289.40$). Much like with heart rate, there was also an effect of time ($F(1, 96) = 4.79$, $p < .05$, $\eta_p^2 = .05$, $MSE = 1289.40$), as the early condition had a lower change in baseline than the stress condition. Lastly, for the STAI-S there was an interaction ($F(1, 96) = 4.36$, $p < .05$, $\eta_p^2 = .04$, $MSE = 1289.40$). Follow-up pairwise comparison revealed that this effect was primarily driven by the stress condition, as the stress-late condition had a higher change in baseline than

the stress-early condition ($t(47) = 2.76$, $p < .01$, $d = 0.79$) while there was no difference in the control-early or control-late condition ($t(51) = 0.08$, $p = .94$, $d = 0.02$).

Table 9
Experiment 3 - Manipulation check results

		Control		Stress	
		Mean	95% CI	Mean	95% CI
Heart Rate (%)	Early	-0.19	[-4.64, 4.27]	10.15	[3.96, 16.32]
	Late	1.56	[-0.69, 3.81]	19.47	[13.74, 25.19]
PANAS (%)	Early	-3.38	[-9.13, 2.36]	33.81	[12.47, 55.15]
	Late	9.02	[-4.14, 22.20]	37.63	[25.20, 50.06]
STAI-S (%)	Early	12.59	[0.79, 24.38]	33.32	[17.79, 48.85]
	Late	13.29	[-0.13, 26.71]	62.68	[47.38, 77.99]

Note: The Early Data includes one additional stress participant, as such the data are slightly different from Chapter 2 (Figure 4). Late data is missing one participant and are slightly different from Chapter 3 (Figure 9).

Modelling

When comparing model fit to model simulation, it appears that most of the model parameters can be recovered effectively (all $r > .77$; figure 18). Although the temperature parameter of the Softmax model had a slightly lower correlation ($r = .77$) than any of the other parameters (all $r > .97$), attempts to produce a better fit through reducing the maximum of the starting parameter space to 20 did not improve the correlation value nor did including other nuisance parameter values such as a right or left side selection bias (as suggested by Wilson and Collins, 2019).

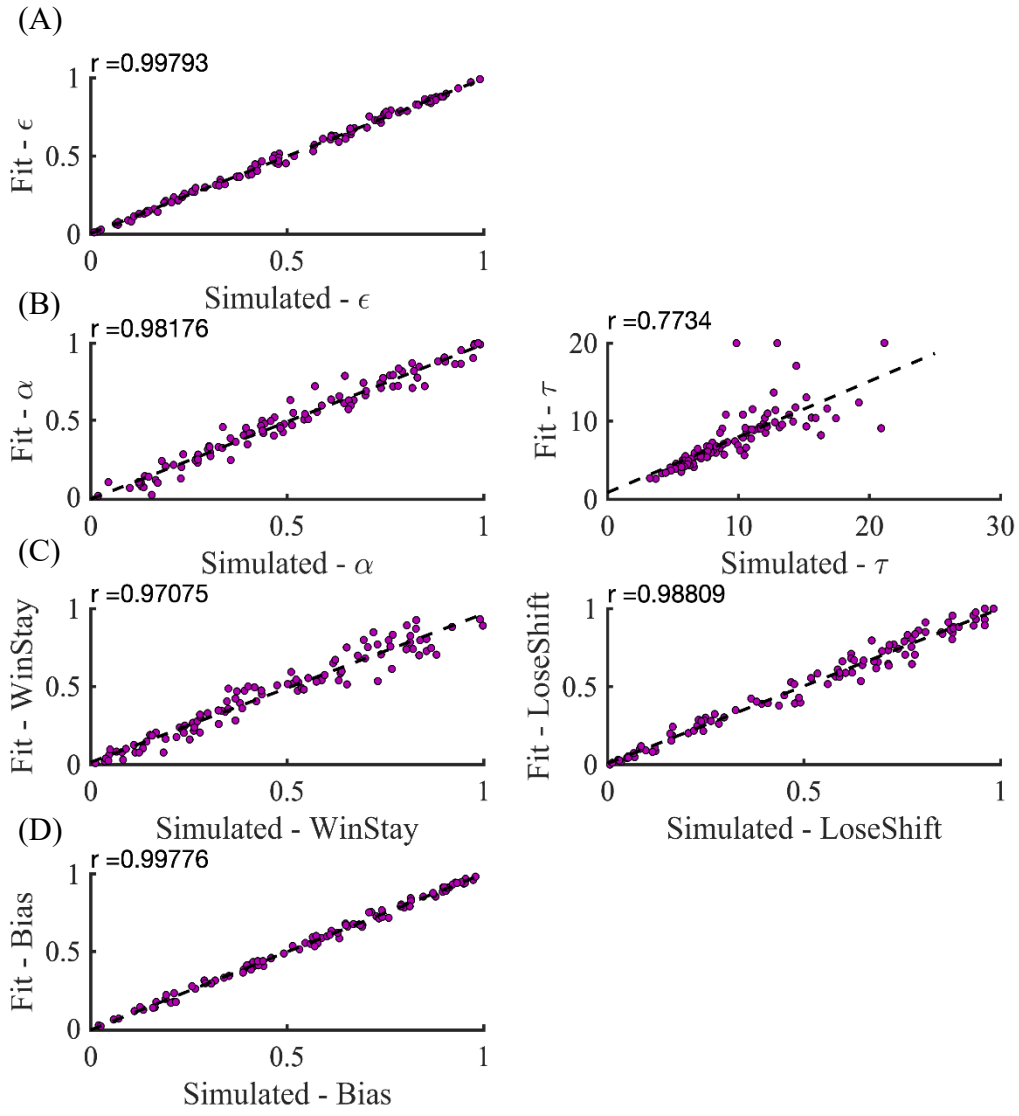

Figure 18. Experiment 3 - Parameter recovery. (A) Greedy model, (B) Softmax model, (C) Win-stay Lose-shift, and (D) the Chance model. The *r* values indicate the Pearson correlation between simulated parameter values and fitted parameter values. α = learning rate, ε = epsilon parameter, and τ = temperature parameter.

When comparing across each model's ability to fit behaviour, both the chance and Win-Stay Lose-Shift models appeared unable to account for behaviour. The mixed ANOVA (2 x 2 x

5) revealed several effects. Most importantly, there was an effect of model ($F(3.44, 351.25) = 228.92$, $p < .001$, $\eta_p^2 = .69$, $MSE = 50.89$) and the only significant interaction was the model by condition interaction ($F(3.44, 351.25) = 3.24$, $p < .05$, $\eta_p^2 = .03$, $MSE = 50.89$). All other main effects and interactions were non-significant (all $p > .06$). Note that Mauchly's Test for Sphericity was violated for both the model and stress by model interaction, as such the degrees of freedom and *p*-values reported are the Greenhouse-Geisser corrected values.

When collapsing across condition and time, follow-up pairwise false-discovery rate corrected *t*-tests for the effect of model revealed that the chance model ($t(105) = 14.57$, $p < .001$, $d = 2.04$) and the Win-Stay Lose-Shift model ($t(105) = 17.38$, $p < .001$, $d = 2.68$) produced lower win percentages than the behavioural data. In contrast, neither the ε-greedy model ($t(105) = 1.42$, $p = .31$, $d = -0.18$) nor the Softmax model ($t(105) = 0.05$, $p = .96$, $d = 0.01$), differed from the behavioural data. When examining the model by condition interaction, there was an effect of acute stress in the behavioural data ($t(104) = 2.86$, $p < .01$, $d = 0.57$) and in the Softmax Model ($t(104) = 2.41$, $p < .05$, $d = 0.47$), but not in any of the other models (all $p > .20$).

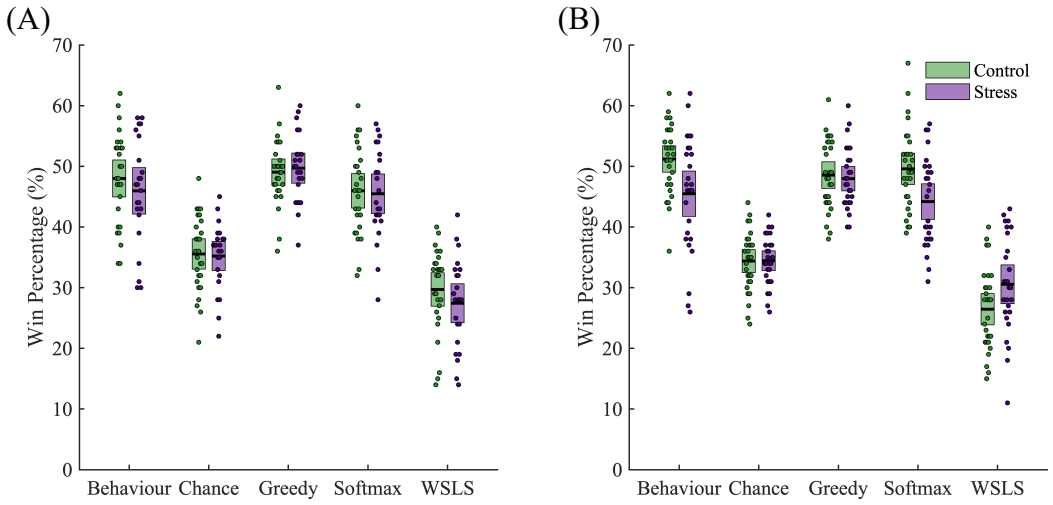

Figure 19. Experiment 3 - Model simulation of performance (A) Early condition and (B) late condition. Figure demonstrates the average points acquired across all 400 trials for both behaviour and each of the four models. Error bars indicate 95% between-subject confidence intervals. Note: The important comparison is between the left most columns (the behavioural data) and each of the four models for both (A) and (B). WSLS = Win-Stay Lose-Shift.

When instead comparing each model's ability to fit trial-by-trial behavioural curves (figure 20 and 21), both the chance and Win-Stay Lose-Shift models appeared unable to produce comparable win-stay percentage and lose-shift percentage curves. That is, for the ε-greedy and the Softmax models, we can see that the 95% confidence intervals have considerable overlap with the behavioural curves for both win-stay percentage and lose-shift percentage. There is a small inability of the ε-greedy model to account for behaviour in lose-shift percentages, which produced slightly higher lose-shift percentages on a trial-to-trial basis than the behaviour. In contrast, both the chance model and the Win-Stay Lose-Shift model show much lower win-stay

percentage than behaviour across both the control and stress conditions and across the early and late conditions, with less difference in the lose-shift percentage curves as the 95% confidence intervals overlap between behaviour and each of the two models.

There was a lack of effect of condition when collapsing across trial number in the win-stay percentages ($F(1, 102) = 2.27$, $p = .13$, $\eta_p^2 = .02$) and lose-shift percentage ($F(1, 102) = 1.60$, $p = .21$, $\eta_p^2 = .02$). Thus, it is unsurprising that there are no apparent differences in the models' ability to fit behaviour when comparing the control and stress conditions across both win-stay and lose-shift percentages. As well, there was no difference across the early and late conditions for both win-stay percentage ($F(1, 102) = .21$, $p = .65$, $\eta_p^2 = .00$) and lose-shift percentage ($F(1, 102) = 0.36$, $p = .55$, $\eta_p^2 = .00$).

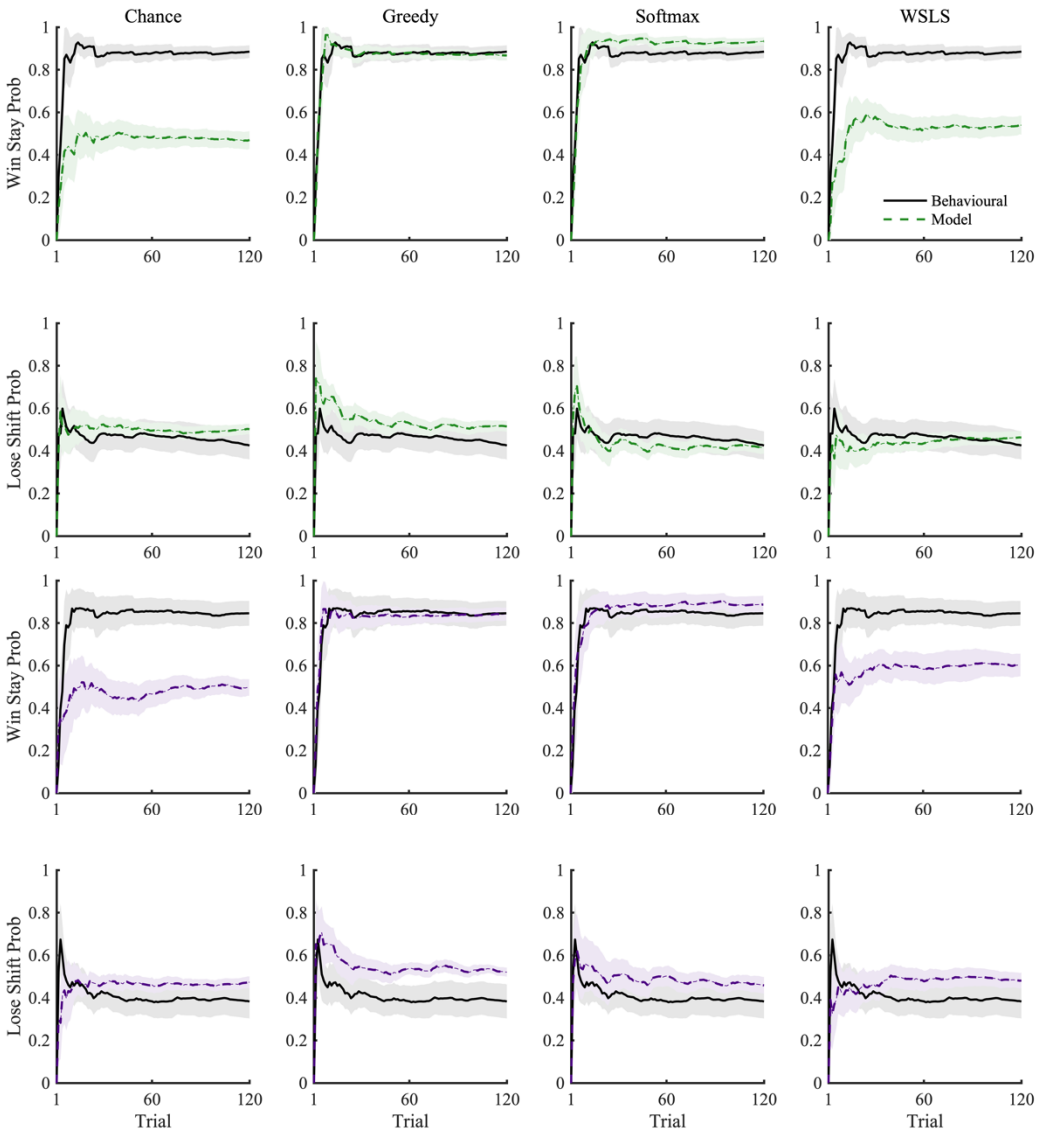

Figure 20. Experiment 3 - Trial-by-trial model performance for the early condition. The black lines indicate the average behaviour of participants while the dotted lines indicate the model averages for both win-stay probabilities and lose-shift probabilities across all one-hundred trials. Green lines are modelled control participants' averages and purple lines are modelled stress participants' averages. Error bars are 95% between-subject confidence intervals. Note: the black lines demonstrate identical data within each row of figures.

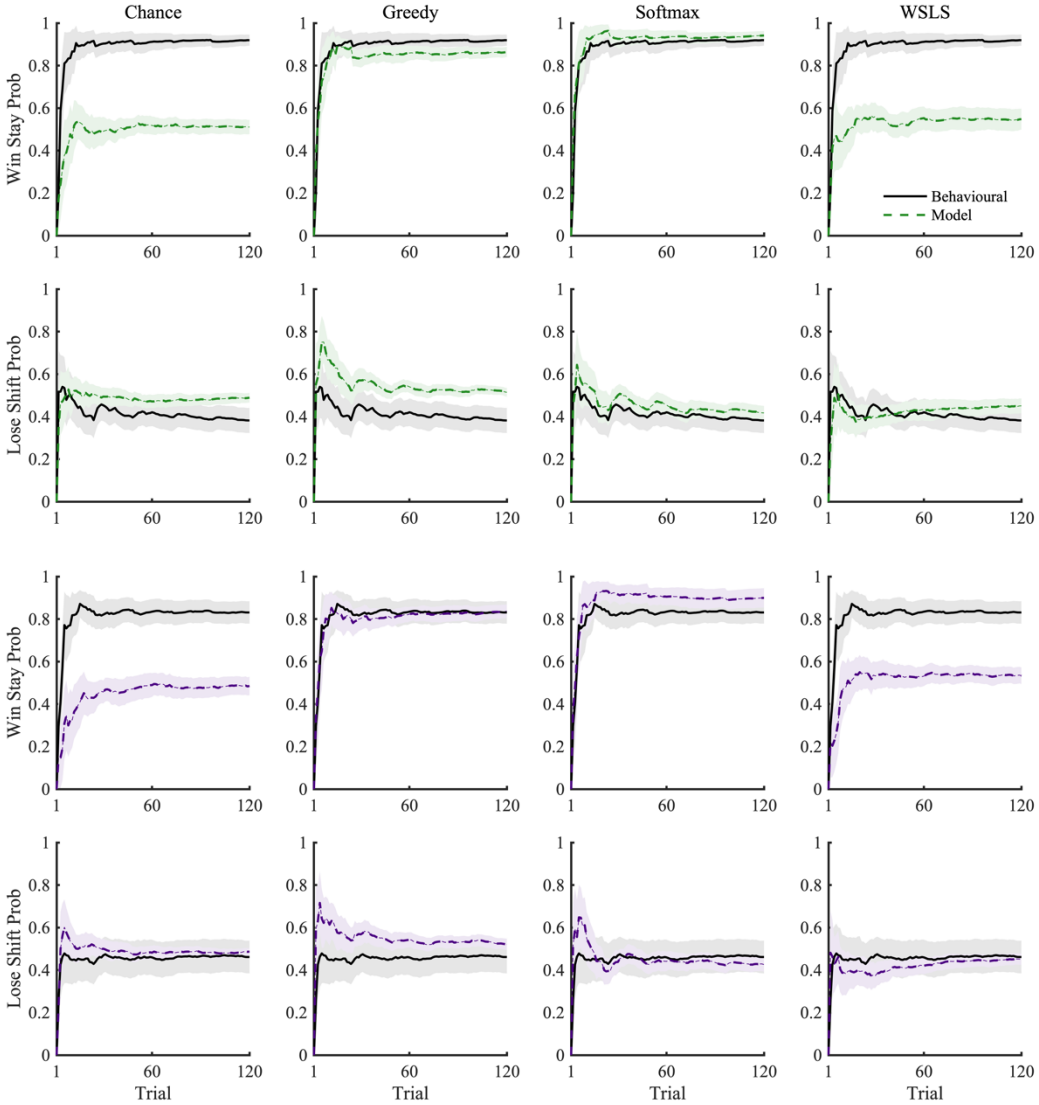

Figure 21. Experiment 3 - Trial-by-trial model performance for the late condition. The black lines indicate the average behaviour of participants while the dotted lines indicate the model averages for both win-stay probabilities and lose-shift probabilities across all one-hundred trials. Green lines are modelled control participants' averages and purple lines are modelled stress participants' averages. Error bars are 95% between-subject confidence intervals. Note: the black lines demonstrate identical data within each row of figures.

Across both the early and late condition, the Softmax model provided the best fit of behaviour (Table 10). The only condition-by-time group where the Softmax model failed to provide a fit greater than for 70% of participants was the Stress Early participants, where it was only the best model for 61% of the participants. Given the inability of the Win-Stay Lose-Shift model to account for behaviour and the lower best fitting classification rate of the ε-greedy and Win-Stay Lose-Shift models, only the Softmax model was used to classify trials as exploration and exploitation.

Generally however, the models were in relative agreement in terms of the overlap of which trials were exploration and exploitation (figure 22). While the Softmax and Greedy models had the highest overlaps of trials at 98%, the Win-Stay Lose-Shift model had lower overlap with the ε-greedy model (74%), and the Softmax mode (73%). In terms of exploration rate, the Win-Stay Lose-Shift model explored the most as 31.30% of trials were explorations, followed by the Softmax Model (17.21%) and then the ε-greedy model (17.88%).

Table 10

Experiment 3 - Best fitting models

	Early		Late	
Model	Control	Stress	Control	Stress
Greedy	19%	13%	10%	11%
Softmax	70%	61%	80%	78%
WSLS	11%	26%	10%	11%
∑	100%	100%	100%	100%

Note: WSLS = Win-Stay Lose-Shift Model. Columns all sum to 100%.

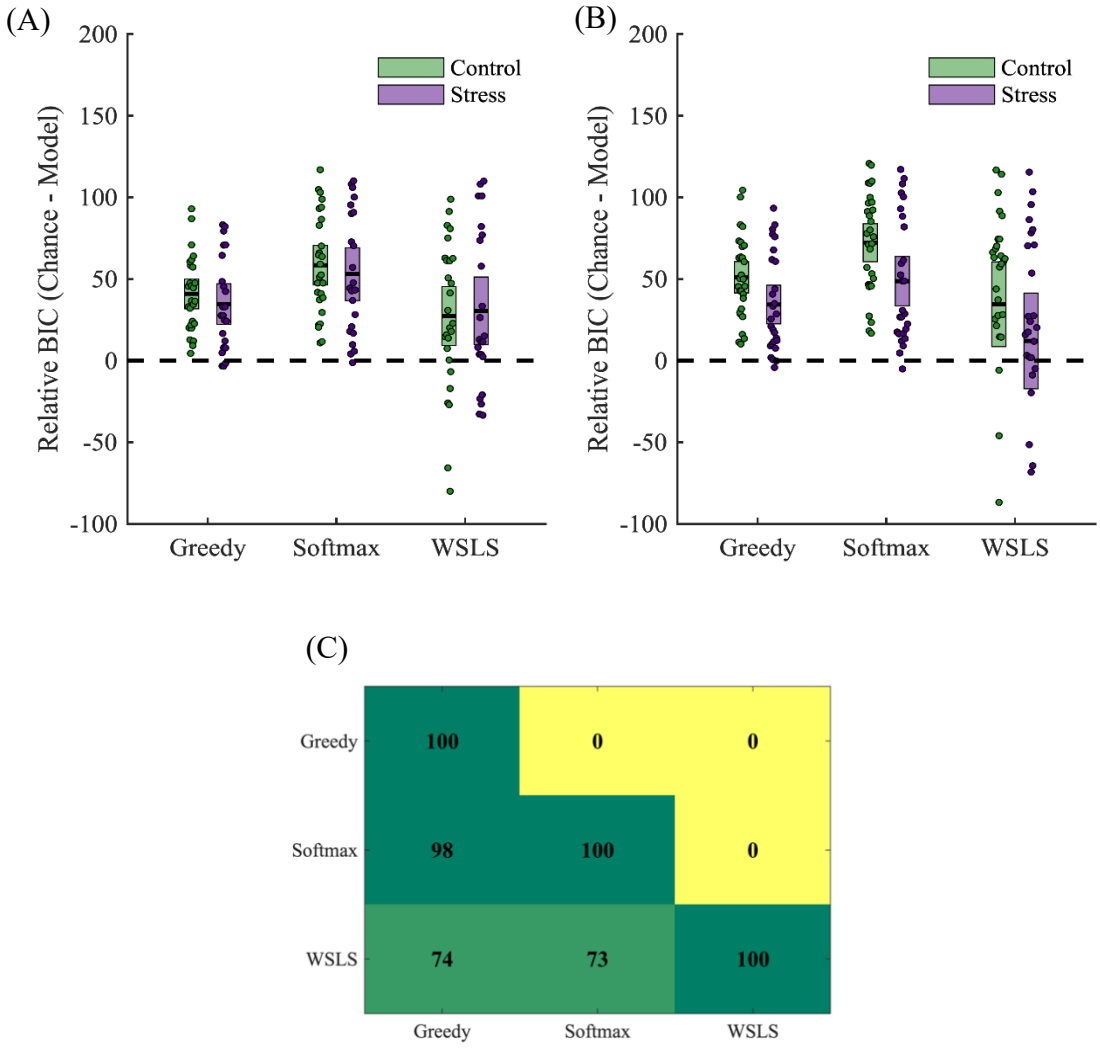

Figure 22. Experiment 3 - Model fit and trial classification.(A) Early BIC and (B) Late BIC. BIC indicates Bayesian Information Criterion from the chance model minus each of the other three models. Higher values indicate a better fit for the model. All models outperform the chance model. Boxes indicate 95% between-subject confidence intervals. (C) The total percentage of explore and exploit trial classification overlap for each non-Chance model.

Behavioural

In the oddball task (Table 11), when examining reaction time there were no effect of condition ($F(1, 102) = 2.15$, $p = .15$, $\eta_p^2 = .02$, $MSE = 4933.72$), time ($F(1, 102) = 0.38$, $p = .54$, $\eta_p^2 = .00$, $MSE = 4933.72$) nor was there an interaction ($F(1, 102) = 1.26$, $p = .27$, $\eta_p^2 = .01$, $MSE = 4933.72$). The same trend was evident for errors, as there was no effect of condition ($F(1, 102) = 0.05$, $p = .83$, $\eta_p^2 = .00$, $MSE = 0.80$), time ($F(1, 102) = 1.64$, $p = .20$, $\eta_p^2 = .02$, $MSE = 0.80$) nor interaction ($F(1, 102) = 2.89$, $p = .09$, $\eta_p^2 = .03$, $MSE = 0.80$).

Table 11

Experiment 3 - Oddball task performance

		Control		Stress	
		Mean	95% CI	Mean	95% CI
Reaction Time (ms)	Early	462.05	[430.98, 493.11]	457.30	[424.69, 489.90]
	Late	485.79	[458.63, 512.95]	450.37	[428.74, 471.99]
Errors	Early	0.26	[0.03, 0.50]	0.52	[-0.09, 1.12]
	Late	0.33	[0.04, 0.62]	0.00	[0.00, 0.00]

In the two-armed bandit there was an effect of the acute stress response across measures of performance (figure 23). There was an effect of condition on exploration ($F(1, 102) = 12.03$, $p < .001$, $\eta_p^2 = .1$, $MSE = 129.18$) but no effect of time ($F(1, 102) = 0.13$, $p = .64$, $\eta_p^2 = .00$, $MSE = 129.18$) and no interaction ($F(1, 102) = 0.41$, $p = .99$, $\eta_p^2 = .01$, $MSE = 129.18$). There was also an effect of condition on win percentage ($F(1, 102) = 7.95$, $p < .01$, $\eta_p^2 = .07$, $MSE = 63.16$) However, there was no effect of time ($F(1, 102) = 0.82$, $p = .37$, $\eta_p^2 = .01$, $MSE = 63.16$) and no

interaction ($F(1, 102) = 0.77$, $p = .38$, $\eta_p^2 = .01$, $MSE = 63.16$). As well, there was an effect of condition on optimal arm choice ($F(1, 102) = 8.57$, $p < .005$, $\eta_p^2 = .08$, $MSE = 129.18$) but no effect of time ($F(1, 102) = 1.26$, $p = .26$, $\eta_p^2 = .01$, $MSE = 129.18$) and no interaction ($F(1, 102) = 0.25$, $p = .62$, $\eta_p^2 = .00$, $MSE = 129.18$).

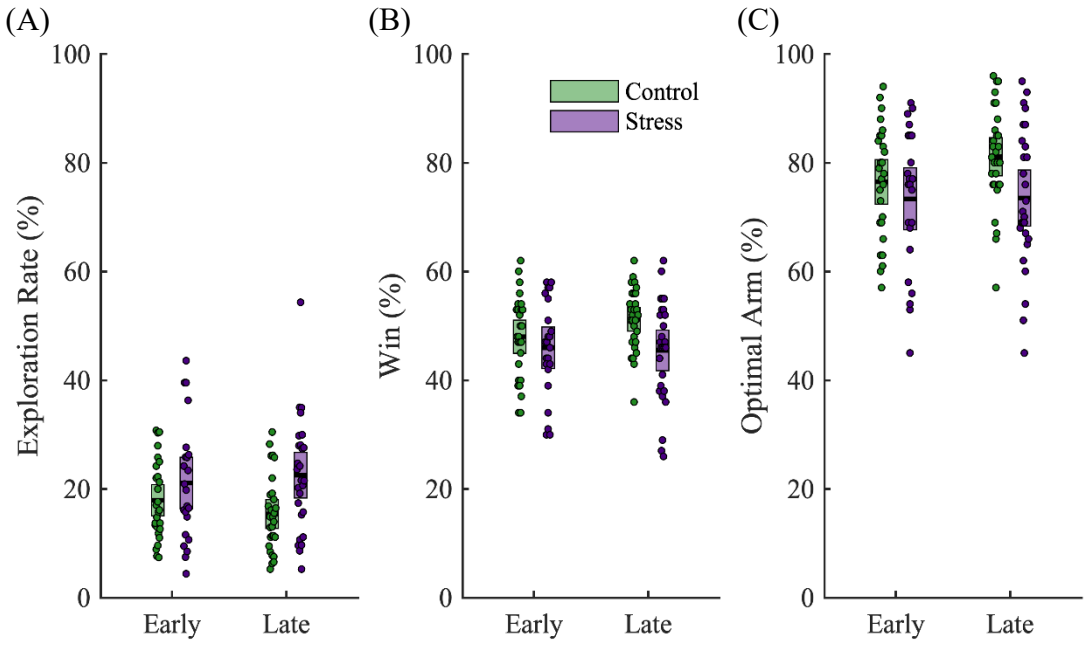

Figure 23. Experiment 3 – Two-armed bandit task performance.(A) Average exploration rate, (B) Win percentage and (C) Optimal Arm Percentage for the control and stress conditions grouped by early and late conditions. Boxes indicate 95% between-subject confidence intervals.

EEG

There was no effect of the acute stress response in the Oddball task (figure 24). Specifically, there was no effect of condition ($F(1, 102) = 2.93$, $p = .09$, $\eta_p^2 = .03$, $MSE = 13.42$), no effect of time ($F(1, 102) = 0.04$, $p = .85$, $\eta_p^2 = .00$, $MSE = 13.42$), and no interaction ($F(1, 102) = 0.09$, $p = .76$, $\eta_p^2 = .00$, $MSE = 13.42$). The additional check of the difference wave (infrequent minus frequent) showed that there was an oddball P300 in all conditions: control-early ($t(26) = 7.39$, $p < .001$, $d = 1.42$), stress-early ($t(22) = 6.78$, $p < .001$, $d = 1.41$), control-late ($t(28) = 8.83$, $p < .001$, $d = 1.64$), and stress-late ($t(26) = 6.81$, $p < .001$, $d = 1.31$).

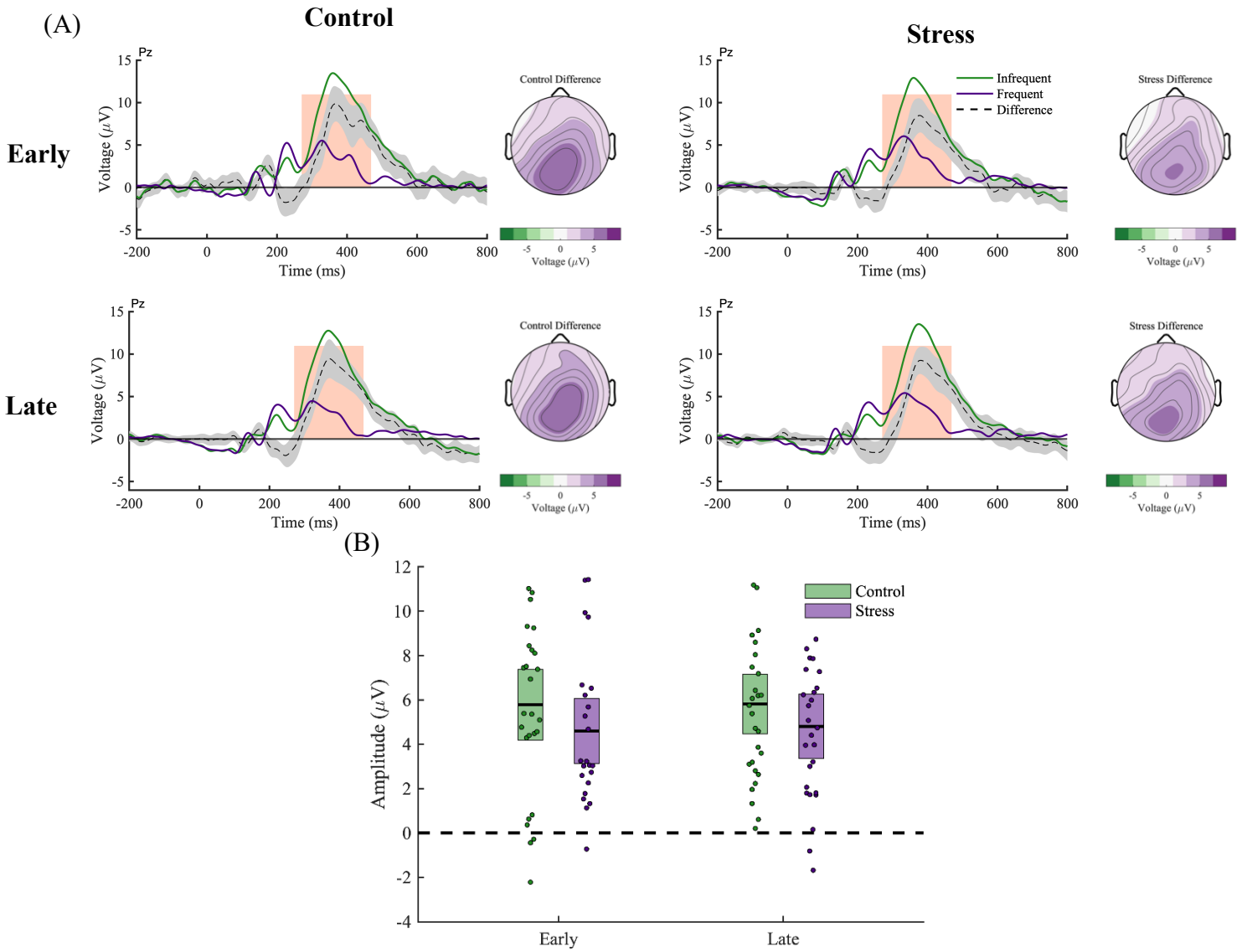

Figure 24. Experiment 3 – Oddball ERP and scalp topographies. (A)The shaded region indicates 95% between-subject confidence intervals. The pink box indicates the region of analysis for the P300. The scalp topographies indicate the difference (infrequent minus frequent trials) within the shaded region. (B) Difference wave averages for the region of analysis. Boxes indicate 95% between-subject confidence intervals.

In the two-armed bandit task, when examining the reward positivity to feedback (figure 25), there was no effect of condition ($F(1, 102) = 1.71$, $p = .19$, $\eta_p^2 = .02$, $MSE = 14.12$), no effect of time ($F(1, 102) = 0.17$, $p = .68$, $\eta_p^2 = .00$, $MSE = 14.12$), and no interaction ($F(1, 102) = 0.00$, $p = .96$, $\eta_p^2 = .00$, $MSE = 14.12$). The additional check of the difference wave (wins minus losses; figure 26) showed that wins were more positive than losses in all conditions. That is, the reward positivity was different from zero in the control-early condition ($t(26) = 3.08$, $p < .005$, $d = 0.59$), the stress-early condition ($t(22) = 4.04$, $p < .001$, $d = 0.92$), the control-late condition ($t(28) = 3.46$, $p < .005$, $d = 0.64$), and the stress-late condition ($t(26) = 4.61$, $p < .001$, $d = 0.90$).

When instead examining the P300 to feedback, there was no effect of condition ($F(1, 102) = 2.16$, $p = .15$, $\eta_p^2 = .02$, $MSE = 13.82$), no effect of time ($F(1, 102) = 3.42$, $p = .06$, $\eta_p^2 = .03$, $MSE = 13.82$), and no interaction ($F(1, 102) = 0.01$, $p = .92$ $\eta_p^2 = .00$, $MSE = 13.82$). The additional check of the difference wave did reveal that wins were more negative than losses for the feedback P300 in the control-early condition ($t(26) = 4.85$, $p < .001$, $d = 0.93$), the stress-early condition ($t(22) = 3.30$, $p < .005$, $d = 0.69$), and the control-late condition ($t(28) = 2.99$, $p < .01$, $d = 0.56$). However, there was no difference between wins and losses for the feedback P300 in the stress-late condition ($t(26) = 1.30$, $p = .21$, $d = 0.25$).

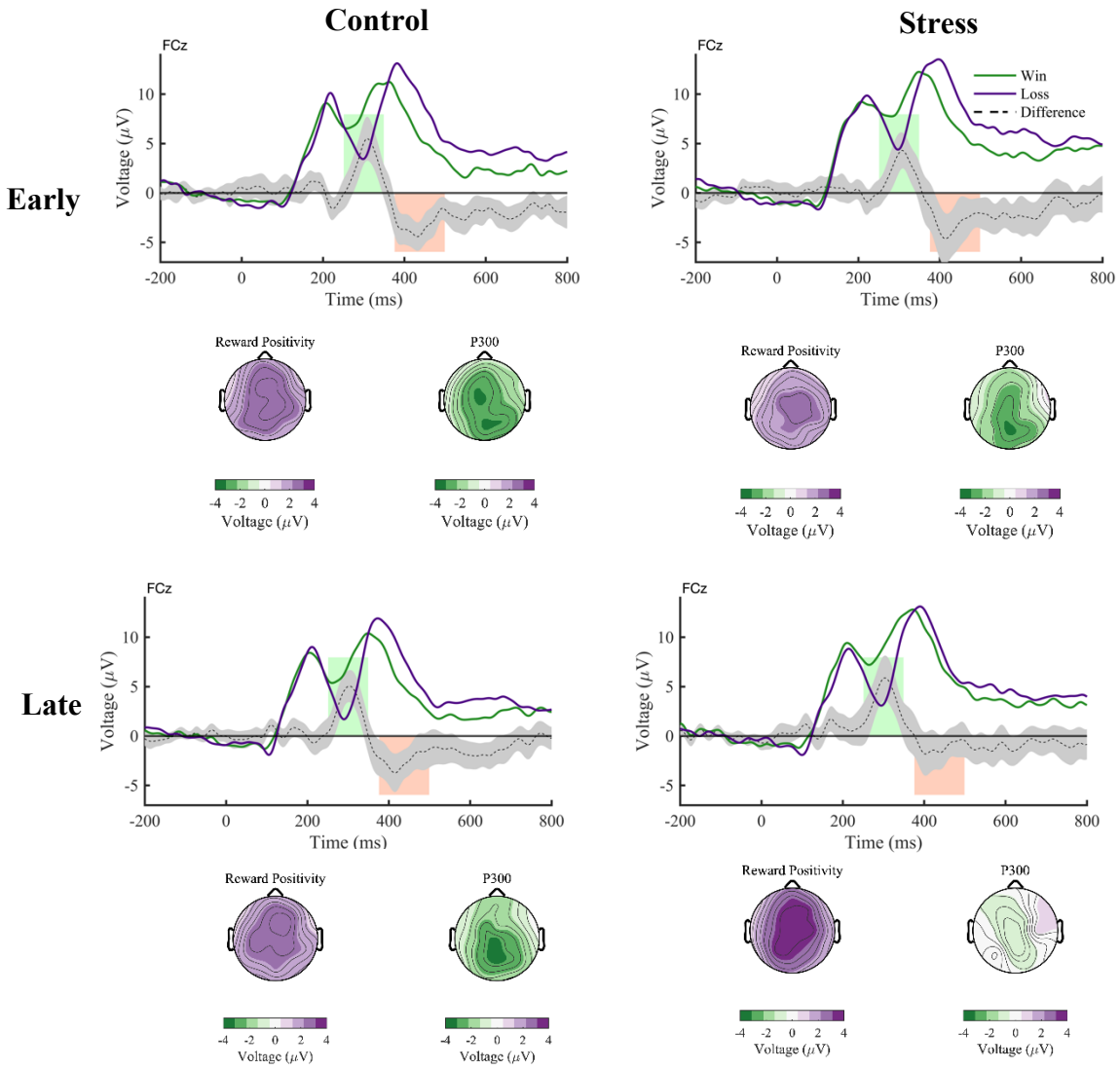

Figure 25. Experiment 3 – Feedback ERP and scalp topographies. The gray shaded region indicates 95% between-subject confidence intervals. The green box indicates the region of analysis for the reward positivity while the pink box indicates the region of analysis for the P300. The scalp topographies indicate the difference (win trials minus loss trials) within the regions of analysis.

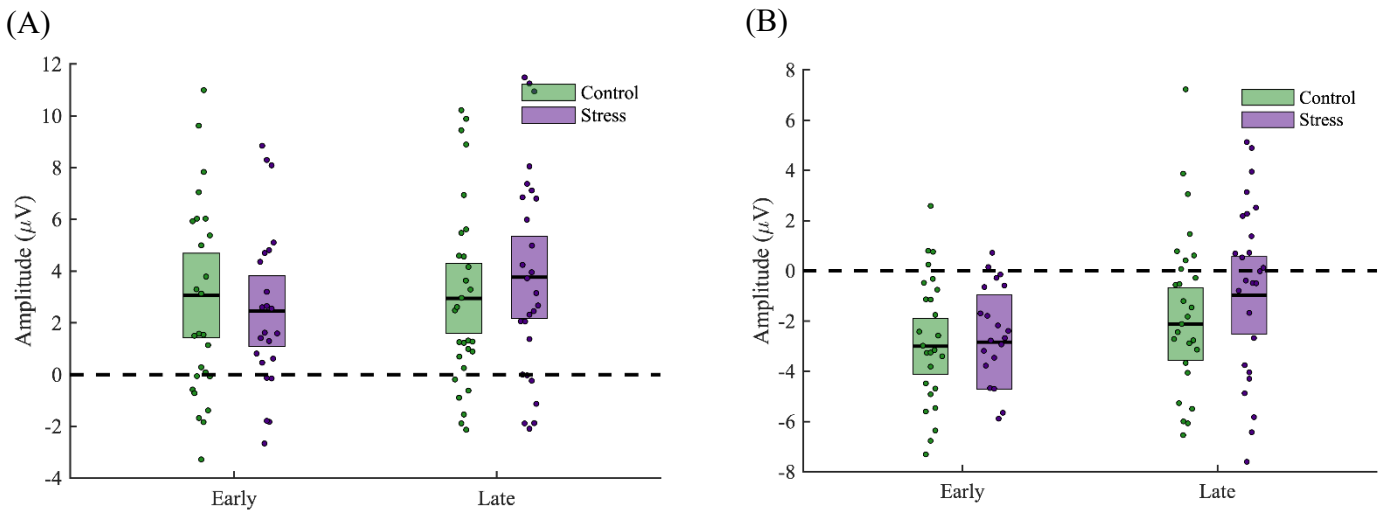

Figure 26. Experiment 3 – Feedback difference wave averages.(A) Reward positivity and (B) P300. Data show the average of the difference wave (win trials minus loss trials) for the regions of analysis at electrode FCz. Boxes indicate 95% between-subject confidence intervals.

When instead examining the P300 to exploration (figure 27), there was an effect of condition ($F(1, 102) = 8.07$, $p < .01$, $\eta_p^2 = .07$, $MSE = 22.04$), no effect of time ($F(1, 102) = 0.17$, $p = .68$, $\eta_p^2 = .00$, $MSE = 22.04$), and no interaction ($F(1, 102) = 0.10$, $p = .75$, $\eta_p^2 = .00$, $MSE = 22.04$). The additional check of the difference wave showed that both the control-early condition ($t(26) = 3.14$, $p < .005$, $d = 0.61$) and control-late condition ($t(28) = 3.34$, $p < .005$, $d = 0.62$) differed from zero. However, neither the stress-early condition ($t(22) = 1.55$, $p = .14$, $d = 0.32$) nor the stress-late condition ($t(26) = 0.38$, $p = .71$, $d = 0.07$) differed from zero.

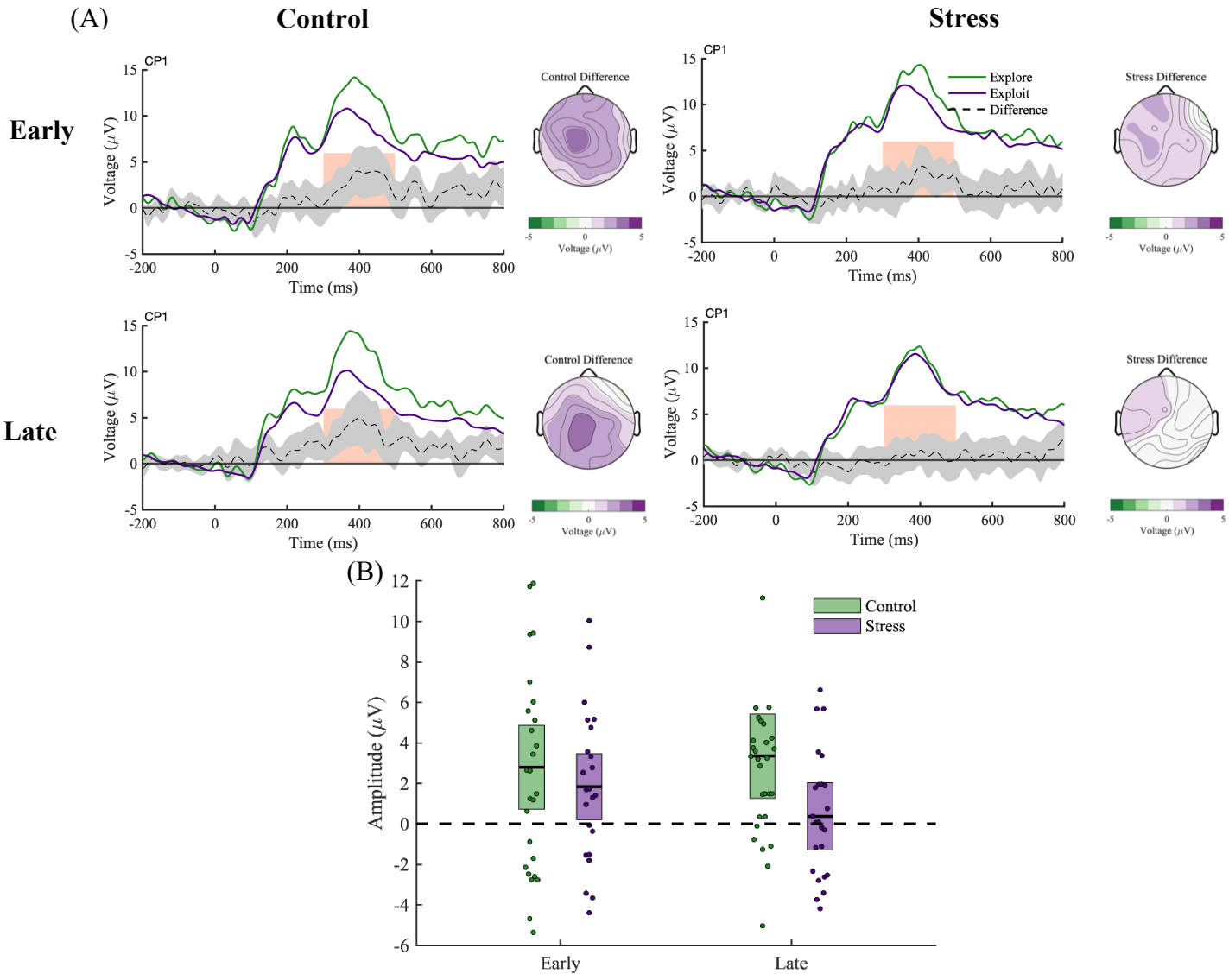

Figure 27. Experiment 3 – Exploration ERPs. (A) Control condition and Stress condition wave forms and topographic maps. The shaded region indicates 95% between-subject confidence intervals. The pink box indicates the region of analysis for the P300. The scalp topographies indicate the difference (exploration trials minus exploitation trials) within the region of analysis. (B) Difference wave average (exploration trials minus exploitation trials) for the region of analysis. Boxes indicate 95% between-subject confidence intervals.

Behavioural and neural correlations

To further tie the effects of the acute stress response on neural signals to modulations in behaviour, I correlated both exploration rate, win percentage and the two components where there was any evidence of a neural effect (the feedback P300; the exploration P300). Unsurprisingly, I found that a higher exploration rate was negatively correlated with a lower win percentage ($r(104) = -.71$, $p < .001$). There was a moderate negative correlation between win percentage the feedback P300 ($r(104) = -.33$, $p < .001$), suggesting participants who won more had a more negative (i.e., larger) feedback P300. There was also a moderate-to-large negative correlation between the exploration P300 and exploration rate ($r(104) = -.46$, $p < .001$), showing that participants who explored more have a lower exploration P300.

Discussion

In experiment 3, I found that the acute stress response not only modulated brain signals associated with feedback learning and uncertainty, but also caused decrements in performance and an increase in exploration rate. In terms of behaviour, I found that the acute stress response reduced win percentage and optimal arm choice, while increasing exploration rate. Much like in experiment 2, I used reinforcement learning models to model participants' behaviour. Both the ε-greedy and Softmax models were able to do a good job of replicating the behavioural effects – both in terms of average performance and trial-by-trial performance. Importantly, the models had a high degree of overlap with their classification of exploration and exploitation trials. I found no effect of the acute stress response on the oddball P300. However, I did observe tentative evidence that acute stress disrupted the feedback P300 while having no impact on the reward positivity. I found that the exploration P300 was disrupted by the acute stress response. In terms

of correlational data, I found that explorations were associated with a lower win percentage, while also showing that a lower exploration P300 was associated with more exploration and that a lower feedback P300 was associated with less wins – thus closely tying the neural effects to behaviour. I will now discuss the implications of these findings in turn.

While in both experiment 1 and 2 I found no effect of the acute stress response on feedback learning behaviour, I believe that this is because those experiments used contextual bandits where win probabilities shifted throughout. In the present experiment the bandit task was learnable – that is, there was a "best' option the participant had to learn to select – and I believe that this change to the task structure made it possible to observe the behavioural effect of the acute stress response. In fact, the finding that the acute stress response reduces bandit task performance supports previous work on this topic. Acute stress has been found to have differential effects on feedback learning – some studies have shown that the acute stress response impairs performance in feedback learning tasks (Paul et al., 2020; Preston et al., 2007) while other work has shown that the acute stress response may enhance learning from positive outcomes (i.e., learning from reward) while impairing learning from negative outcomes (i.e., learning from punishment; Lighthall et al., 2013). One possibility for why I observed a different finding from Lighthall and colleagues is that they only examined performance when no feedback was presented – that is, participants were given a chance to learn the feedback probabilities of different stimuli and had to reach an adequate level of performance, but their performance was assessed when no feedback was given (i.e., following training). As well, Lighthall and colleagues found that acute stress reduced sensitivity to feedback more generally during training. In the case of my experiment, participants continued to receive feedback for their choices throughout the task. In addition, in a study finding no effect of the acute stress response on feedback learning

(Glienke et al., 2015), the authors also examined performance following learning. Neuroimaging work has shown that the acute stress response reduces reward sensitivity in brain regions tied to feedback learning such as the striatum (in the case of the striatum acute stress reduced loss sensitivity as well) and the midcingulate cortex (Lincoln et al., 2019). My suggestion is that the acute stress response disrupted feedback learning signals and the ability of participants to use feedback information, which I argue represents a more top-down impact of the acute stress response on feedback learning, although perhaps this is only the case when feedback is presented.

I also found that the acute stress response increased the exploration rate of participants. In humans, prior work has shown that acute stress impairs the explore-exploit dilemma (Lenow et al., 2017). However, in that experiment, the authors found that participants tended to over-exploit while in my experiment I found participants over-explored. What might explain this difference? It seems more likely that the acute stress response does not lead to someone being overly exploitative or overly exploratory – that is the direction of the effect depends on the environment. In the patch foraging task by Lenow and colleagues, the participants over-exploitative behaviour was sub-optimal – that is they were staying too long at a patch and reducing their total reward. In the two-armed bandit task used here, participants were being overly exploratory to the detriment of behaviour. Thus, perhaps differences in the utility of exploration in the two paradigms might have caused the difference in effects observed. Relatedly, perhaps a clue can be gleaned from work showing that acute stress typically causes people to use cognitive strategies which require less effort (Bogdanov et al., 2021; Schwabe & Wolf, 2011, 2013). Perhaps in the present work it could be that participants in the stress condition are exploring more often due to not considering the task structure – for example, they

receive loss feedback on the bandit that wins more often and switch to the other bandit – when in fact, the participant should instead consider that the bandit that wins more often still will produce losses on 40% of trials. That participants in the stress condition selected the optimal arm less gives credence to this suggestion.

In the present work, I found tentative evidence that the acute stress response disrupted the feedback P300 in the stress-late condition but had no impact on the reward positivity. The stress-related disruption of the feedback P300 replicates what I saw in experiment 2 and mirrors prior work showing that the P300 is disrupted following acute stress (Banis & Lorist, 2012; Paul et al., 2020). I believe that this disruption of the feedback P300 suggests a modulation of the ability of participants to use feedback in a top-down manner to guide behaviour (i.e., a more downstream part of the feedback signal), as Wu & Zhou (2009) have argued that the feedback P300 is sensitive to factors such as magnitude, valence, and expectancy while the reward positivity might solely reflect outcome evaluation. Importantly, the amplitude of the feedback P300 was negatively correlated with the win percentage of participants. Thus, given that participants in the stress-late condition had a lower win percentage, it seems that acute stress disrupted the feedback P300 in turn leading to worse performance.

However, that I found no effect of the acute stress response on the reward positivity doesn't replicate experiment 1 and does not support prior work showing that the reward positivity is decreased following acute stress (Banis & Lorist, 2012; Burani et al., 2021; Paul et al., 2020; X. Zhang et al., 2020). While it should be noted that some work has observed an increase in the reward positivity under acute stress (Glienke et al., 2015), the lack of consistency in my findings with the literature is concerning. With regards to previous work finding that acute stress decreased the reward positivity, it should be noted that three of the above papers examined

participants used a within-subject design and did not use psychological stressors (Banis & Lorist, 2012; Burani et al., 2021; X. Zhang et al., 2020) which might not elicit robust cortisol responses (Dickerson & Kemeny, 2004). Thus, it's possible that a lack of between-group experiment power might have reduced my ability to detect an effect of the acute stress response on the reward positivity. Alternatively, these differential effects from the literature could be due to design differences such as my use of a psychological stressor in experiment 3.

Relatedly, the lack of power to detect small or medium effects might explain why there was an effect of the acute stress response on the reward positivity in experiment 1 but no effect in the present experiment. However, there were some minor methodological differences which could also have contributed to the effect not replicating. One methodological difference between the tasks was that in experiment 1 the task used was a contextual bandit while in experiment 3 the task used was a learnable bandit. Thus another possibility is that the effect of the acute stress response on the reward positivity in the contextual bandit (experiment 1) could be related to work showing that acute stress causes participants to be vulnerable to fatigue effects leading to task disengagement (e.g., Shafiei et al., 2012). Specifically, the two-armed bandit task in experiment 3 was completed quickly as it took five to seven minutes on average to complete. In contrast, the four-armed bandit took between fifteen to twenty minutes to complete in experiment 1. That is, the effects on the reward positivity are only observable in longer time-frame tasks. Thus, the lack of effect of the reward positivity in experiment 3 (in contrast to experiment 1) that I observed here could possibly be related to the difference in terms of the length of the tasks. Future work could thus consider examining changes to the reward positivity on a block-by-block timeframe while employing longer tasks.

Importantly, in the present work I found that the acute stress response disrupted the exploration P300, perhaps leading to worse performance and greater exploration. This suggests that the acute stress response modulated uncertainty in both stress conditions. These findings replicate the tenuous evidence from experiment 1 – although it does not replicate experiment 2. I believe that the lack of replication could be due to low power to detect an effect of the acute stress response. In fact, it could be that the acute stress response increases the variability of the uncertainty signal leading to a decrease in power to detect an effect of trial type. I also found that the exploration P300 and the exploration rate were negatively correlated. In other words, participants who explored more had a lower exploration P300 – which tied behaviour to the underlying neurophysiology. Thus, perhaps following acute stress – where participants have a disrupted uncertainty signal – their assessment of task uncertainty does not function as expected (as occurs in substance abuse – J. C. Yu et al., 2020) and participants subsequently explore more often than they should. That is, participants under acute stress are not experiencing differences in uncertainty between explorations and exploitations and are subsequently failing to update their task model or worldview as one would expect (Dayan & Yu, 2006), perhaps related to modulations to norepinephrine under acute stress. The disruption of the exploration P300 provides one possible neural mechanism for why the acute stress response modulates the explore-exploit dilemma, although why this occurs remains unclear. In addition, given that exploration is associated with worse performance, I believe that the disruption of the uncertainty signal, and the change in exploration and task performance are closely linked. That is, participants who have been exposed to an acute stressor have disrupted uncertainty signals, and that the disruption of the uncertainty signal impairs the ability of participants to manage the explore-exploit dilemma and to perform the task. Moreover, the fact that the oddball P300 was

not impacted by the acute stress response suggests that the exploration P300 likely indexes a distinct cognitive process from the oddball P300, a finding I will return to in the general discussion.

One of the goals of the present experiment was to determine whether the assessment of cognitive performance relative to the stressor had any interaction. In fact, time of assessment might not have an impact on how the acute stress response impacts the explore-exploit dilemma. However, why might there be an effect of the acute stress response on the uncertainty signal at both early and late time assessments? My suggestion is related to both SAM and HPA axis activity acting on the uncertainty signal. Specifically, the SAM axis primarily impacts norepinephrine and phasic NE plays a key role in the uncertainty signal (Hassall, 2019). As well, there is evidence that the HPA axis also impacts norepinephrine through the activity of cortisol. Thus, it seems that both the arousal related changes caused by the SAM axis and the delayed activity of cortisol under the HPA axis both impact the uncertainty signal.

Future work should consider better separating out the immediate and delayed effects of the acute stress response. For example, the collection of salivary alpha-amylase would provide a more direct measure of SAM axis activity. The better understanding of the interaction of time of assessment and acute stress could have implications in the literature. For example, some studies have adopted designs where acute stress is repeatedly induced throughout a task with noise stressors (e.g., Banis & Lorist, 2012). Given that there is little evidence that noise stress consistently activates the HPA axis and correspondingly elevates cortisol (Dickerson & Kemeny, 2004) and given the lack of any delay following the stressor used in both Banis et al., (2014) & Banis & Lorist, (2012) – it could be argued that these studies are examining the impact of the SAM axis on cognition. More generally, it is evident that greater power is required to fully

investigate the differences in SAM and HPA axis activity in the explore-exploit dilemma – which the use of within subject designs could accomplish.

In the present experiment, the manipulation check data revealed an effect of time on both state anxiety and the heart rate data. As well, there was even an interaction between time and stress condition for state anxiety. Given that the data are taken from experiment 1 and 2, it is worth considering if there were any differences in how the experiments were conducted. In fact, one unique aspect of experiment 2 was that the experiment was conducted during the COVID-19 pandemic while in-door mask mandates were in-place. As such, members of the Trier Social Stress Test panel had masks on during the interview – which could explain why the stress conditions in experiment 2 had greater anxiety related changes. However, what might explain the effect of time on heart rate, where there was no interaction (early participants simply had lower changes in baseline)? One possibility is that during the experiment 2 participants were informed that cortisol would be collected immediately prior to the placebo or stressor task. The collection of cortisol might also elicit some level of nervousness in participants simply from the collection procedure or perhaps because cortisol was collected from them in a biosafety level 2 certified room (which included several biohazard signs with large warnings on them).

In sum, the present work provides evidence that the acute stress response modulates uncertainty signals in the brain, may weakly modify feedback signals, and impacts task performance and exploration rate. This in turn provides a neural mechanism for how the acute stress response impacts the explore-exploit dilemma while also highlighting the importance of both uncertainty and feedback signals to when participants decide to explore. Disruption of these signals due to the acute stress response leads to less adaptive behaviour. In experiment 4, I now change my emphasis from acute stress to chronic stress. While chronic stress has not been

considered in the previous experiments, Lenow and Colleagues (2017) found that chronic stress does modulate the explore-exploit dilemma by reducing exploration rate. As such, to further investigate this finding, I turn now to a consideration of the impact of chronic stress on the explore-exploit dilemma.

Chapter 5: Experiment 4

Introduction

Chronic stress is functionally distinct from acute stress and impacts both SAM axis and HPA axis reactivity. Unlike acute stress which can be adaptive in some scenarios and allows for the re-distribution of cognitive resources to deal with challenges, the effects of chronic stress appear to be purely negative. For example, there are many links between chronic stress and disease, including heart-disease (Vitaliano et al., 2002), obesity (Dallman et al., 2003), and cancer (Reiche et al., 2004). Given that acute stress requires harnessing resources to deal with a threat, chronic mobilization causes pathology due to the system being overly taxed (McEwen, 2007). A broad definition of chronic stress is a stressful event/events (i.e., a threat to allostasis) that occurs repeatedly over the course of days to years and has long lasting (genomic) changes to the brain and body. In human research, typically questionnaires such as the perceived stress scale (S. Cohen et al., 1983) or the Trier Inventory for the Assessment of Chronic Stress (Schulz & Scholtz, 1999) are used to measure chronic stress levels.

Work on the neurobiological effects of chronic stress suggest it leads to decreases in the density of pre-frontal cortex dendrites while conversely leading to growth of dendrites in the amygdala (McEwen, 2007). More generally, it seems that chronic stress changes broad parts of the pre-frontal cortex (Arnsten, 2009; McEwen & Gianaros, 2011). Chronic stress seems to lead to increases in circulating glucocorticoids and thus modulates HPA axis activity (Schulz, Kirschbaum, Prüßner, & Hellhammer, 1998; Wüst, Federenko, Hellhammer, & Kirschbaum, 2000). However, there is also evidence that chronic stress impacts the SAM Axis as well, as

chronic stress was found to be associated with higher levels of salivary alpha-amylase (Nater et al., 2007; Vineetha et al., 2014).

One key factor in the stress response is the emotional impact of stress. While stress elicits a number of different emotions (Lazarus, 2006), a commonly elicited emotion is an increase in feelings of anxiety and the two constructs are closely linked (Fink, 2016). This makes some amount of sense given that stress is a response to a threat (Lupien, McEwen, Gunnar, & Heim, 2009) and anxiety itself reflects heightened arousal to some perceived threat (Lieb, 2005). In fact, stress increases feelings of anxiety immediately following a stressor while these feelings return to baseline ten to fifteen minutes post-stress (van Gerven et al., 2016). As well, both acute stress and anxiety overlap in the neural circuits they impact, as both acute stress and anxiety act on the basolateral amygdala and pre-frontal cortex (Daviu et al., 2019). Moreover, Daviu and colleagues demonstrated that acute stress increases feelings of anxiety through activity of the locus-coeruleus norepinephrine system. To put it another way, acute stress increases anxiety immediately following a stressor, which appears to be closely tied to activity of the SAM axis and norepinephrine and demonstrates stress reactivity. Thus, when considering the effect of acute stress, it might then be important to also consider separately how anxiety itself impacts the explore-exploit dilemma.

Given the close links between anxiety and stress, it is unsurprising that anxiety also impacts many of the same brain regions that stress does. Generally, it is suggested that anxiety is closely associated with increases in arousal and activity in attentional networks in the brain such as the fronto-parietal network (Saviola et al., 2020) and the amygdala (Bishop, Duncan, Brett, et al., 2004). This makes some amount of sense when one considers the behavioural impact of anxiety such as increased responses to fearful stimuli (Bishop, Duncan, & Lawrence, 2004), an

increase in responses to threatening stimuli (Macleod & Mathews, 1988), and a preference for choices with less uncertainty (Starcevic & Berle, 2006). Generally, high anxiety is associated with reductions in activity in the pre-frontal cortex which in turn signals deficiencies in the ability to implement cognitive control (Bishop, 2009). Importantly, high levels of anxiety appear to be closely associated with activity of the SAM axis, and thus more closely associated with norepinephrine (Redmond Jr. & Huang, 1979). More specifically, anxiety has been shown to be closely related to locus-coeruleus norepinephrine activity, as the locus-coeruleus norepinephrine system innervates brain regions that are tied to anxiety and anxiogenic behaviour such as the amygdala and the pre-frontal cortex (Sara, 2009). Moreover, increases in norepinephrine are associated with an increase in feelings of fear related to the stress response (Morilak et al., 2005).

The primary goal with experiment 4 was to investigate how chronic stress and anxiety impact common electrophysiological measures of decision making extracted from a feedback learning task (a two-armed bandit task) and a context updating task (a visual Oddball task). Chronic stress shares many of the same neural targets as the acute stress response, most importantly the pre-frontal cortex (Cook & Wellman, 2004). As well, I should note here that when I refer to chronic stress, I am generally referring to a chronic exposure to stressors and a history of those exposures. However, it is important to note that the Trier Inventory for the Assessment of Chronic Stress also may index the stress response as well. However, for clarity and because I used the summary scale of the Trier Inventory for the Assessment of Chronic Stress, throughout when discussing the impact of chronic stress, I will refer to the impact of chronic stress exposure rather than the stress response (which was the focus of the previous three experiments). As well, anxiety (specifically, state anxiety) is an important emotional response to stress as stress leads to an increase in feelings of anxiety (Grillon et al., 2007). Thus, in the

present experiment I investigated how anxiety impacted both the reward positivity and the P300, and by extension, whether changes to these components under stress might reflect changes to explore-exploit decision making. I should note that the anxiety questionnaire I used (the State-Trait Anxiety Inventory – Spielberger, 2010) allowed for the separation of state anxiety (i.e., how anxious the participant is in the moment) and trait anxiety (i.e., how anxious the participant generally feels).

Previous investigations of the effect of anxiety and chronic stress on the reward positivity and P300 have mostly been focused on the effect of anxiety on the P300. More specifically, increases in state anxiety led to a decrease in P300 amplitude (Rossi & Pourtois, 2017), which the authors suggested reflected impairments in goal-directed processing. Relatedly, high trait anxiety has been shown to decrease P300 amplitudes when compared to people with low trait anxiety in an Oddball task (Huang et al., 2009). These two findings in conjunction suggest that anxiety (whether state or trait) reduces P300 amplitude – although how these neural changes impact behaviour is unclear and merits further investigation. In addition, trait anxiety has also been shown to reduce reward positivity amplitude (Takács et al., 2015) which the authors suggest reflects changes to how people value outcomes due to anxiety. Prior has examined the impact of chronic stress exposures on the reward positivity (e.g., Ethridge et al., 2018; Freeman et al., 2022; Suor et al., 2021; or see Kujawa et al., 2020 for a review). Typically then, chronic stress exposures seem to impair reward processing and sensitivity, and these investigations have examined a variety of forms of chronic stress exposures such as relational victimization (Ethridge et al., 2018), childhood maltreatment (Suor et al., 2021), and a maternal history of depression (Freeman et al., 2022). As well, there has been work attempting to use machine learning to classify chronic stress states using EEG (Baumgartl et al., 2020; Peng et al., 2013).

However, no studies have investigated the impact of the Trier Inventory for the Assessment for Chronic Stress and whether scores on this survey have any relationship to changes to the reward positivity. As well, there have been no investigations no investigations of the effect of state anxiety on the reward positivity – suggesting this area is ripe for investigation.

The second goal of this project was to conduct a computational modelling investigation on participants' behaviour in the two-armed bandit task. To do this, I applied a reinforcement learning model to participant's behavioural data from the two-armed bandit (R. S. Sutton & Barto, 2018). This model allowed for the classification of trials as either explorations or exploitations. I adopted this approach because it could provide a direct investigation of whether explore-exploit decision making is modulated by any of these measures (chronic stress, state anxiety, and trait anxiety).

My expectations for each of the two goals were as follows. For goal one, as per previous work, I expected that both trait and state anxiety would be negatively correlated with P300 amplitude (Huang et al., 2009; Rossi & Pourtois, 2017), and that trait anxiety would also be negatively correlated with reward positivity amplitude (Takács et al., 2015). Moreover, given work highlighting that higher state anxiety impairs reward learning (Hein et al., 2021) – I expected higher state anxiety to correlate with lower reward positivity amplitudes. Given research demonstrating a relationship between chronic stress exposures and attenuated reward processing signals (Ethridge et al., 2018; Freeman et al., 2022; Kujawa et al., 2020; Suor et al., 2021) and given that chronic stress exposures inhibit reward related dopaminergic activity (Cabib & Puglisi-Allegra, 2012), I expected that chronic stress would reduce the amplitude of the reward positivity. As an exploratory analysis I also examined the feedback P300 and the exploration P300 and the impact of anxiety and chronic stress exposure on both measures.

Relatedly, given that chronic stress exposure changes attentional abilities and dendritic morphology in rats (Liston et al., 2006) and that the oddball P300 is closely linked to attentional processes and working memory (Polich, 2007) I also expected that the oddball P300 amplitude to be reduced by chronic stress. For goal two, given that previous work has shown that chronic stress leads to an over-exploitative strategy (Lenow et al., 2017) and that certain forms of trait anxiety may reduce exploration (Fan et al., 2021), I would expect that both trait anxiety and chronic stress would reduce exploration rate. I had no strong prediction about how state anxiety would impact exploration rate.

Method

Participants

For this study, I recruited a large sample of university students (n = 309; 201 females, 104 males, 4 chose not to declare; age range = 17 to 46 years, mean age = 21.42, 95% CI [20.93, 21.92]) and measured participants' EEG after they completed a series of questionnaires. In order to assess chronic stress and anxiety, I had participants complete two questionnaires: (1) the Trier Inventory for the Assessment of Chronic Stress (TICS; Becker, Schulz, & Schlotz, 2004), and (2) the State Trait Anxiety Inventory (STAI; Spielberger, 2010). Participants also had to complete an Oddball task and a two-armed bandit task. The order of the two tasks was counterbalanced. An important note is that participants were not explicitly recruited for this study, the data collected here was the result of participants who came into lab for other studies who completed these two tasks. All participants completed these two tasks after finishing the questionnaires but before any other cognitive tasks.

Apparatus and Procedure

Questionnaires. I had participants complete the TICS questionnaire to measure their chronic stress levels. Specifically, the Trier Inventory for the Assessment of Chronic Stress measures stress exposures and stress responses in the past three months across 7 different sub-scales related to chronic stress (Becker et al., 2004). For my purposes, I solely examined the chronic stress summary score, which included 12 questions. Scores on the chronic stress summary score can vary from 0 (very low chronic stress) to 48 (very high chronic stress), although in my sample the scores ranged from 0 to 44. I also had participants complete the State Trait Anxiety Inventory (Spielberger, 2010) in order to measure their anxiety levels. In this case, the STAI measures both state (STAI-S, acute) and trait (STAI-T, long-term) anxiety levels. Higher scores on each sub-scale indicate higher levels of either state and trait anxiety, while lower scores indicate lower levels of either state or trait anxiety. Scores on the State Trait Anxiety Inventory – State scale and State Trait Anxiety Inventory – Trait scale range from 20 to 80 for both scales – however, in my sample I observed a range of scores between 20 to 66 for the State Trait Anxiety Inventory – State scale and between 20 to 71 for the State Trait Anxiety Inventory – Trait scale.

Participants completed identical tasks to experiment 3. That is, they completed the Oddball task and the two-armed bandit task (see figure 28).

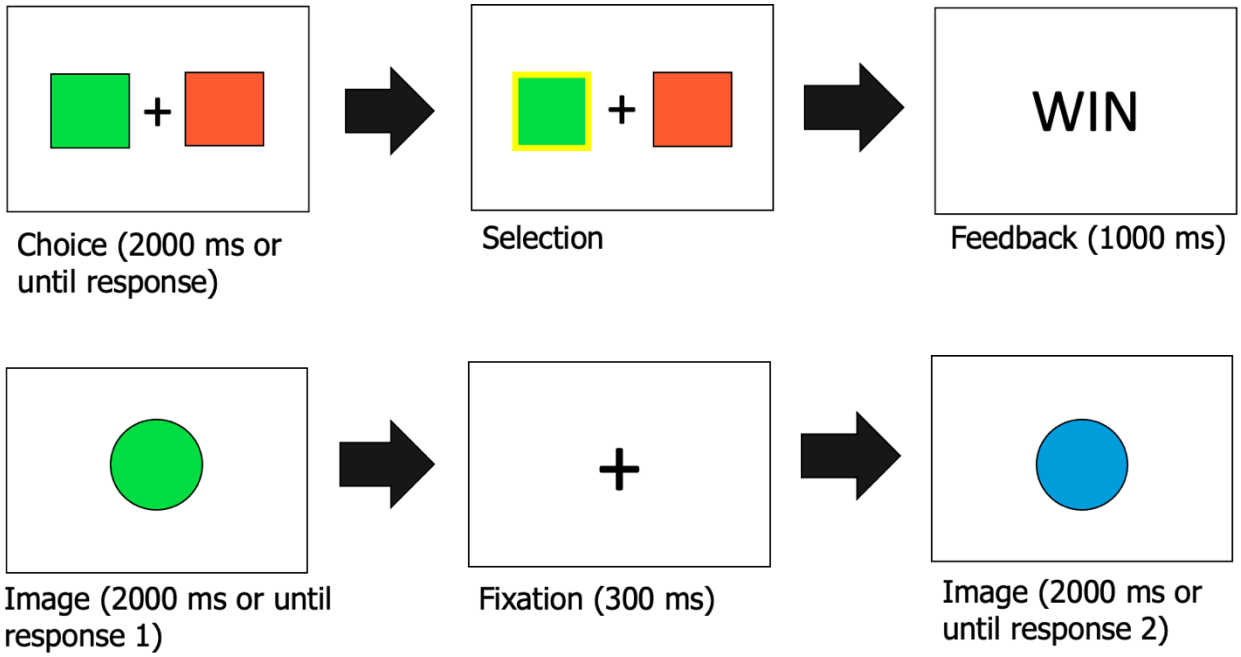

Figure 28. Experiment 4 - Tasks (A) Two-armed bandit paradigm and (B) The visual oddball paradigm. Note figure is identical to figure 17.

Computational Model

The second goal of this project was to determine whether any effects were apparent in the data with regards to the exploration rate in the two-armed bandit task. As such, I used a reinforcement learning model to extract out explore-exploit behaviour. Importantly, given the results from Experiment 3, I did not attempt to validate across multiple models in this experiment. That is, because the Softmax model provided the best fit for the majority of the participants in experiment 3, and because I used an identical task, I only examined the Softmax model in the present experiment. All model parameter optimization occurred identically to experiment 3.

Data Processing

Data were processed identically to the previous experiments. The final data had an artifact rejection rate of 17.60%, 95% CI = [16.36, 18.84] of trials in the Oddball task, and 21.92%, 95% CI = [20.61, 23.23] of trials in the two-armed bandit task.

Data Analysis

Behaviour. In the Oddball task, I examined the number of errors participants made and their reaction time to make a keypress. In the two-armed bandit task, I examined both win-percentage (total wins divided by total valid trials) and exploration percentage (explore trials divided by the sum of explore and exploit trials). I then correlated each measure of performance with the three questionnaire scores. For each set of correlations within each survey and the behavioural measures, I used a false-discovery rate correction for the *p* values (Benjamini-Hochberg, 1995).

ERP. First, as a manipulation check, the presence of the oddball P300, the reward positivity, the feedback P300, and the exploration P300 were assessed using a difference wave comparison. For all components I used the identical electrodes and windows as per experiment 3. To verify the existence of each component, I used the mean difference of the conditional waveforms, confidence intervals, and Cohen's *d*, and tested each difference from zero using a one-sample *t*-test.

Questionnaires. The primary goal of this experiment was to investigate the relationship between scores on the Trier Inventory for the Assessment of Chronic Stress and the State Trait Anxiety Inventory, and four ERP components: the oddball P300, the reward positivity, the feedback P300, and the exploration P300. For the participants' scores across the two questionnaires (their score on the Trier Inventory for the Assessment of Chronic Stress, their

score on the State Trait Anxiety Inventory – State scale, and their score the State Trait Anxiety Inventory – Trait scale), I computed the Pearson correlation between these measures and the three ERP components. For each set of correlations within each survey and the ERP components, I used a false-discovery rate correction for the *p* values (Benjamini & Hochberg, 1995).

Results

Behaviour

The only correlations between behaviour and any of the questionnaires were between anxiety and measures from the Oddball task (Table 12). Specifically, there was a small correlation between the number of errors made in the Oddball task and state anxiety ($r(307) = .18$, $t = 2.51$, $p < .05$). It should be noted, however, that participants made very few errors in the oddball task, on average less than one per participant. In addition, there as a small correlation between state anxiety and reaction time ($r(307) = .14$, $t = 3.18$ $p < .01$). While there was a small correlation between trait anxiety and errors made in the Oddball task, it did not survive FDR correction. All other correlations were negligible (all $| r | < .10$, $p > .20$).

Table 12

Experiment 4 - Behavioural Performance

		Chronic Stress		Trait Anxiety		State Anxiety	
		Pearson *r*	95% CI	Pearson *r*	95% CI	Pearson *r*	95% CI
Oddball	Errors	.07	[-.04, .18]	.13	[.02, .24]	.18*	[.07, .28]
	Reaction Time	-.01	[-.11, .10]	.00	[-.10, .12]	.14*	[.03, .25]
Bandit	Win Percentage	-.01	[-.12, .11]	.00	[-.11, .11]	-.03	[-.13, .09]
	Exploration Percentage	-.01	[-.12, .10]	-.03	[-.14, .08]	.00	[-.11, .11]

Note: * indicates correlations which survived FDR adjustment to significance testing

ERP Manipulation Checks

As expected, for both tasks I observed clear ERPs. That is, for the oddball task, I found that the infrequent waveform was larger than the frequent waveform in the region of analysis, and the difference wave ($\bar{X}$ = 6.31 μV, 95% CI [5.89, 6.72]) was different from zero ($t(308)$ = 29.87, $p < .001$, d = 1.69). In the two-armed bandit, I observed a clear waveform and topographic map indicative of the reward positivity (see figure 29) – the win waveform was more positive than the loss waveform within 200 to 400 ms post-feedback (win minus loss; $\bar{X}$ = 3.14 μV, 95% CI [2.68, 3.59]) was different from zero ($t(308)$ = 14.36, $p < .001$, d = 0.82). The difference for the feedback P300 was also different from zero (win minus loss: $\bar{X}$ = -1.91 μV, 95% CI [-2.40, -1.44]; ($t(308)$ = 7.84, $p < .001$, d = -0.45). For the exploration P300 in the two-armed bandit task, the difference wave (explore minus exploit; $\bar{X}$ = 1.53 μV, 95% CI [1.00, 2.05]) was different from zero ($t(308)$ = 5.71, $p < .001$, d = .33).

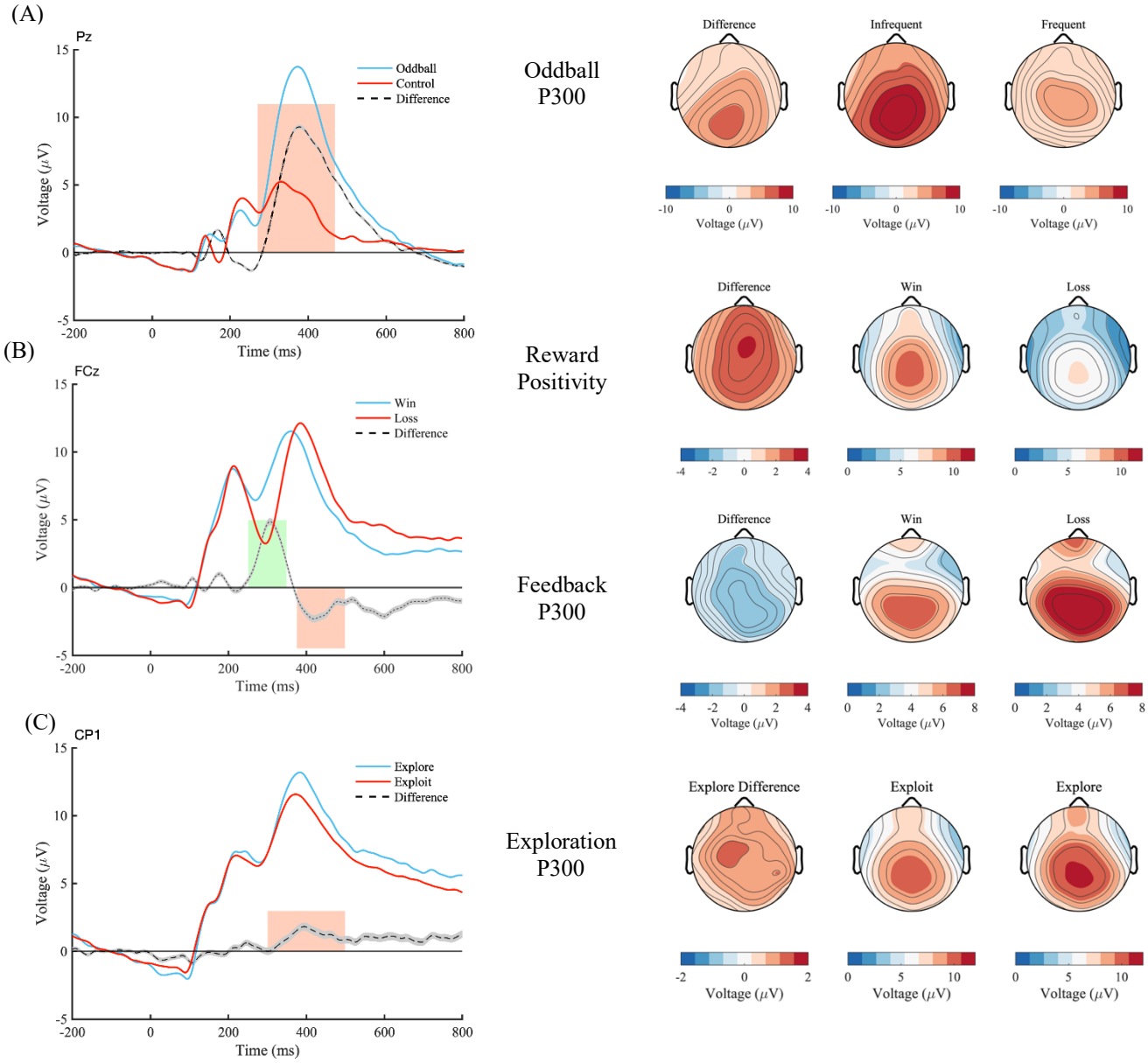

Figure 29. Experiment 4 - ERP results. (A) The oddball P300 in the Oddball task, (B) the reward positivity in the two-armed bandit task, (C) the exploration P300 in the two-armed bandit task. Topographic maps demonstrate the average of the conditional waveforms and difference waveforms within the component windows. Error bars are 95% between-subject confidence intervals.

Correlations

The correlation analysis revealed only a small, negative correlation between state anxiety and the P300 difference wave amplitude in the Oddball task (Table 13). Only the State Trait Anxiety Inventory - State scale scores and P300 amplitude correlation survived FDR adjustment ($r(307) = -.14$, $p < .01$). While the small, negative correlation between trait anxiety and the P300 was significant originally ($p < .05$), it did not survive the FDR adjustment. All other correlations were negligible (all $| r | < .10$, all $p > .25$), including between chronic and P300 amplitude, and between the reward positivity and all questionnaires.

Table 13

Experiment 4 - ERP component amplitude and questionnaire correlations

		Chronic Stress		Trait Anxiety		State Anxiety	
		Pearson *r*	95% CI	Pearson *r*	95% CI	Pearson *r*	95% CI
Oddball	Oddball P300	-.01	[-.12, .10]	-.11	[-.22, .00]	-.14*	[-.25, -.03]
Bandit	RewP	.03	[-.08, .14]	.07	[-.05, .18]	-.05	[-.16, .07]
	Feedback P300	.03	[-.08, 14]	.10	[-.02, 0.20]	.04	[-.07, 0.16]
	Exploration P300	-.04	[-.15, .07]	-.04	[-.15, .07]	.00	[-.11, .11]

Note: * indicates correlations which survived FDR adjustment to significance testing

Discussion

For experiment 4, I found that only anxiety was associated with changes to neurophysiology. Specifically, for my first goal of determining the relationship between chronic stress exposure, state anxiety, trait anxiety, and the reward positivity and the P300 signals (oddball, feedback, exploration), I found that state anxiety had a negative correlation with oddball P300 amplitude. As well, there was a correlation between state anxiety and both errors and reaction time within the Oddball task. No other correlations between questionnaire scores (including chronic stress) were associated with changes in the oddball P300, the feedback P300, the exploration P300, or the reward positivity. For my second goal, the investigation of explore-exploit decision making, I found no impact of either chronic stress or anxiety on exploration rate.

Generally, there were few effects of either chronic stress exposure, state anxiety, or trait anxiety on behaviour. Importantly, there was no effect of any of these measures on performance in the two-armed bandit task, including on exploration rate. There is preliminary evidence from a recent pre-print that trait anxiety may reduce exploratory behaviour (Fan et al., 2021) – and thus it is worth asking why there was no evidence for effects of trait anxiety in the present work. In fact, the authors conducted a factor analysis to isolate "somatic trait anxiety" using multiple surveys which measure anxiety (one of which, it should be noted, was the State-Trait Anxiety Inventory), and found that only somatic trait anxiety impact exploratory behaviour. Their other measures (cognitive anxiety, negative affect, and low self-esteem) had no impact. In the study by Fan and colleagues, somatic trait anxiety was a collection of questions from the State-Trait Anxiety Inventory and the State-Trait Inventory for Cognitive and Somatic Anxiety (Ree et al., 2008). Moreover, the State-Trait Inventory for Cognitive and Somatic Anxiety Scale contains a greater proportion of somatic anxiety items. Thus, perhaps I found no effect because the trait

anxiety construct which they measured – tied directly to the physical experience of anxiety – was a distinct form of trait anxiety from the form I measured here.

In addition, while there is evidence that chronic stress scores may reduce exploratory behaviour (Lenow et al., 2017), I did not observe any relationship between chronic stress and exploratory behaviour in the present work. Specifically, the authors found that chronic stress scores reduced exploration. An important distinction is that I used the Trier Inventory for the Assessment of Chronic Stress in the present work, while Lenow and colleagues used the Perceived Stress Scale (S. Cohen et al., 1983). One major difference between the two surveys is that the Trier Inventory for the Assessment of Chronic Stress uses a slightly longer timeframe (three months) while the Perceived Stress Scale only examines participants' perceptions of stress in the last month. As well, the Perceived Stress Scale assesses subjective stress responses while the Trier Inventory for the Assessment of Chronic Stress measures both the frequency of chronic stress exposure (how often) and subjective stress responses. This could suggest that the choice of survey influences the observed relationship as the two surveys are assessing different constructs.

Perhaps more importantly however, there were differences in the type of tasks used. That is, Lenow and colleagues used a patch-foraging task which stands in contrast to the two-armed bandit I used here. Importantly, there are different forms of exploratory behaviour (directed and random exploration; Wilson et al., 2014). It seems possible then, that the lack of impact of chronic stress exposure in this experiment could be because the type of exploratory behaviour I examined and Lenow and colleagues examined was different. Specifically, the two-armed bandit task might emphasize higher levels of directed exploration (which is specifically tied to reducing uncertainty about an option). Given participants in the present experiment were aware of the task structure before beginning (that one bandit wins more often), when exploring they likely would

be reducing relative uncertainty about one of the options to determine if it is the bandit that wins more often. In contrast, when making decisions to explore in the patch foraging task, participants are aware that the patches randomly differ in both reward levels and reward depletion rates randomly (there was no underlying structural differences between the patches). Thus, when participants are deciding to explore in the patch foraging task, their decisions to explore would likely be tied to total task uncertainty (they do not get to pick the next patch and there is no relative uncertainty between patches) and thus random exploration. Thus, it seems the two-armed bandit might be more likely to favour directed exploration while the patch foraging task is more likely to favour random exploration. Future work examining the impact of chronic stress on the explore-exploit dilemma should consider adopting tasks such as the Horizons task (Wilson et al., 2014) which allows for the differentiation of directed and random exploration.

Broadly, I found no relationship between chronic stress exposures on any of the four ERP components I examined (the oddball P300, the reward positivity, the feedback P300, and the exploration P300). Prior work examining the reward positivity has shown that a variety of measures of chronic stress exposures attenuate reward processing signals (Ethridge et al., 2018; Freeman et al., 2022; Suor et al., 2021). Thus, it is worth asking why there was no relationship between chronic stress exposures and any of the components examined here. One difference is the studies cited above examined chronic stress through constructs such as relational victimization (Ethridge et al., 2018), maternal depression (Freeman et al., 2022), and childhood maltreatment (Suor et al., 2021), while in the present work I examined chronic stress exposures as per scores on the Trier Inventory for the Assessment of Chronic Stress. Thus, perhaps how chronic stress is measured determines the relationship between stress exposure and reward positivity amplitude – which makes some amount of sense as chronic stress is a multi-faceted

construct involving both emotional appraisals, individual experiences, and the frequency of occurrence. Relatedly, these is evidence of interactions between individual differences and the impact of chronic stress on reward processing. For example, Suor et al., (2021) only found a negative relationship between reward positivity amplitude and childhood maltreatment in mothers with a history of major depressive disorder. As well, Ethridge and colleagues (2018) only found a relationship between relational victimization but not physical victimization to reward processing. In addition, Ethridge and colleagues specifically considered the neural response to gains rather than the reward positivity. Thus, perhaps the reliance on the reward positivity difference wave – as was used here in experiment 4 – might not always be the revealing approach. Broadly then, this suggests that individual differences (and disorders such as depression) are important moderators for the impact of chronic stress on event-related potential components and reward processing in general. Future work should thus ensure that previously identified individual differences are considered as they clearly impact the relationship between chronic stress and reward processing. Thus, in the present work, it could be that the lack of relationship observed might be due to these factors not being considered adequately. Alternatively, the examination of specific neural responses to gain or losses might also be more appropriate.

However, what about the exploration P300 – an uncertainty signal? Work examining reinforcement learning and chronic stress has shown that chronic stress levels are associated with how subjects weight uncertainty during a reinforcement learning task– although it should be noted that the authors split subjects into high and low chronic stress levels per their scores on a survey (Lei et al., 2021). Interestingly, the authors only found an impact of high chronic stress on uncertainty in female participants – suggesting there may be sex differences in the impact of

chronic stress (echoing prior work under acute stress - Lighthall et al., 2009). Again, there is an important methodological difference between their study and the present work. One difference is that they used the Perceived Stress Scale while I used the Trier Inventory for the Assessment of Chronic Stress (more specifically the summary score from the Trier Inventory for the Assessment of Chronic Stress). As mentioned, there is evidence that the two surveys provide indices of different measures of chronic stress as the perceived stress scale focuses on subjective reactions to stressful demands while the Trier Inventory for the Assessment of Chronic Stress instead focuses on the number of stressful environmental and cognitive experiences (Federenko et al., 2006). Another difference is that the authors modeled expected uncertainty rather than unexpected uncertainty. Specifically, unexpected uncertainty is the type of uncertainty which myself and others (Hassall, 2019) have argued the exploration P300 represents (c.f., Dayan & Yu, 2006; A. J. Yu & Dayan, 2005). In contrast, expected uncertainty has been argued to be closely tied to the activity of acetylcholine while unexpected uncertainty is tied to the activity of norepinephrine (A. J. Yu & Dayan, 2003). Thus, the difference in how chronic stress was measured and the difference in the uncertainty measures used could help explain the lack of relationship between chronic stress and uncertainty signals in the present work.

I also found that state anxiety was associated with a reduction in the amplitude of the oddball P300. This is unsurprising, as previous work has observed that both higher state anxiety (Rossi & Pourtois, 2017) are associated with reductions in the amplitude of the P300. In addition, that there was a negative association between state anxiety on the oddball P300 also mirrors previous work that has only examined participants who have scored high and low on anxiety measures (Huang et al., 2009). Interestingly, there was an association between higher state anxiety and an increase in both reaction time and errors made in the Oddball task. Given that the

oddball paradigm provides a measure of attention (S. Sutton et al., 1965), and that anxiety typically represents a heightened arousal to threat (Lieb, 2005), these behavioural and electrophysiological findings could be indicative of modulations to attentional systems under high levels of anxiety (e.g., Saviola et al., 2020) as per previous work (Huang et al., 2009; Rossi & Pourtois, 2017). In addition, that anxiety did not impact all three measures of the P300 in the same direction provides further evidence that the P300 is not a uniform construct and clearly reflects different cognitive processes related to different task demands and contexts (see chapter 6 for an extended discussion of this idea).

Perhaps more broadly, this work highlights that researchers should use caution when only sampling from the tails of the distribution to understand anxiety or chronic stress and their relationship to cognition. It is common practice to exclude or only examine participants who score low or high on measures of anxiety (e.g., Hajcak et al., 2003; Huang et al., 2009; Luo et al., 2018; Pedersen & Larson, 2016; Rossi & Pourtois, 2017; Xia et al., 2017; L. Zhang et al., 2015) and on measures of chronic stress (Lei et al., 2021). As is unsurprising, when a researcher samples from the tails of the distribution, she would observe a larger effect. Thus, it seems probable that researchers adopting this approach may be over-estimating the impact of anxiety and chronic stress on cognition or neurophysiology. While the argument likely is that by doing so you would be more closely approximating the clinical impact of those diagnosed with an anxiety or chronic stress disorder, some of these findings may be artificially elevating the size of the effect through their sampling methods. As such, researchers interested in using this approach should consider whether the collection of larger sample sizes might be more appropriate.

In sum the present work provides evidence from a large sample of undergraduate students that chronic stress does not appear to have an impact on feedback learning, contextual updating,

or uncertainty. In addition, there was no impact of chronic stress or anxiety on exploration. However, this work does provide evidence that anxiety is associated with changes in contextual updating signals, which I have argued is likely because they are directly tied to attentional systems – which typically show greater activation during periods of high anxiety (Saviola et al., 2020). While promising, it is evident that further research is needed on the role of chronic stress and anxiety in the explore-exploit dilemma.

General Discussion

Experiment Summaries

The main goal of my dissertation was to determine the impact of stress on the explore-exploit dilemma (Table 14). Specifically, I used electroencephalography (EEG) to measure feedback and uncertainty signals in the brain to determine whether there was evidence for how stress impacts the cognitive processes involved in making decisions to explore or exploit. In experiment 1, I found tentative evidence that the acute stress response disrupted both the feedback signal (the reward positivity) and the uncertainty signal (the switch/stay P300). In experiment 2, I found that the acute stress response only modulated the feedback P300 but not the exploration P300. Perhaps unsurprisingly given the task design and the use of contextual bandit tasks, in both experiments 1 and 2 I found no effect of the acute stress response on behaviour or exploration rate. However, in experiment 3, I again found tentative evidence that the acute stress response disrupts both the feedback P300, and stronger evidence that acute stress disrupts the exploration P300. Importantly, I found that acute stress increased exploration rate and decreased task performance (win percentage and optimal arm choice). In experiment 4, I sought to use a different approach and examined chronic stress exposure and anxiety – finding that chronic stress exposure had no meaningful impact on behaviour or neurophysiology while state anxiety was associated with a reduction in the oddball P300. For experiment 1 I used switch-stay behaviour as a proxy for exploration-exploitation while in experiments 2 through 4, I instead used computational models taken from the reinforcement learning literature to model exploratory decisions (Daw et al., 2006; R. S. Sutton and Barto, 2018).

Taken together these findings suggest possible neural mechanisms for how the acute stress response impacts decisions to explore while also suggesting some important methodological implications for work on the explore-exploit dilemma and stress more generally. I will now discuss some interpretations of these findings and how this works fits in a broader theoretical position not just in the stress and cognition literature, but on theories of the P300 as well. Given the lack of any evidence of a relationship between chronic stress exposure and my neurophysiological and behavioural measures, here then I will typically be referring to the acute stress response (due to acute stress exposure from the Trier Social Stress Test) and its effects for the rest of the dissertation, except where noted.

Table 14

Experiment summaries

Experiment and Task	Behavioural Findings	Neural Findings
1 – Contextual bandit	None	Acute stress disrupted the reward positivity (tentative) and the switch P300 (tentative)
2 – Contextual bandit	None	Acute stress reduced the feedback P300
3 – Learnable bandit & Oddball	Acute stress reduced exploration rate and win percentage	Acute stress disrupted the feedback P300 (tentative) and the exploration P300
4 – Learnable bandit & Oddball	Anxiety (state) was associated with more errors and slower reaction time in the Oddball task	Anxiety (state) was associated with a lower oddball P300

When examining the impact of acute stress on feedback signals, some interesting trends emerge. Specifically, I found effects of the acute stress response on both the reward positivity and the feedback P300 – although not in the same experiment. It is important to note – likely the lack of consistency in effects is due to the low power to detect effects of the acute stress response on the feedback signals. However, there might be additional contributing factors. For example, I hypothesized that the effect of the acute stress response on the reward positivity was because I used a contextual bandit in Experiment 1 with win/loss feedback. In contrast, I used a contextual bandit with point feedback in experiment 2 and a learnable bandit in Experiment 3. Given the use of point feedback in experiment 2, there was no clear reward positivity, making it impossible to determine if the acute stress response impacted the reward positivity in experiment 2. Thus, one possible methodological issue – in addition to the lack of power – is that the lack of effect of the acute stress response on the reward positivity in experiment 3 is due to the probabilistic nature of the learning task. In experiment 1, there was no specific learnable bandit – the highest value bandit shifted throughout the task and participants were aware of this shifting. That is, the acute stress response modulated the reward positivity under conditions of uncertainty (experiment 1) rather than a condition where the bandits are learnable (experiment 3). Although a full discussion of the lack of power as a limitation will be returned to later, in the following two paragraphs I will discuss some methodological differences in the literature which could contribute to the inconsistency in findings both within my dissertation and within the literature.

Prior work using probabilistic learning tasks has shown that acute stress exposure reduces the reward positivity (Banis & Lorist, 2012; Burani et al., 2021; Paul et al., 2020; X. Zhang et al., 2020) – however, there are key methodological differences in the prior investigations and the present work. In two of the papers, the authors used stressors which may not elicit a robust

cortisol response (Dickerson & Kemeny, 2004) including a noise stressor (Banis & Lorist, 2012), and the threat of shock paradigm (X. Zhang et al., 2020) – perhaps tying these effects on the reward positivity to arousal related processes and the SAM axis. Relatedly, Burani and colleagues used a cold-pressor task but they had participants dunk their hands in the cold water before every block making it difficult to differentiate the arousal-based, SAM axis effects from HPA axis effects.

An alternative explanation for this difference in the effect of the acute stress response on the reward positivity between experiment 1 (disruption) and experiment 3 (no effect) could be related to work showing that acute stress exposure causes participants to be vulnerable to fatigue effects in turn leading to task disengagement (Shafiei et al., 2012). The two-armed bandit task in experiment 3 only took five to seven minutes to complete, while the four-armed bandit in experiment 1 took between fifteen to twenty to complete. That is, the effects of the acute stress response on the reward positivity may only be observable in longer time-frame tasks and the different findings on the effect of the acute stress response on the reward positivity between experiment 1 and 3 may be related to the different task lengths employed in the present work. I also note that evidence showing reductions in the reward positivity have been from tasks which occurred over longer time frames – between 25 minutes (Banis & Lorist, 2012) and 36 minutes (Paul et al., 2020).

Importantly, I argued previously that the effect of the acute stress response on the P300 seen in experiment 2 (strong) and 3 (tentative) reflects the fact that the acute stress response might be modulating feedback signals by disrupting top-down assessments of feedback, which it has been hypothesized that the feedback P300 indexes (Wu & Zhou, 2009). In fact, prior work has shown that the acute stress response reduces the feedback P300 (Banis & Lorist, 2012; Paul

et al., 2020), although as mentioned above there are methodological distinctions between my task and their tasks such as the choice of stressor (e.g., Banis & Lorist, 2012). Thus, I believe that the acute stress response is modulating the top-down attentional signals for feedback, in turn disrupting feedback learning behaviour by preventing effective behavioural adjustments to feedback (as was seen in experiment 3). The effect of the acute stress response disrupting "top-down" P300 signals can be interpreted within the framework of research showing stress causes participants to adopt less costly cognitive strategies (Otto et al., 2013; Schwabe & Wolf, 2013) and increases the avoidance of cognitive effort (Bogdanov et al., 2021). Thus, the disruption of the top-down feedback signals in experiments 2 and 3 might be a reflection in the brain of the disruption of controlled processes related to feedback use – although it is evident further research is necessary.

Another important effect revealed in my dissertation is the (at times tentative) effect of the acute stress response on uncertainty signals in the brain. Work has argued that the exploration P300 may be a neural interrupt signal tied to likelihood and uncertainty (Hassall, 2019). That is, Hassall suggested that the exploration P300 reflects a burst of norepinephrine (c.f. Nieuwenhuis et al., 2005) which occurs when the participants' expectation of the task structure is disrupted due to unexpected uncertainty (Dayan & Yu, 2006). In the present work, it appears that the acute stress response disrupts the uncertainty signal – at least in some situations (experiment 1 and 3). I argue that the uncertainty signal might be disrupted because participants are either not fully engaged with the task (which occurs under acute stress – Sänger et al., 2014) or it could be that their internal assessments of uncertainty are not updating properly leading to them not fully integrating how certain they of the task structure (Dayan and Yu, 2006), similar to what occurs in substance abuse (J. C. Yu et al., 2020). A drawback in the present work is that I did not formally

model uncertainty. There are computational approaches – such as the Kalman Filter (Kalman, 1961) and Upper-confidence bound algorithms (R. S. Sutton & Barto, 2018) – which can be used to model uncertainty directly (Gershman, 2019). Future work could thus tie the impact of the acute stress response on the uncertainty signals (the exploration P300) to specific modulations on uncertainty extracted from a computational model to better understand the role of uncertainty signals in exploration.

While the results from experiments 1 through 3 highlight the neural impact of the acute stress response on uncertainty and feedback, in only experiment 3 was there an effect of the acute stress response on behaviour. Importantly, the behavioural results from experiment 3 showed that the acute stress response increased exploration while decreasing win percentage and optimal arm choice. My claim is that because the feedback signals and the uncertainty signals were disrupted, participants explored more often. That is, the participants were unable to incorporate feedback information effectively nor incorporate assessments of task uncertainty, which I argue led them to explore more often to the detriment of task performance. However, the causal direction of the impact of stress on feedback and uncertainty signals and the relationship to behaviour is unclear. That is, it could be that stress disrupted the neural signal of uncertainty (exploration P300) leading to more explorations and lower points, or it could be that the acute stress response disrupted the feedback signals which reduced learning, in turn impacting exploration. Of course, it could be that acute stress impacts both cognitive processes concurrently. The finding from experiment 3 could suggest both signals are needed to guide exploratory behavior effectively. Broadly however, to resolve this issue requires a need for more work in this area given the paucity of research on the topic of how stress impacts decisions to explore and exploit. Importantly, these findings from experiment 1 through 3 highlight possible

neural signals for exploration and the role of feedback and assessments of uncertainty in exploration, thus providing neural signals sensitive to these two cognitive processes which are key aspects of how humans solve the explore-exploit dilemma. As will be discussed below, these findings could have implications for our understanding of disorders where the explore-exploit dilemma is modulated (Addicott et al., 2017) or in the further refinement of artificial intelligence algorithms (R. S. Sutton & Barto, 2018). As well, this work adds to the growing literature on the impact of stress on decision making (e.g., Starcke & Brand, 2016).

The fact that in this work there was only a behavioural effect in experiment 3 requires further discussion. That is, although there were neural effects of the acute stress response in experiment 1 and 2 – some tentative as in the case of experiment 1 – there were no behavioural effects. I suggested that this might be due to the use of contextual bandit tasks in experiment 1 and 2 – that is the contextual bandit task was not an ideal task for eliciting behavioural differences. One piece of evidence for this was the fact that participants optimal arm choice was quite stable across the 400 trials in both experiment 1 and 2. The lack of ability to tie the neural findings to behaviour is concerning – suggesting they could be incidental and not tied to how participants solve or manage the explore-exploit dilemma. However, it is important to consider what EEG allows us to the investigate. In addition to the excellent temporal resolution, EEG as tool can be used for covert processes such as measuring changes to neural processing in populations such as infants or participants who can't make responses (as highlighted in Luck, 2014). In addition, EEG is can be used in attentional paradigms and can reveal further information that might be masked by traditional behaviour measures such as reaction time (e.g., Handy et al., 2001). As well, tasks such as the gambling task where positive feedback is presented on 50% of trials no matter the selection (e.g., Ethridge et al., 2020) would also by the

design of the task not produce a behaviour effect. I believe that the contextual bandits could be akin to the gambling task used by Ethridge and colleagues, but it is important to note that the findings from experiment 1 and 2 might simply be incidental and not related to changes in the explore-exploit dilemma given the lack of effect of acute stress on behavioural performance and exploration rate.

I should note that my finding that exploration increased following the acute stressor runs counter to prior work from humans showing that the acute stress response decreases exploration rate (Lenow et al., 2017). I believe that it is participants' ability to correctly solve the explore-exploit dilemma which is impacted due to the acute stress response – rather than acute stress causing a simple increase or decrease in exploration. Interpreted in this way, the behavioural effect from experiment 3 may not run counter to the results obtained by Lenow et al. However, why would there be differences in the direction of the effect of the acute stress response depending on the paradigm? I suggested in the discussion of experiment 3 that this could be due to a difference in the utility or ease of exploration in the two paradigms. A related explanation is that the exploratory behaviour examined in Lenow et al., (2017) and the exploratory behaviour examined in the present work may be different forms of exploration.

As mentioned, a useful framework for distinguishing types of exploratory behaviour is the distinction between directed and random exploration (Wilson et al., 2014). On one hand, directed exploration is suggested to be related to a person's desire to reduce the uncertainty of other options in their environment. On the other hand, random exploration is when a person explores due to a simple stochastic process related to decision noise. Given that participants were presented with both bandits in experiment 3, and that they were informed that one bandit produced more wins than the other, I believe that participants might have relied on directed

exploration more than random exploration as the participants were exploring to reduce uncertainty to learn which bandit was better. In contrast, in the patch foraging task used by Lenow, participants are faced with a single option (one patch) from which they get some reward, with the knowledge that the reward amount will be declining following each selection of that patch. As well, while they know that the patches differ in the amount of reward and the rate of decline, they are informed that the rate of decline would vary randomly between patches, and it could be that participants are exploring more randomly rather than to specifically reduce uncertainty in the patch environment. While evidence suggests that people use both types of exploratory behaviours to solve the explore-exploit dilemma (Wilson et al., 2014), perhaps the learnable bandit used in experiment 3 and the patch foraging task used by Lenow produce different demands for the type of exploratory behaviour. I believe that future work should consider examining how the acute stress response impacts directed and random exploration to help resolve this dilemma.

One interesting result from the present work is the differential effects of acute and chronic stress on the P300 in a variety of contexts. Currently, there does not exist a comprehensive theory of the P300 signal (Polich, 2020). A particularly prominent theory is that the P300 reflects context updating and working memory signals (Polich, 2007). However, more recent research has argued for a role of the P300 in cognitive processes such as surprise (Mars et al., 2008), learning (Fischer & Ullsperger, 2013; Nassar, Bruckner, et al., 2019), feedback processing (Wu & Zhou, 2009), uncertainty (Kopp et al., 2016), and exploration (Hassall, 2019). Here, I found tentative evidence that the acute stress response modulated uncertainty signals (the exploration P300) in experiment 1 and experiment 3. I also found tentative evidence that the acute stress response modulated the feedback P300 elicited in the bandit tasks in experiment 2

and 3. Importantly, in experiment 3, I found no effect of the acute stress response on the oddball P300. If interpreted narrowly, my findings do not support the P300 as solely a context updating signal tied to working memory when viewing a stimulus. I say this because the effect of the acute stress response on the P300 varied depending on whether feedback, stimulus presentation, or exploration was examined. If interpreted more broadly, than the context updating theory might be able to explain the elevated P300 to exploration as a distinct context (an exploration) tied to a representation of the task where exploitation was the default. However, work has shown that in tasks where there are more explorations than exploitations, the P300 is still elevated to exploration (Hassall & Krigolson, 2020) – suggesting that even when exploitations may not be the default state they still have smaller amplitude than explorations. In addition, that the effect of the acute stress response on the P300 was different whether feedback, exploration, or stimulus presentation were examined, suggests that the P300 indexes different cognitive processes dependent on the task and trial demands.

While the present data may be unable to answer the broad question of what the P300 signal means generally, it could be that there is no overarching cognitive process which the P300 represents. Given that multiple neural generators are tied to the P300 dependent on context and task demands (Johnson Jr., 1993), the P300 likely does not provide an index of a uniform cognitive process. At the very least, it seems that the P300 signal is sensitive to multiple cognitive processes and demands. In the present task, the feedback P300 may be sensitive to learning while the exploration P300 may be sensitive to uncertainty. That is, in experiment 3, the feedback P300 was tenuously modulated by acute stress and was correlated with task performance (a larger feedback P300 indicated better performance). In turn the exploration P300 was negatively correlated with decisions to explore as a larger exploration P300 was associated

with fewer explorations (and explorations were sub-optimal in experiment 3), suggesting the two neural signals were sensitive to different behavioural demands. While the idea of the P300 reflecting context updating is a useful framework, that acute stress did not consistently modulate the P300 across (or even within) tasks should cause one to pause when accepting the claim that the P300 is a context updating signal broadly.

Relatedly, given the *uncertainty* of what the P300 signal represents, it might be useful to re-visit my claim that the P300 to decisions to explore or exploit does in fact reflect some measure of uncertainty in both learnable and contextual bandit tasks. To put it another way, should the exploration P300 in the learnable bandit specifically be considered to provide an index of uncertainty? Briefly, prior work has shown that the P300 is larger under conditions of uncertainty (Kopp et al., 2016). As well, decisions to explore are tied to reducing uncertainty in an environment (Aston-Jones & Cohen, 2005) and the exploration P300 has been argued to reflect unexpected uncertainty related to norepinephrine (Hassall, 2019; c.f. A. J. Yu & Dayan, 2005). Generally, decisions made in contextual bandits are considered as decisions made under uncertainty (E. Schulz et al., 2018). This make some amount of sense as the decision maker is aware that throughout the task the win probabilities of the bandits are changing – that is there is a natural task uncertainty inherent to the design as there is always a chance one of the other options becomes the ideal choice due to changing reward probabilities. Importantly however, work by Gershman (2019) has specifically shown that uncertainty is also present in learnable two-armed bandit tasks. Using computational models, Gershman showed that both relative and total uncertainty are present in the task structure of learnable bandits. That is, participants in learnable bandits explore both to reduce uncertainty about other options and randomly due to task uncertainty. Thus, I believe that the P300 to decisions to explore and exploit does reflect an

uncertainty signal in the present work across all experiments. In fact, work has shown that the exploration P300 is sensitive to the prior likelihood of choices, as more exploration trials in a row reduced the exploration P300 signal, adding further credence to my suggestion (Hassall, 2019).

Limitations

The first major limitation is the lack of power to detect effects. As an example, across the three acute stress experiments (experiments 1 to 3), the effect of the acute stress response on the feedback P300 varied from η_p^2 = .00 (experiment 1), η_p^2 = .14 (experiment 2), and η_p^2 = .02 (experiment 3). These experiments were thus underpowered to detect small or medium effects of acute stress – and given the findings from the present experiments, the effect of the acute stress response on both behaviour and neurophysiology might be small or medium. In fact, I typically relied on an examination of the difference wave to infer disruption of the EEG signals – which took advantage of the within subject comparison and concomitant power increase. As such, many of the above conclusions regarding how acute stress impacts feedback and uncertainty signals need to be considered as tentative or suggestive. One appealing solution to this problem of a lack power to detect small or medium effects would be to instead test participants in a within-subject design – which would boost power. In fact, within-subject designs are adopted in the stress literature (e.g., Burani et al., 2021). Of course, researchers must be careful to try to minimize order effects, ensure adequate delay following the acute stressor (if the activity of the HPA axis and cortisol are of interest), and make sure to counter-balance when participants are exposed to the placebo task or the stressor. As will be highlighted below, another option to increase power is to collect and analyze cortisol following the placebo or acute stressor. The

collection of cortisol allows for a principled approach for removing potential outliers – participants who may not have responded to the stressor or participants who responded to the placebo stressor. However, and as will be expanded on below, cortisol only provides one measure of the acute stress response.

The second major limitation, which is closely related to the limitation of the lack power to detect small and medium effects, is the lack of behavioural effects of the acute stress response in experiment 1 and 2. I have argued that this limitation is likely due to the task design – contextual bandits may not be an appropriate task to distinguish between group effects on behaviour (such as exploration rate and task performance). That is, the contextual bandit task itself is too random. In fact, when I conducted a power analysis to detect a behavioural effect in experiment 2, the small effect size meant that 393 subjects would be needed in each condition for a desired power of 80% – and it should be noted I used a liberal approach by choosing the behaviour (reaction time) which had the largest effect size in experiment 1. What might be a solution to this issue other than collecting 400 participants in each group for contextual bandits – a more difficult proposition when collecting neuroimaging data? Given that in experiment 3 I was able to tie the neural findings to behaviour as there was a behavioural effect of acute stress, future work examining group differences could consider either adopting block-based designs where both contextual and stationary bandits are used while ensuring an adequate amount of trials for exploration and exploitations and wins and losses (Boudewyn et al., 2018). By adopting such as a design, experimenters would be able to assess both learnable bandits and contextual (i.e., non-learnable) bandits while also having a variation in task uncertainty.

The third major limitation of the present dissertation is the issue of component overlap. Specifically, both the feedback P300 and the exploration P300 overlapped in the time windows

used in my analyses. While I attempted to avoid this by examining frontal and posterior electrodes for the two P300 signals and collapsing across the other trial type for both components (feedback: collapsed across explore/exploit; exploration: collapsed across win/loss), future work should attempt to separate out wins and losses under exploratory decisions and wins and losses under exploitative decisions. Unfortunately, due to low trial counts in experiment 3 for the exploration condition, I was unable to test whether there were interactions between the type of feedback presented and the decision to explore or exploit. To put it another way, it could be that rather than acute stress tentatively impacting the feedback P300, it could be that the effect is driven by the effect of acute stress on the exploration P300 (or vice-versa). To counter this claim however, there is evidence that decisions to explore and decisions to exploit both have the same relationship between win and loss feedback as there was no interaction between exploration and feedback on the P300 (Hassall et al., 2019). That is, the authors found that there was a P300 difference in the same direction between wins and losses for both exploration and exploitation. Future work investigating the overlap between P300 components should consider either including a larger number of trials to ensure adequate power to detect a difference (as Hassall and colleagues did) or adopt a signal processing approach such as Principal Component Analysis to unmix these signals and account for component overlap (e.g., Spencer et al., 2001).

The fourth major limitation of the present work is that no cortisol data were reported. Although I included both emotional and physiological measures to determine if stress levels were elevated by the TSST – there is evidence that cortisol may drive some of the effects of stress on cognition (Joëls et al., 2011). As a pertinent example, in Lenow and colleague's investigation, the authors did not observe a between-condition effect of acute stress on explore-exploit decision making but did observe that cortisol was associated with reductions in

exploration rate. Importantly, there is evidence that some participants do not react to laboratory stressors such as the Trier Social Stress Test per salivary cortisol (Kirschbaum et al., 1993). For example, Paul and colleagues (2020) found that out of a sample of 25 participants in the acute stress condition, 8 (32%) of participants were non-responders and out of a sample of 25 in the control condition, 4 (16%) of participants were responders. Paul and colleagues then excluded these participants from the analysis. While these numbers do vary from experiment to experiment, and the type of method used to classify someone as a non-responder/responder matters (Miller et al., 2013), the use of cortisol can detect these non-responders to stress and responders to the placebo condition, in turn allowing for greater control of outliers. In addition, the collection of cortisol would allow for a stronger link between the effects of the acute stress response on explore-exploit decisions and the HPA axis to be made. However, it is important to not assume that just because a participant does not have a cortisol change relative to an acute stressor that they are not stressed out. For example, while females on hormonal birth control may have a reduced cortisol response, they likely still experience the stressor as being an aversive and stressful event and perhaps have compensatory hormonal or cognitive changes that are not observed when solely examining cortisol. That is, acute stress is a multifaceted construct which involves appraisals of the stressor and emotional changes. Thus, while cortisol could provide one means of removing outliers, researchers must consider the entirety of the acute stress response. Perhaps instead comparing low and high cortisol responders as sub-groups might be a more principled way to examine the strength of the impact of acute stress without relying on removing data.

Future Research

Several future directions seem evident for the present line of research. Perhaps one of the more obvious is to investigate how the acute stress response modulates oscillatory brain components such as frontal theta oscillations when participants are managing the explore-exploit dilemma given their role in cognitive control (e.g., Cavanagh & Frank, 2014; Ferguson et al., 2021). While it was not investigated in the present work, the use of oscillatory measures could be revealing. For example, the use of theta oscillations as a control mechanism recruited to deal with uncertainty, and induce exploration, might provide an explanation for some of the effects of the acute stress response observed here. That is, it is not simply that the feedback and uncertainty signals are impacted by the acute stress response, but it is also the ability of participants to recruit control processes which is disrupted. Oscillatory measures could provide more insight into the relationship between stress, uncertainty, and feedback – especially given that theta oscillations are typically larger following a loss to signal the recruitment of control (Cavanagh, Zambrano-Vazquez, et al., 2012). Theta oscillations could provide a particularly fruitful future direction as there is clear evidence of a role of both cognitive control and theta signals in exploratory behaviour which was specifically tied to uncertainty using a computational model (Cavanagh, Figueroa, et al., 2012). As well, the use of oscillatory measures might reveal more information about reward processing to compliment event-related potential analyses (Weinberg et al., 2019). This approach has been used previously to reveal nuanced effects of the acute stress response on reward processing (Ethridge et al., 2020), and thus the analysis of oscillatory measures seems a

fruitful future direction for understanding the explore-exploit dilemma following exposure to an acute stressor.

Another future direction is to not simply focus on explore or exploit signals at feedback presentation but during stimulus presentation or selection. In fact, the examination at stimulus selection is a common approach used in previous investigations of exploratory behaviour using EEG (e.g., Cavanagh, Figueroa, et al., 2012; Hassall et al., 2013), and thus seems ripe for investigation under acute stress. For example, it could be that not only does acute stress modulate the decisions people make and how their brain responds to feedback, but it could also be that acute stress impacts how people process the stimuli while they make their choice leading to modulations in the explore-exploit dilemma. As one example, the claim that following exposure to an acute stressor and activation of the acute stress response, participants might not be considering task structure or all available options during exploration could be strengthened through evidence of modulations to parietal alpha oscillations (which have been tied to attention; Klimesch et al., 2007) during selection. In fact, given that exploration and exploitation have been shown to differ in their recruitment of attentional processes (Walker et al., 2019), the use of the excellent temporal resolution of EEG to examine the explore-exploit dilemma – and the impact of stress – at stimulus presentation would be an exciting future direction.

Relatedly, future work should consider adopting tasks that specifically manipulate uncertainty and the type of exploratory decisions available to participants. One could imagine an experiment where participants are exposed to varying degrees of reward certainty on a block-by-block basis. In fact, work varying outcome predictability to induce distinct certainty levels (i.e., certain and uncertain blocks) has been integral in showing how attention is modified under exploration (Walker et al., 2019 c.f. Le Pelley et al., 2011), and this “learned predictiveness” task

would be ideal for future work. Another possible option would be to use a task such as the leap-frog bandit task (Knox et al., 2011) or the Horizons task (Wilson et al., 2014), as both tasks allow for different exploratory behaviours to be disentangled. For example, the use of the Horizons task separates out directed exploration and random exploration, and there is evidence that people will use both strategies concurrently (Wilson et al., 2014) but that there are different brain regions involved in these exploratory decisions (Zajkowski et al., 2017). Thus, future work examining the impact of acute stress on the explore-exploit dilemma could use these approaches to glean a better understanding of just how exploratory decisions are modified under stress.

Implications

Before I conclude, it is worth considering the broad context of the present work and why the interplay of stress and the explore-exploit dilemma is a valuable area of investigation. Thus, why might modulations to the explore-exploit dilemma under stress matter? Understanding how behaviour is impacted by stress is important – stress exposure is tied to numerous negative health outcomes (e.g., Taylor, 2010), and is implicated in mental health issues such as Alzheimer's dementia (Marin et al., 2011) and general age-related decline (Marshall et al., 2015) – and a greater understanding of the impacts of stress exposure can contribute to how treatments are developed for mental and physical health disorders. Importantly, problems in solving the explore-exploit dilemma are tied to disorders involving both impulsivity and risk-taking (Addicott et al., 2017). For example, there is evidence that relapse following substance-abuse might be predicted by performance on an explore-exploit task (L. S. Morris et al., 2016). Importantly, stress exposure can lead to relapses in people who suffer from substance-use disorders (Sinha, 2001). Thus, understanding the basic mechanisms of how stress (both chronic

and acute) impacts the explore-exploit dilemma in the brain might contribute to developing treatment for substance abuse problems. For example, one could imagine the addition of neurofeedback in clinical work to help train the uncertainty and feedback signals involved in the explore-exploit dilemma, perhaps strengthening these pathways, and helping substance abusers better avoid falling victim to relapse. Thus, developing a greater understanding of both the cognitive impact of stress and the impact of stress specifically on exploration and exploitation has consequences outside of the laboratory.

More generally, a better understanding the explore-exploit dilemma is a worthwhile goal in of itself. Exploration is a complex cognitive process. It involves attention, learning, reward valuation, and assessments of risk. More importantly, exploration is a cognitive process which governs many aspects of our lives from deciding what restaurant to eat at, what food to order at the restaurant, to even life questions such as deciding whether to pursue graduate school. Understanding not just how a human chooses to explore but the scale on which it decides to explore can thus help explain human behaviour and the choices humans make. Moreover, solving the explore-exploit dilemma is fundamental for algorithms in reinforcement learning and artificial intelligence (e.g., Yasui et al., 2019). As massive gains are made in these areas, we can see the effect of these algorithms on our daily lives in terms of social media (e.g., Zimmer et al., 2019), advertising (Miralles-Pechuán et al., 2018), and even healthcare (Jothi et al., 2015) despite notable issues with these algorithms including racial bias (Obermeyer et al., 2019). Thus, understanding the mechanisms of exploration, and how different cognitive processes and their corresponding neural signals are involved in solving the explore-exploit will likely help further refine and develop artificial intelligence algorithms.

Conclusions

In sum, the present dissertation provides evidence from a series of experiments attempting to understand the impact of stress on explore-exploit decision making. Specifically, it seems that both uncertainty and feedback signals are modulated by the acute stress response which in turn impacts how participants manage the explore-exploit trade-off. Thus, for the first time, neural evidence has been presented to explain why the explore-exploit dilemma is disrupted following an acute stressor. Importantly, task design had a clear impact on the effects observed and more specifically whether exploration and performance are impacted by the acute stress response. These findings also highlight the diverse role of the P300 signal in cognition and suggest that current theories of the functional significance of the P300 do not adequately explain the role of the P300 in cognition. However, further research is needed to understand the role of chronic stress on the explore-exploit dilemma. While promising, perhaps the main suggestion from this work is to highlight a clear need for further research on the explore-exploit dilemma under stress.

References

Addicott, M. A., Pearson, J. M., Sweitzer, M. M., Barack, D. L., & Platt, M. L. (2017). A Primer on Foraging and the Explore/Exploit Trade-Off for Psychiatry Research. *Neuropsychopharmacology*, *42*(10), 1931–1939. https://doi.org/10.1038/npp.2017.108

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov & F. Caski (Eds.), *Proceedings of the Second International Symposium on Information Theory* (pp. 267–281). Akademiai Kiado.

Allen, A. P., Kennedy, P. J., Cryan, J. F., Dinan, T. G., & Clarke, G. (2014). Biological and psychological markers of stress in humans: Focus on the Trier Social Stress Test. *Neuroscience and Biobehavioral Reviews*, *38*, 94–124. https://doi.org/10.1016/j.neubiorev.2013.11.005

Arnsten, A. F. T. (2009). stress signalling pathways and PFC. *Molecular Biology*, *26*(2), 148. https://doi.org/10.1038/nrn2648.Stress

Arnsten, A. F. T. (2015). Stress weakens prefrontal networks: Molecular insults to higher cognition. *Nature Neuroscience*, *18*(10), 1376–1385. https://doi.org/10.1038/nn.4087

Arnsten, A. F. T., & Goldman-Rakic, P. S. (1998). Noise stress impairs prefrontal cortical cognitive function in monkeys: Evidence for a hyperdopaminergic mechanism. *Archives of General Psychiatry*, *55*(4), 362–368. https://doi.org/10.1001/archpsyc.55.4.362

Aston-Jones, G., & Cohen, J. D. (2005). An Integrative Theory of Locus Coeruleus-Norepinephrine Function: Adaptive Gain and Optimal Performance. *Annu. Rev. Neurosci*, *28*, 403–450. https://doi.org/10.1146/annurev.neuro.28.061604.135709

Aylward, J., Valton, V., Ahn, W., 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*, 1116–1123. https://doi.org/10.1038/s41562-019-0628-0

Baker, T. E., & Holroyd, C. B. (2009). Which Way Do I Go? Neural Activation in Response to Feedback and Spatial Processing in a Virtual T-Maze. *Cerebral Cortex*, *19*(8), 1708–1722. https://doi.org/10.1093/cercor/bhn223

Baker, Travis E., & Holroyd, C. B. (2011). Dissociated roles of the anterior cingulate cortex in reward and conflict processing as revealed by the feedback error-related negativity and N200. *Biological Psychology*, *87*(1), 25–34. https://doi.org/10.1016/j.biopsycho.2011.01.010

Banis, S., Geerligs, L., & Lorist, M. M. (2014). Acute stress modulates feedback processing in men and women: Differential effects on the feedback-related negativity and theta and beta power. *PLoS ONE*, *9*(4). https://doi.org/10.1371/journal.pone.0095690

Banis, S., & Lorist, M. M. (2012). Acute noise stress impairs feedback processing. *Biological Psychology*, *91*(2), 163–171. https://doi.org/10.1016/j.biopsycho.2012.06.009

Baumgartl, H., Fezer, E., & Buettner, R. (2020). Two-level classification of chronic stress using machine learning on resting-state EEG recordings. *26th Americas Conference on Information Systems, AMCIS 2020*, *August*, 0–9.

Becker, P., Schulz, P., & Schlotz, W. (2004). Persönlichkeit, chronischer stress und körperliche gesundheit: Eine prospektive studie zur überprüfung eines systemischen anforderungs-ressourcen-modells. *Zeitschrift Für Gesundheitspsychologie*, *12*(1), 11–23.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1), 289–300.

Berghorst, L. H., Bogdan, R., Frank, M. J., & Pizzagalli, D. A. (2013). Acute stress selectively

reduces reward sensitivity. *Frontiers in Human Neuroscience*, *7*(MAR), 1–15. https://doi.org/10.3389/fnhum.2013.00133

Berridge, K. C. (2000). Reward learning: Reinforcement, incentives, and expectations. *Psychology of Learning and Motivation - Advances in Research and Theory*, *40*, 223–278. https://doi.org/10.1016/s0079-7421(00)80022-5

Bishop, S. J. (2009). Trait anxiety and impoverished prefrontal control of attention. *Nature Neuroscience*, *12*(1), 92–98. https://doi.org/10.1038/nn.2242

Bishop, S. J., Duncan, J., Brett, M., & Lawrence, A. D. (2004). Prefrontal cortical function and anxiety : controlling attention to threat-related stimuli. *Nature Neuroscience*, *7*(2), 184–188. https://doi.org/10.1038/nn1173

Bishop, S. J., Duncan, J., & Lawrence, A. D. (2004). State Anxiety Modulation of the Amygdala Response to Unattended Threat-Related Stimuli. *The Journal of Neuroscience*, *24*(46), 10364–10368. https://doi.org/10.1523/JNEUROSCI.2550-04.2004

Bogdan, R., & Pizzagalli, D. A. (2006). Acute Stress Reduces Reward Responsiveness: Implications for Depression. *Biological Psychiatry*, *60*(10), 1147–1154. https://doi.org/10.1016/j.biopsych.2006.03.037

Bogdanov, M., Nitschke, J. P., LoParco, S., Bartz, J. A., & Otto, A. R. (2021). Acute Psychosocial Stress Increases Cognitive-Effort Avoidance. *Psychological Science*, *32*(9), 1463–1475. https://doi.org/10.1177/09567976211005465

Bornstein, S. R., Engeland, W. C., Ehrhart-Bornstein, M., & Herman, J. P. (2008). Dissociation of ACTH and glucocorticoids. *Trends in Endocrinology and Metabolism*, *19*(5), 175–180. https://doi.org/10.1016/j.tem.2008.01.009

Boudewyn, M. A., Luck, S. J., Farrens, J. L., & Kappenman, E. S. (2018). How many trials does

it take to get a significant ERP effect? It depends. *Psychophysiology*, *55*(6), 1–16. https://doi.org/10.1111/psyp.13049

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436.

Burani, K., Gallyer, A., Ryan, J., Jordan, C., Joiner, T., & Hajcak, G. (2021). Acute stress reduces reward-related neural activity: Evidence from the reward positivity. *Stress*, *0*(0), 1–7. https://doi.org/10.1080/10253890.2021.1929164

Butts, K. A., Weinberg, J., Young, A. H., & Phillips, A. G. (2011). Glucocorticoid receptors in the prefrontal cortex regulate stress-evoked dopamine efflux and aspects of executive function. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(45), 18459–18464. https://doi.org/10.1073/pnas.1111746108

Byrne, K. A., Cornwall, A. C., & Worthy, D. A. (2019). Acute stress improves long-term reward maximization in decision-making under uncertainty. *Brain and Cognition*, *133*(March), 84–93. https://doi.org/10.1016/j.bandc.2019.02.005

Byrne, K. A., Peters, C., Willis, H. C., Phan, D., Cornwall, A., & Worthy, D. A. (2020). Acute stress enhances tolerance of uncertainty during decision-making. *Cognition*, *205*(August), 104448. https://doi.org/10.1016/j.cognition.2020.104448

Cabib, S., & Puglisi-Allegra, S. (2012). The mesoaccumbens dopamine in coping with stress. *Neuroscience and Biobehavioral Reviews*, *36*(1), 79–89. https://doi.org/10.1016/j.neubiorev.2011.04.012

Cannon, W. B. (1914). The emergency function of the medulla in pain and the major emotions. *American Physiological Society*, *33*(2), 356–372.

Cannon, W. B. (1929). Organization for physiological homeostasis. *Physiological Reviews*, *9*(3), 399–431. https://doi.org/10.1152/physrev.1929.9.3.399

Carvalheiro, J., Conceiç, V. A., Mesquita, A., & Seara-cardoso, A. (2021). Acute stress impairs reward learning in men Joana. *Brain and Cognition*, *147*(105657). https://doi.org/10.1016/j.bandc.2020.105657

Cavanagh, J. F., Figueroa, C. M., Cohen, M. X., & Frank, M. J. (2012). Frontal theta reflects uncertainty and unexpectedness during exploration and exploitation. *Cerebral Cortex*, *22*(11), 2575–2586. https://doi.org/10.1093/cercor/bhr332

Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in Cognitive Sciences*, *18*(8), 414–421. https://doi.org/10.1016/j.tics.2014.04.012

Cavanagh, J. F., Zambrano-Vazquez, L., & Allen, J. J. B. (2012). Theta lingua franca: a common mid-frontal substrate for action monitoring processes. *Psychophysiology*, *49*(2), 220–238. https://doi.org/10.1111/j.1469-8986.2011.01293.x

Charnov, E. L. (1976). Optimal foraging, the marginal value theorem. *Theoretical Population Biology*, *9*(2), 129–136. https://doi.org/10.1016/0040-5809(76)90040-X

Cohen, J. D., McClure, S. M., & Yu, A. J. (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1481), 933–942. https://doi.org/10.1098/rstb.2007.2098

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior*, *24*(4), 385–396.

Constantino, S. M., & Daw, N. D. (2015). Learning the opportunity cost of time in a patch-foraging task. *Cognitive, Affective and Behavioral Neuroscience*, *15*(4), 837–853. https://doi.org/10.3758/s13415-015-0350-y

Cook, S. C., & Wellman, C. L. (2004). Chronic stress alters dendritic morphology in rat medial

prefrontal cortex. *Journal of Neurobiology*, *60*(2), 236–248. https://doi.org/10.1002/neu.20025

Costa, V. D., Tran, V. L., Turchi, J., & Averbeck, B. B. (2014). Dopamine modulates novelty seeking behavior during decision making. *Behavioural Neuroscience*, *128*(5), 556–566. https://doi.org/10.1016/j.physbeh.2017.03.040

Critchley, H. D., Mathias, C. J., & Dolan, R. J. (2001). Neural activity in the human brain relating to uncertainty and arousal during anticipation. *Neuron*, *29*(2), 537–545. https://doi.org/10.1016/S0896-6273(01)00225-2

Dallman, M. F., Pecoraro, N., Akana, S. F., La Fleur, S. E., Gomez, F., Houshyar, H., Bell, M. E., Bhatnagar, S., Laugero, K. D., & Manalo, S. (2003). Chronic stress and obesity: A new view of "comfort food." *Proceedings of the National Academy of Sciences of the United States of America*, *100*(20), 11696–11701. https://doi.org/10.1073/pnas.1934666100

Daviu, N., Bruchas, M. R., Moghaddam, B., Sandi, C., & Beyeler, A. (2019). Neurobiological links between stress and anxiety. *Neurobiology of Stress*, *11*(100191), 1–9. https://doi.org/10.1016/j.ynstr.2019.100191

Daw, N. D. (2011). Trial-by-trial data analysis using computational models. *Decision Making, Affect, and Learning: Attention and Performance XXIII*, 1–26. https://doi.org/10.1093/acprof:oso/9780199600434.003.0001

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

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

https://doi.org/10.1080/09548980601004024

De Kloet, E. R., Joëls, M., & Holsboer, F. (2005). Stress and the brain: From adaptation to disease. *Nature Reviews Neuroscience*, *6*(6), 463–475. https://doi.org/10.1038/nrn1683

Dedovic, K., D'Aguiar, C., & Pruessner, J. C. (2009). What stress does to your brain: A review of neuroimaging studies. *Canadian Journal of Psychiatry*, *54*(1), 6–15. https://doi.org/10.1177/070674370905400104

Dedovic, K., Wadiwalla, M., Engert, V., & Pruessner, J. C. (2009). The Role of Sex and Gender Socialization in Stress Reactivity. *Developmental Psychology*, *45*(1), 45–55. https://doi.org/10.1037/a0014433

Delaney, L. D., Fink, G., & Harmon, C. P. (2021). Effects of Stress on Economic Decision-Making: Evidence from Laboratory Experiments. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2420705

Delorme, A., & Makeig, S. (2004). EEGLAB : an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*, 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Dickerson, S. S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research. *Psychological Bulletin*, *130*(3), 355–391. https://doi.org/10.1037/0033-2909.130.3.355

Dierolf, A. M., Fechtner, J., Böhnke, R., Wolf, O. T., & Naumann, E. (2017). Influence of acute stress on response inhibition in healthy men: An ERP study. *Psychophysiology*, *54*(5), 684–695. https://doi.org/10.1111/psyp.12826

Donchin, E. (1981). Surpise! ... Suprise? *Psychophysiology*, *18*(5), 493–513.

Donchin, E., & Coles, M. G. H. (1988). Is the P300 component a manifestation of context

updating? *Behavioral and Brain Sciences*, *11*(3), 357–374. https://doi.org/10.1017/S0140525X00058027

Duncko, R., Cornwell, B., Cui, L., Merikangas, K. R., & Grillon, C. (2007). Acute exposure to stress improves performance in trace eyeblink conditioning and spatial learning tasks in healthy men. *Learning & Memory*, *14*(5), 329–335. https://doi.org/10.1101/lm.483807.applied

Elzinga, B. M., & Roelofs, K. (2005). Cortisol-induced impairments of working memory require acute sympathetic activation. *Behavioral Neuroscience*, *119*(1), 98–103. https://doi.org/10.1037/0735-7044.119.1.98

Ethridge, P., Ali, N., Racine, S. E., Pruessner, J. C., & Weinberg, A. (2020). Risk and Resilience in an Acute Stress Paradigm: Evidence From Salivary Cortisol and Time-Frequency Analysis of the Reward Positivity. *Clinical Psychological Science*, *8*(5), 872–889. https://doi.org/10.1177/2167702620917463

Ethridge, P., Sandre, A., Dirks, M. A., & Weinberg, A. (2018). Past-year relational victimization is associated with a blunted neural response to rewards in emerging adults. *Social Cognitive and Affective Neuroscience*, *13*(12), 1259–1267. https://doi.org/10.1093/scan/nsy091

Fan, H., Gershman, S. J., & Phelps, E. A. (2021). *Trait somatic anxiety is associated with reduced directed exploration and underestimation of uncertainty*. PsyArXiv. https://psyarxiv.com/yx6sb/

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyzes using G * Power 3.1: Tests for correlation and regression analyzes. *Behavior Research Methods*, *41*, 1149-1160.

Federenko, I. S., Schlotz, W., Kirschbaum, C., Bartels, M., Hellhammer, D. H., & Wüst, S.

(2006). The heritability of perceived stress. *Psychological Medicine*, *36*(3), 375–385. https://doi.org/10.1017/S0033291705006616

Ferguson, T. . D., Williams, C. C., Skelton, R. W., & Krigolson, O. E. (2019). Passively learned spatial navigation cues evoke reinforcement learning reward signals. *Cognition*, *189*. https://doi.org/10.1016/j.cognition.2019.03.015

Ferguson, T. D., Bub, D. N., Masson, M. E. J., & Krigolson, O. E. (2021). The role of cognitive control and top-down processes in object affordances. *Attention, Perception, and Psychophysics*, *83*(5), 2017–2032. https://doi.org/10.3758/s13414-021-02296-z

Fink, G. (2016). Stress, Definitions, Mechanisms, and Effects Outlined: Lessons from Anxiety. In G. Fink (Ed.), *Stress: Concepts, Cognition, Emotion, and Behavior: Handbook of Stress Series* (1st ed., Vol. 1, Issue April, pp. 3–11). Academic Press.

Finlay, J. M., Zigmond, M. J., & Abercrombie, E. D. (1995). Increased dopamine and norepinephrine release in medial prefrontal cortex induced by acute and chronic stress: Effects of diazepam. *Neuroscience*, *64*(3), 619–628. https://doi.org/10.1016/0306-4522(94)00331-X

Fischer, A. G., & Ullsperger, M. (2013). Real and fictive outcomes are processed differently but converge on a common adaptive mechanism. *Neuron*, *79*(6), 1243–1255. https://doi.org/10.1016/j.neuron.2013.07.006

Foote, S. L., & Morrison, J. H. (1987). Extrathalamic modulatiomn of cortical function. *Annual Review of Neuroscience*, *10*, 67–95.

Freeman, C., Ethridge, P., Banica, I., Sandre, A., Dirks, M. A., Kujawa, A., & Weinberg, A. (2022). Neural response to rewarding social feedback in never-depressed adolescent girls and their mothers with remitted depression: Associations with multiple risk indices. *Journal*

of Psychopathology and Clinical Science, *131*(2), 141–151. https://doi.org/10.1037/abn0000728

Gershman, S. J. (2019). Uncertainty and exploration. *Decision*, *6*(3), 277–286. https://doi.org/10.1037/dec0000101.Uncertainty

Glienke, K., Wolf, O. T., & Bellebaum, C. (2015). The impact of stress on feedback and error processing during behavioral adaptation. *Neuropsychologia*, *71*, 181–190. https://doi.org/10.1016/j.neuropsychologia.2015.04.004

Godoy, L. D., Rossignoli, M. T., Delfino-Pereira, P., Garcia-Cairasco, N., & Umeoka, E. H. de L. (2018). A comprehensive overview on stress neurobiology: Basic concepts and clinical implications. *Frontiers in Behavioral Neuroscience*, *12*(July), 1–23. https://doi.org/10.3389/fnbeh.2018.00127

Goldstein, L. E., Rasmusson, A. M., Bunney, B. S., & Roth, R. H. (1996). Role of the amygdala in the coordination of behavioral, neuroendocrine, and prefrontal cortical monoamine responses to psychological stress in the rat. *Journal of Neuroscience*, *16*(15), 4787–4798. https://doi.org/10.1523/jneurosci.16-15-04787.1996

Goodman, W. K., Janson, J., & Wolf, J. M. (2017). Meta-analytical assessment of the effects of protocol variations on cortisol responses to the Trier Social Stress Test. *Psychoneuroendocrinology*, *80*, 26–35. https://doi.org/10.1016/j.psyneuen.2017.02.030

Grillon, C., Duncko, R., Covington, M. F., Kopperman, L., & Kling, M. A. (2007). Acute Stress Potentiates Anxiety in Humans. *Biological Psychiatry*, *62*, 1183–1186. https://doi.org/10.1016/j.biopsych.2007.06.007

Groeneweg, F. L., Karst, H., de Kloet, E. R., & Joëls, M. (2011). Rapid non-genomic effects of corticosteroids and their role in the central stress response. *Journal of Endocrinology*,

209(2), 153–167. https://doi.org/10.1530/JOE-10-0472

Groombridge, C. J., Kim, Y., Maini, A., Smit, D. V., & Fitzgerald, M. C. (2019). Stress and decision-making in resuscitation: A systematic review. *Resuscitation*, *144*(September), 115–122. https://doi.org/10.1016/j.resuscitation.2019.09.023

Hajcak, G., Holroyd, C. B., Moser, J. S., & Simons, R. F. (2005). Brain potentials associated with expected and unexpected good and bad outcomes. *Psychophysiology*, *42*(2), 161–170. https://doi.org/10.1111/j.1469-8986.2005.00278.x

Hajcak, G., Mcdonald, N., & Simons, R. F. (2003). Anxiety and error-related brain activity. *Biological Psychology*, *64*, 77–90. https://doi.org/10.1016/S0301-0511(03)00103-0

Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2007). It's worse than you thought: The feedback negativity and violations of reward prediction in gambling tasks. *Psychophysiology*, *44*(6), 905–912. http://10.0.4.87/j.1469-8986.2007.00567.x

Handy, T. C., Green, V., Mangun, G. R., & Klein, R. M. (2001). Combined expectancies: Event-related potentials reveal the early benefits of spatial attention that are obscured by reaction time measures. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(2), 303–317. https://doi.org/10.1037/0096-1523.27.2.303

Hassall, C. D. (2019). *The neural correlates of exploration*. University of Victoria.

Hassall, C. D., Holland, K., & Krigolson, O. E. (2013). What do I do now? An electroencephalographic investigation of the explore/exploit dilemma. *Neuroscience*, *228*, 361–370. https://doi.org/10.1016/j.neuroscience.2012.10.040

Hassall, C. D., & Krigolson, O. E. (2020). Neuropsychologia Feedback processing is enhanced following exploration in continuous environments. *Neuropsychologia*, *146*(107538). https://doi.org/10.1016/j.neuropsychologia.2020.107538

Hassall, C. D., McDonald, C. G., & Krigolson, O. E. (2019). Ready, set, explore! Event-related potentials reveal the time-course of exploratory decisions. *Brain Research*, *1719*(May), 183–193. https://doi.org/10.1016/j.brainres.2019.05.039

Hein, T. P., Fockert, J. De, & Herrojo, M. (2021). NeuroImage State anxiety biases estimates of uncertainty and impairs reward learning in volatile environments. *NeuroImage*, *224*(117424). https://doi.org/10.1016/j.neuroimage.2020.117424

Herman, J. P., Ostrander, M. M., Mueller, N. K., & Figueiredo, H. (2005). Limbic system mechanisms of stress regulation: Hypothalamo-pituitary- adrenocortical axis. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *29*(8), 1201–1213. https://doi.org/10.1016/j.pnpbp.2005.08.006

Hermans, E. J., Henckens, M. J. A. G., Joëls, M., & Fernández, G. (2014). Dynamic adaptation of large-scale brain networks in response to acute stressors. *Trends in Neurosciences*, *37*(6), 304–314. https://doi.org/10.1016/j.tins.2014.03.006

Het, S., Rohleder, N., Schoofs, D., Kirschbaum, C., & Wolf, O. T. (2009). Neuroendocrine and psychometric evaluation of a placebo version of the "Trier Social Stress Test." *Psychoneuroendocrinology*, *34*(7), 1075–1086. https://doi.org/10.1016/j.psyneuen.2009.02.008

Hines, E. A., & Brown, G. E. (1932). A standard stimulus for measuring vasomotor reactions: its application in the study of hypertension. *Proceedings of the Staff Meeting of the Mayo Clinic*.

Holroyd, C. B., & Coles, M. G. H. (2002). The Neural Basis of Human Error Processing: Reinforcement Learning, Dopamine, and the Error-Related Negativity. *Psychological Review*, *109*(4), 679–709. https://doi.org/10.1037/0033-295X.109.4.679

Holroyd, C. B., Krigolson, O. E., & Lee, S. (2011). Reward positivity elicited by predictive cues. *NeuroReport*, *22*(5), 249–252. https://doi.org/10.1097/WNR.0b013e328345441d

Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., & Camerer, C. F. (2005). Neuroscience: Neural systems responding to degrees of uncertainty in human decision-making. *Science*, *310*(5754), 1680–1683. https://doi.org/10.1126/science.1115327

Huang, Y., Bai, L., Ai, H., Li, W., Yu, C., Liu, J., & Luo, Y. (2009). Influence of trait-anxiety on inhibition function : Evidence from ERPs study. *Neuroscience Letters*, *456*, 1–5. https://doi.org/10.1016/j.neulet.2009.03.072

Jepma, M., Murphy, P. R., Nassar, M. R., Rangel-Gomez, M., Meeter, M., & Nieuwenhuis, S. (2016). Catecholaminergic Regulation of Learning Rate in a Dynamic Environment. *PLoS Computational Biology*, *12*(10), 1–24. https://doi.org/10.1371/journal.pcbi.1005171

Joëls, M., Fernandez, G., & Roozendaal, B. (2011). Stress and emotional memory: A matter of timing. *Trends in Cognitive Sciences*, *15*(6), 280–288. https://doi.org/10.1016/j.tics.2011.04.004

Johnson Jr., R. (1993). On the neural generators of the P300 component of the event-related potential. *Psychophysiology*, *30*, 90–97.

Jothi, N., Rashid, N. A., & Husain, W. (2015). Data Mining in Healthcare - A Review. *Procedia Computer Science*, *72*, 306–313. https://doi.org/10.1016/j.procs.2015.12.145

Kalman, R. E. (1961). A new approach to linear filtering and prediction theory. *Transactions of the ASME-Journal of Basic Engineering*, *82*(Series D), 35–45.

Kappenman, E. S., & Luck, S. J. (2016). Techniques and Methods Best Practices for Event-Related Potential Research in Clinical Populations. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *1*(2), 110–115. https://doi.org/10.1016/j.bpsc.2015.11.007

Kelly, S. P., & O'Connell, R. G. (2013). Internal and external influences on the rate of sensory evidence accumulation in the human brain. *Journal of Neuroscience*, *33*(50), 19434–19441. https://doi.org/10.1523/JNEUROSCI.3355-13.2013

Kirschbaum, C., Pirke, K. M., & Hellhammer, D. H. (1993). The "Trier social stress test" - A tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology*, *28*(1–2), 76–81. https://doi.org/10.1159/000119004

Kirschbaum, C., Wust, S., & Hellhammer, D. (1992). Consistent Sex Differences in Cortisol Responses to Psychological Stress. *Psychosomatic Medicine*, *54*, 648–657.

Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: The inhibition–timing hypothesis. *Brain Research Reviews*, *53*(1), 63–88. https://doi.org/10.1016/J.BRAINRESREV.2006.06.003

Knox, W. B., Otto, A. R., Stone, P., & Love, B. C. (2011). The nature of belief-directed exploratory choice in human decision-making. *Frontiers in Psychology*, *2*(JAN), 1–12. https://doi.org/10.3389/fpsyg.2011.00398

Kopp, B., Seer, C., Lange, F., Kluytmans, A., Kolossa, A., Fingscheidt, T., & Hoijtink, H. (2016). P300 amplitude variations, prior probabilities, and likelihoods: A Bayesian ERP study. *Cognitive, Affective and Behavioral Neuroscience*, *16*(5), 911–928. https://doi.org/10.3758/s13415-016-0442-3

Kreis, I., Zhang, L., Moritz, S., & Pfuhl, G. (2021). Spared performance but increased uncertainty in schizophrenia: Evidence from a probabilistic decision-making task. *Schizophrenia Research*, *June*. https://doi.org/10.1016/j.schres.2021.06.038

Krigolson, O. E. (2017). Event-related brain potentials and the study of reward processing: Methodological considerations. *International Journal of Psychophysiology*, *July*, 0–1.

https://doi.org/10.1016/j.ijpsycho.2017.11.007

Kudielka, B. M., Hellhammer, H., & Kirschbaum, C. (2007). Ten years of research with the trier social stress test. *Social Neuroscience*, *January 2007*, 56–83.

Kujawa, A., Klein, D. N., Pegg, S., & Weinberg, A. (2020). Developmental trajectories to reduced activation of positive valence systems: A review of biological and environmental contributions. *Developmental Cognitive Neuroscience*, *43*, 100791. https://doi.org/10.1016/j.dcn.2020.100791

Lataster, J., Collip, D., Ceccarini, J., Haas, D., Booij, L., van Os, J., Pruessner, J., Van Laere, K., & Myin-Germeys, I. (2011). Psychosocial stress is associated with in vivo dopamine release in human ventromedial prefrontal cortex: A positron emission tomography study using [18F]fallypride. *NeuroImage*, *58*(4), 1081–1089. https://doi.org/10.1016/j.neuroimage.2011.07.030

Lazarus, R. S. (2006). *Stress and emotion: A new synthesis* (R. S. Lazarus (ed.)). Springer Publishing Co.

Le Pelley, M. E., Beesley, T., & Griffiths, O. (2011). Overt Attention and Predictiveness in Human Contingency Learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *37*(2), 220–229. https://doi.org/10.1037/a0021384

Leder, J., Häusser, J. A., & Mojzisch, A. (2013). Stress and strategic decision-making in the beauty contest game. *Psychoneuroendocrinology*, *38*(9), 1503–1511. https://doi.org/10.1016/j.psyneuen.2012.12.016

Lei, H., Mochizuki, Y., Chen, C., Hagiwara, K., Hirotsu, M., Matsubara, T., & Nakagawa, S. (2021). Sex difference in the weighting of expected uncertainty under chronic stress. *Scientific Reports*, *11*(1), 1–10. https://doi.org/10.1038/s41598-021-88155-1

Lenow, J. K., Constantino, S. M., Daw, N. D., & Phelps, E. A. (2017). Chronic and acute stress promote overexploitation in serial decision making. *Journal of Neuroscience*, *37*(23), 5681–5689. https://doi.org/10.1523/JNEUROSCI.3618-16.2017

Lewandowsky, S., & Farrell, S. (2018). *Computational modeling in cognition: Principles and practice.* SAGE Publications Inc.

Lieb, R. (2005). Anxiety disorders: clinical presentation and epidemiology. In J. E. Barrett (Ed.), *Handbook of Experimental Pharmacology* (pp. 405–432). Springer.

Lighthall, N. R., Gorlick, M. A., Schoeke, A., Frank, M. J., & Mather, M. (2013). Stress modulates reinforcement learning in younger and older adults. *Psychology and Aging*, *28*(1), 35–46. https://doi.org/10.1037/a0029823

Lighthall, N. R., Mather, M., & Gorlick, M. A. (2009). Acute Stress Increases Sex Differences in Risk Seeking in the Balloon Analogue Risk Task. *PLoS ONE*, *4*(7), 1–6. http://10.0.5.91/journal.pone.0006002

Lighthall, N. R., Sakaki, M., Vasunilashorn, S., Nga, L., Somayajula, S., Chen, E. Y., Samii, N., & Mather, M. (2012). Gender differences in reward-related decision processing under stress. *Social Cognitive and Affective Neuroscience*, *7*(4), 476–484. https://doi.org/10.1093/scan/nsr026

Lincoln, S. H., Pisoni, A., Bondy, E., Kumar, P., Singleton, P., Hajcak, G., Pizzagalli, D. A., & Auerbach, R. P. (2019). Altered reward processing following an acute social stressor in adolescents. *PLoS ONE*, *14*(1), 1–13. https://doi.org/10.1371/journal.pone.0209361

Liston, C., Miller, M. M., Goldwater, D. S., Radley, J. J., Rocher, A. B., Hof, P. R., Morrison, J. H., & McEwen, B. S. (2006). Stress-induced alterations in prefrontal cortical dendritic morphology predict selective impairments in perceptual attentional set-shifting. *Journal of*

Neuroscience, *26*(30), 7870–7874. https://doi.org/10.1523/JNEUROSCI.1184-06.2006

Littman, M. L. (2015). Reinforcement learning improves behaviour from evaluative feedback. *Nature*, *521*(7553), 445–451. https://doi.org/10.1038/nature14540

Liu, J. J. W., Ein, N., Peck, K., Huang, V., Pruessner, J. C., & Vickers, K. (2017). Sex differences in salivary cortisol reactivity to the Trier Social Stress Test (TSST): A meta-analysis. *Psychoneuroendocrinology*, *82*, 26–37. https://doi.org/10.1016/J.PSYNEUEN.2017.04.007

Luo, P., Zhuang, M., Jie, J., Wu, X., & Zheng, X. (2018). State anxiety down-regulates empathic responses: Electrophysiological evidence. *Frontiers in Human Neuroscience*, *12*(December), 1–8. https://doi.org/10.3389/fnhum.2018.00502

Lupien, S. J., Maheu, F., Tu, M., Fiocco, A., & Schramek, T. E. (2007). The effects of stress and stress hormones on human cognition: Implications for the field of brain and cognition. *Brain and Cognition*, *65*(3), 209–237. https://doi.org/10.1016/j.bandc.2007.02.007

Lupien, S. J., McEwen, B. S., Gunnar, M. R., & Heim, C. (2009). Effects of stress throughout the lifespan on the brain, behaviour and cognition. *Nature Reviews Neuroscience*, *10*(6), 434–445. https://doi.org/10.1038/nrn2639

Macleod, C., & Mathews, A. (1988). Anxiety and the allocation of attention to threat Anxiety and the Allocation of Attention to Threat. *The Quarterly Journal of Experimental Psychology*, *40A*(4), 653–670. https://doi.org/10.1080/14640748808402292

Mansouri, F. A., Buckley, M. J., Mahboubi, M., & Tanaka, K. (2015). Behavioral consequences of selective damage to frontal pole and posterior cingulate cortices. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(29), E3940–E3949. https://doi.org/10.1073/pnas.1422629112

Marin, M. F., Lord, C., Andrews, J., Juster, R. P., Sindi, S., Arsenault-Lapierre, G., Fiocco, A. J., & Lupien, S. J. (2011). Chronic stress, cognitive functioning and mental health. *Neurobiology of Learning and Memory*, *96*(4), 583–595. https://doi.org/10.1016/j.nlm.2011.02.016

Mars, R. B., Debener, S., Gladwin, T. E., Harrison, L. M., Haggard, P., Rothwell, J. C., & Bestmann, S. (2008). Trial-by-trial fluctuations in the event-related electroencephalogram reflect dynamic changes in the degree of surprise. *Journal of Neuroscience*, *28*(47), 12539–12545. https://doi.org/10.1523/JNEUROSCI.2925-08.2008

Marshall, A. C., Cooper, N. R., Segrave, R., & Geeraert, N. (2015). The effects of long-term stress exposure on aging cognition: A behavioral and EEG investigation. *Neurobiology of Aging*, *36*(6), 2136–2144. https://doi.org/10.1016/j.neurobiolaging.2015.02.026

McCall, J. G., Al-Hasani, R., Siuda, E. R., Hong, D. Y., Norris, A. J., Ford, C. P., & Bruchas, M. R. (2015). CRH Engagement of the Locus Coeruleus Noradrenergic System Mediates Stress-Induced Anxiety. *Neuron*, *87*(3), 605–620. https://doi.org/10.1016/j.neuron.2015.07.002

McEwen, B. S. (2007). Physiology and neurobiology of stress and adaptation: Central role of the brain. *Physiological Reviews*, *87*(3), 873–904. https://doi.org/10.1152/physrev.00041.2006

McEwen, B. S., & Gianaros, P. J. (2011). Stress- and Allostasis-Induced Brain Plasticity. *Annual Review of Medicine*, *62*(1), 431–445. https://doi.org/10.1146/annurev-med-052209-100430

McEwen, B. S., & Wingfield, J. C. (2003). The concept of allostasis in biology and biomedicine. *Hormones and Behavior*, *43*(1), 2–15. https://doi.org/10.1016/S0018-506X(02)00024-7

Miller, R., Plessow, F., Kirschbaum, C., & Stalder, T. (2013). Classification criteria for distinguishing cortisol responders from nonresponders to psychosocial stress: Evaluation of

salivary cortisol pulse detection in panel designs. *Psychosomatic Medicine*, *75*(9), 832–840. https://doi.org/10.1097/PSY.0000000000000002

Miltner, W. H. R., Braun, C. H., & Coles, M. G. H. (1997). Event-Related Brain Potentials Following Incorrect Feedback i n a Time-Estimation Task: Evidence for a "Generic" Neural System for Error Detection. *Journal of Cognitive Neuroscience*, *9*(6), 788–798.

Miralles-Pechuán, L., Ponce, H., & Martínez-Villaseñor, L. (2018). A novel methodology for optimizing display advertising campaigns using genetic algorithms. *Electronic Commerce Research and Applications*, *27*, 39–51. https://doi.org/10.1016/j.elerap.2017.11.004

Morilak, D. A., Barrera, G., Echevarria, D. J., Garcia, A. S., Hernandez, A., Ma, S., & Petre, C. O. (2005). Role of brain norepinephrine in the behavioral response to stress. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *29*(8), 1214–1224. https://doi.org/10.1016/j.pnpbp.2005.08.007

Morris, L. (2005). The process of decision-making by stressed social workers: To stay or leave the workplace. *International Review of Psychiatry*, *17*(5), 347–354. https://doi.org/10.1080/09540260500238488

Morris, L. S., Baek, K., Kundu, P., Harrison, N. A., Frank, M. J., & Voon, V. (2016). Biases in the Explore-Exploit Tradeoff in Addictions: The Role of Avoidance of Uncertainty. *Neuropsychopharmacology*, *41*(4), 940–948. https://doi.org/10.1038/npp.2015.208

Nassar, M. R., Bruckner, R., & Frank, M. J. (2019). Statistical context dictates the relationship between feedback-related EEG signals and learning. *ELife*, *8*, 1–26. https://doi.org/10.7554/eLife.46975

Nassar, M. R., McGuire, J. T., Ritz, H., & Kable, J. W. (2019). Dissociable forms of uncertainty-driven representational change across the human brain. *Journal of Neuroscience*, *39*(9),

1688–1698. https://doi.org/10.1523/JNEUROSCI.1713-18.2018

Nater, U. M., La Marca, R., Florin, L., Moses, A., Langhans, W., Koller, M. M., & Ehlert, U. (2006). Stress-induced changes in human salivary alpha-amylase activity - Associations with adrenergic activity. *Psychoneuroendocrinology*, *31*(1), 49–58. https://doi.org/10.1016/j.psyneuen.2005.05.010

Nater, U. M., Rohleder, N., Schlotz, W., Ehlert, U., & Kirschbaum, C. (2007). Determinants of the diurnal course of salivary alpha-amylase. *Psychoneuroendocrinology*, *32*(4), 392–401. https://doi.org/10.1016/j.psyneuen.2007.02.007

Nielsen, S. E., Ahmed, I., & Cahill, L. (2014). Postlearning stress differentially affects memory for emotional gist and detail in naturally cycling women and women on hormonal contraceptives. *Behavioral Neuroscience*, *128*(4), 482–493. https://doi.org/10.1037/a0036687

Nieuwenhuis, S., Aston-Jones, G., & Cohen, J. D. (2005). Decision making, the P3, and the locus coeruleus-norepinephrine system. *Psychological Bulletin*, *131*(4), 510–532. https://doi.org/10.1037/0033-2909.131.4.510

Nieuwenhuis, S., Holroyd, C. B., Mol, N., & Coles, M. G. H. (2004). Reinforcement-related brain potentials from medial frontal cortex: Origins and functional significance. *Neuroscience and Biobehavioral Reviews*, *28*(4), 441–448. https://doi.org/10.1016/j.neubiorev.2004.05.003

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, *366*(6464), 447–453. https://doi.org/10.1126/science.aax2342

Ossewaarde, L., Qin, S., Van Marle, H. J. F., van Wingen, G. A., Fernández, G., & Hermans, E.

J. (2011). Stress-induced reduction in reward-related prefrontal cortex function. *NeuroImage*, *55*(1), 345–352. https://doi.org/10.1016/j.neuroimage.2010.11.068

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 of the United States of America*, *110*(52), 20941–20946. https://doi.org/10.1073/pnas.1312011110

Pacak, K., Palkovits, M., Kopin, I. J., & Goldstein, D. S. (1995). Stress-induced norepinephrine release in the hypothalamic paraventricular nucleus and pituitary-adrenocortical and sympathoadrenal activity: In vivo microdialysis studies. In *Frontiers in Neuroendocrinology* (Vol. 16, Issue 2, pp. 89–150). https://doi.org/10.1006/frne.1995.1004

Palminteri, S., Khamassi, M., Joffily, M., & Coricelli, G. (2015). Contextual modulation of value signals in reward and punishment learning. *Nature Communications*, *6*, 8096–9009. https://doi.org/10.1038/ncomms9096

Paul, M. (2019). *The Influence of Stress on Behavioral and Electrophysiological Correlates of Feedback-Based Learning (Doctoral dissertation)*. https://doi.org/https://doi.org/10.13154/294-6541

Paul, M., Bellebaum, C., Ghio, M., Suchan, B., & Wolf, O. T. (2020). Stress effects on learning and feedback-related neural activity depend on feedback delay. *Psychophysiology*, *57*(2), 1–18. https://doi.org/10.1111/psyp.13471

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. https://doi.org/10.1016/j.neuron.2013.04.037

Pearce, D., & Yamamoto, K. R. (1993). Mineralocorticoid and glucocorticoid receptor activities

distinguished by nonreceptor factors at a composite response element. *Science*, *259*(5098), 1161–1165. https://doi.org/10.1126/science.8382376

Pedersen, W. S., & Larson, C. L. (2016). State anxiety carried over from prior threat increases late positive potenital amplitude during an instructed emotion regulation task. *Emotion*, *16*(5), 719–729. https://doi.org/10.1037/emo0000154.State

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442.

Peng, H., Hu, B., Zheng, F., Fan, D., Zhao, W., Chen, X., Yang, Y., & Cai, Q. (2013). A method of identifying chronic stress by EEG. *Personal and Ubiquitous Computing*, *17*(7), 1341–1347. https://doi.org/10.1007/s00779-012-0593-3

Petzold, A., Plessow, F., Goschke, T., & Kirschbaum, C. (2010). Stress reduces use of negative feedback in a feedback-based learning task. *Behavioral Neuroscience*, *124*(2), 248–255. https://doi.org/10.1037/a0018930

Platt, M. L., & Huettel, S. A. (2008). Risky business: The neuroeconomics of decision making under uncertainty. *Nature Neuroscience*, *11*(4), 398–403. https://doi.org/10.1038/nn2062

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019

Polich, J. (2020). 50+ years of P300: Where are we now? *Psychophysiology*, *57*(7), e13616.

Polich, J., & Kok, A. (1995). Cognitive and biological determinants of P300: an integrative review. *Biological Psychology*, *41*(2), 103–146. https://doi.org/10.1016/0301-0511(95)05130-9

Porcelli, A. J., & Delgado, M. R. (2009). Acute stress modulates risk taking in financial decision making. *Psychological Science*, *20*(3), 278–283. https://doi.org/10.1111/j.1467-

9280.2009.02288.x

Porcelli, A. J., & Delgado, M. R. (2017). Stress and decision making: effects on valuation, learning, and risk-taking. *Current Opinion in Behavioral Sciences*, *14*, 33–39. https://doi.org/10.1016/j.cobeha.2016.11.015

Porcelli, A. J., Lewis, A. H., & Delgado, M. R. (2012). Acute stress influences neural circuits of reward processing. *Frontiers in Neuroscience*, *6*(NOV), 1–9. https://doi.org/10.3389/fnins.2012.00157

Preston, S. D., Buchanan, T. W., Stansfield, R. B., & Bechara, A. (2007). Effects of Anticipatory Stress on Decision Making in a Gambling Task. *Behavioral Neuroscience*, *121*(2), 257–263. https://doi.org/10.1037/0735-7044.121.2.257

Proudfit, G. H. (2015). The reward positivity: From basic research on reward to a biomarker for depression. *Psychophysiology*, *52*(4), 449–459. https://doi.org/10.1111/psyp.12370

Pruessner, J. C., Champagne, F., Meaney, M. J., & Dagher, A. (2004). Dopamine Release in Response to A Psychological Stress in Humans and Its Relationship to Early Life Maternal Care: A Positron Emission Tomography Study Using [11C] Raclopride. *Journal of Neuroscience*, *24*(11), 2825–2831. https://doi.org/10.1523/JNEUROSCI.3422-03.2004

Pruessner, J. C., Dedovic, K., Khalili-Mahani, N., Engert, V., Pruessner, M., Buss, C., Renwick, R., Dagher, A., Meaney, M. J., & Lupien, S. (2008). Deactivation of the Limbic System During Acute Psychosocial Stress: Evidence from Positron Emission Tomography and Functional Magnetic Resonance Imaging Studies. *Biological Psychiatry*, *63*(2), 234–240. https://doi.org/10.1016/j.biopsych.2007.04.041

Qin, S., Hermans, E. J., van Marle, H. J. F., Luo, J., & Fernández, G. (2009). Acute Psychological Stress Reduces Working Memory-Related Activity in the Dorsolateral

Prefrontal Cortex. *Biological Psychiatry*, *66*(1), 25–32. https://doi.org/10.1016/j.biopsych.2009.03.006

Redmond Jr., D. E., & Huang, Y. H. (1979). New evidence for a locus coeruleus-norepinephrine connection with anxiety. *Life Sciences*, *25*(26), 2149–2162.

Ree, M. J., French, D., MacLeod, C., & Locke, V. (2008). Distinguishing cognitive and somatic dimensions of state and trait anxiety: Development and validation of the State-Trait Inventory for Cognitive and Somatic Anxiety (STICSA). *Behavioural and Cognitive Psychotherapy*, *36*(3), 313–332. https://doi.org/10.1017/S1352465808004232

Reiche, E. M., Odebrecht, S., Nunes, V., & Morimoto, H. K. (2004). Stress, depression, the immune system, and cancer. *Oncology*, *5*, 617–625. https://doi.org/10.1016/S1470-2045(04)01597-9

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science*, *306*(5695), 443–447. https://doi.org/10.1126/science.1100301

Robinson, O. J., Overstreet, C., Charney, D. R., Vytal, K., & Grillon, C. (2013). Stress increases aversive prediction error signal in the ventral striatum. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(10), 4129–4133. https://doi.org/10.1073/pnas.1213923110

Rossi, V., & Pourtois, G. (2017). Someone's lurking in the dark : The role of state anxiety on attention deployment to threat-related stimuli. *Biological Psychology*, *122*, 21–32. https://doi.org/10.1016/j.biopsycho.2015.10.014

Sambrook, T. D., & Goslin, J. (2015). A Neural reward prediction error revealed by a meta-analysis of ERPs using great grand averages. *Psychological Bulletin*, *141*(1), 213–235.

https://doi.org/10.1037/bul0000006

San Martín, R. (2012). Event-related potential studies of outcome processing and feedback-guided learning. *Frontiers in Human Neuroscience*, *6*(OCTOBER 2012), 1–40. https://doi.org/10.3389/fnhum.2012.00304

Sänger, J., Bechtold, L., Schoofs, D., Blaszkewicz, M., & Wascher, E. (2014). The influence of acute stress on attention mechanisms and its electrophysiological correlates. *Frontiers in Behavioral Neuroscience*, *8*(OCT), 1–13. https://doi.org/10.3389/fnbeh.2014.00353

Sapolsky, R. M., Krey, L. C., & McEwen, B. S. (1986). The neuroendocrinology of stress and aging: the glucocorticoid cascade hypothesis. *Endocrine Reviews*, *7*(3), 284–301.

Sara, S. J. (2009). The locus coeruleus and noradrenergic modulation of cognition. *Nature Reviews Neuroscience*, *10*(3), 211–223. https://doi.org/10.1038/nrn2573

Saviola, F., Pappaianni, E., Monti, A., Grecucci, A., Jovicich, J., & Pisapia, N. De. (2020). Trait and state anxiety are mapped differently in the human brain. *Scientific Reports*, *10*(11112), 1–11. https://doi.org/10.1038/s41598-020-68008-z

Schuermann, B., Endrass, T., & Kathmann, N. (2012). Neural correlates of feedback processing in decision making under risk. *Frontiers in Human Neuroscience*, *6*(JUNE 2012), 1–10. https://doi.org/10.3389/fnhum.2012.00204

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science (New York, N.Y.)*, *275*(5306), 1593–1599. https://doi.org/10.1126/SCIENCE.275.5306.1593

Schulz, E., Bhui, R., Love, B. C., Brier, B., Todd, M. T., & Gershman, S. J. (2019). Structured, uncertainty-driven exploration in real-world consumer choice. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(28), 13903–13908.

https://doi.org/10.1073/pnas.1821028116

Schulz, E., Konstantinidis, E., & Speekenbrink, M. (2018). Putting bandits into context: How function learning supports decision making. *Journal of Experimental Psychology: Learning Memory and Cognition*, *44*(6), 927–943. https://doi.org/10.1037/xlm0000463

Schulz, P., Kirschbaum, C., Prüßner, J., & Hellhammer, D. (1998). Increased free cortisol secretion after awakening in chronically stressed individuals due to work overload. *Stress Medicine*, *14*(2), 91–97. https://doi.org/10.1002/(SICI)1099-1700(199804)14:2<91::AID-SMI765>3.0.CO;2-S

Schulz, P., & Scholtz, W. (1999). Trierer Inventar zur Erfassung von chronischem Sre (TICS): Skalenkonstruktion, teststatistische Überprüfung und Validierung der Skala Arbeitsüberlastung. *Diagnostica*.

Schwabe, L., Bohringer, A., Chatterjee, M., & Schachinger, H. (2008). *Effects of pre-learning stress on memory for neutral, positive and negative words: Different roles of cortisol and autonomic arousal*. https://doi.org/10.1016/j.nlm.2008.02.002

Schwabe, L., & Schächinger, H. (2018). Ten years of research with the Socially Evaluated Cold Pressor Test: Data from the past and guidelines for the future. *Psychoneuroendocrinology*, *92*(March), 155–161. https://doi.org/10.1016/j.psyneuen.2018.03.010

Schwabe, L., & Wolf, O. T. (2011). Stress-induced modulation of instrumental behavior: From goal-directed to habitual control of action. *Behavioural Brain Research*, *219*(2), 321–328. https://doi.org/10.1016/j.bbr.2010.12.038

Schwabe, L., & Wolf, O. T. (2013). Stress and multiple memory systems: From “thinking” to “doing.” *Trends in Cognitive Sciences*, *17*(2), 60. https://doi.org/10.1016/j.tics.2012.12.001

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions.

Journal of Neurology, Neurosurgery, and Psychiatry, *20*, 11–21. https://doi.org/10.1176/jnp.12.1.103-a

Selye, H. (1950). *The physiology and pathology of exposure to stress*. ACTA Inc.

Shackman, A. J., Maxwell, J. S., McMenamin, B. W., Greischar, L. L., & Davidson, R. J. (2011). Stress potentiates early and attenuates late stages of visual processing. *Journal of Neuroscience*, *31*(3), 1156–1161. https://doi.org/10.1523/JNEUROSCI.3384-10.2011

Shackman, A. J., Salomons, T. V., Slagter, H. A., Fox, A. S., Winter, J. J., & Davidson, R. J. (2011). The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nature Reviews Neuroscience*, *12*(3), 154–167. https://doi.org/10.1038/nrn2994

Shafiei, N., Gray, M., Viau, V., & Floresco, S. B. (2012). Acute stress induces selective alterations in cost/benefit decision-making. *Neuropsychopharmacology : Official Publication of the American College of Neuropsychopharmacology*, *37*(10), 2194–2209. https://doi.org/10.1038/npp.2012.69

Shields, G. S. (2020). Psychoneuroendocrinology Stress and cognition : A user's guide to designing and interpreting studies. *Psychoneuroendocrinology*, *112*(104475). https://doi.org/10.1016/j.psyneuen.2019.104475

Shields, G. S., Sazma, M. A., & Yonelinas, A. P. (2016). The effects of acute stress on core executive functions: A meta-analysis and comparison with cortisol. *Neuroscience and Biobehavioral Reviews*, *68*, 651–668. https://doi.org/10.1016/j.neubiorev.2016.06.038

Sinha, R. (2001). How does stress increase risk of drug abuse and relapse? *Psychopharmacology*, *158*(4), 343–359. https://doi.org/10.1007/s002130100917

Spencer, K., Dien, J., & Donchin, E. (2001). Spatiotemporal analysis of the late ERP to deviant stimuli. *Psychophysiology*, *38*, 343–358. https://doi.org/10.1111/1469-8986.3820343

Spielberger, C. D. (2010). State-Trait anxiety inventory. *The Corsini Encyclopedia of Psychology*, 1–1.

Squire, L. R. (2009). The Legacy of Patient H.M. for Neuroscience. *Neuron*, *61*(1), 6–9. https://doi.org/10.1016/j.neuron.2008.12.023

Starcevic, V., & Berle, D. (2006). Cognitive specificity of anxiety disorders: A review of selected key constructs. *Depression and Anxiety*, *23*, 51–61. https://doi.org/10.1002/da

Starcke, K., Agorku, J. D., & Brand, M. (2017). Exposure to Unsolvable Anagrams Impairs Performance on the Iowa Gambling Task. *Frontiers in Behavioral Neuroscience*, *11*, 114. https://doi.org/10.3389/fnbeh.2017.00114

Starcke, K., & Brand, M. (2012). Decision making under stress: A selective review. *Neuroscience and Biobehavioral Reviews*, *36*(4), 1228–1248. https://doi.org/10.1016/j.neubiorev.2012.02.003

Starcke, K., & Brand, M. (2016). Effects of stress on decisions under uncertainty: A meta-analysis. *Psychological Bulletin*, *142*(9), 909–933. https://doi.org/10.1037/bul0000060

Starcke, K., Wolf, O. T., Markowitsch, H. J., & Brand, M. (2008). Anticipatory stress influences decision making under explicit risk conditions. *Behavioral Neuroscience*, *122*(6), 1352–1360. https://doi.org/10.1037/a0013281

Sun, Y., Hunt, S., & Sah, P. (2015). Norepinephrine and Corticotropin-Releasing Hormone: Partners in the Neural Circuits that Underpin Stress and Anxiety. *Neuron*, *87*(3), 468–470. https://doi.org/10.1016/j.neuron.2015.07.022

Suor, J. H., Granros, M., Calentino, A. E., Luan Phan, K., & Burkhouse, K. L. (2021). The interplay of childhood maltreatment and maternal depression in relation to the reward positivity in youth. *Development and Psychopathology*, 1–11.

https://doi.org/10.1017/S0954579421000857

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed.). MIT Press.

Sutton, S., Braren, M., Zubin, J., & John, E. R. (1965). Evoked potential correlates of stimulus uncertainty. *Science*, *150*(3700), 1187–1188.

Takács, Á., Kóbor, A., Janacsek, K., & Honbolygó, F. (2015). High trait anxiety is associated with attenuated feedback-related negativity in risky decision making. *Neuroscience Letters*, *600*, 188–192. https://doi.org/10.1016/j.neulet.2015.06.022

Taylor, S. E. (2010). Mechanisms linking early life stress to adult health outcomes. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(19), 8507–8512. https://doi.org/10.1073/pnas.1003890107

Thoma, M. V., Kirschbaum, C., Wolf, J. M., & Rohleder, N. (2012). Acute stress responses in salivary alpha-amylase predict increases of plasma norepinephrine. *Biological Psychology*, *91*(3), 342–348. https://doi.org/10.1016/j.biopsycho.2012.07.008

Ulrich-Lai, Y. M., & Herman, J. P. (2009). Effects of Work-Related Stress: : University of Cambridge. *Nat Rev Neurosci*, *10*(6), 397–409. https://doi.org/10.1038/nrn2647.Neural

van Gerven, D. J. H., Ferguson, T., & Skelton, R. W. (2016). Acute stress switches spatial navigation strategy from egocentric to allocentric in a virtual Morris water maze. *Neurobiology of Learning and Memory*, *132*. https://doi.org/10.1016/j.nlm.2016.05.003

van Oort, J., Tendolkar, I., Hermans, E. J., Mulders, P. C., Beckmann, C. F., Schene, A. H., Fernández, G., & van Eijndhoven, P. F. (2017). How the brain connects in response to acute stress: A review at the human brain systems level. *Neuroscience and Biobehavioral Reviews*, *83*(October), 281–297. https://doi.org/10.1016/j.neubiorev.2017.10.015

Vineetha, R., Pai, K. M., Vengal, M., Gopalakrishna, K., & Narayanakurup, D. (2014). Usefulness of salivary alpha amylase as a biomarker of chronic stress and stress related oral mucosal changes - a pilot study. *Journal of Clinical and Experimental Dentistry*, *6*(2). https://doi.org/10.4317/jced.51355

Vining, R. F., McGinley, R. A., Maksvytis, J. J., & Ho, K. Y. (1983). Salivary cortisol: A better measure of adrenal cortical function than serum cortisol. *Annals of Clinical Biochemistry*, *20*(6), 329–335. https://doi.org/10.1177/000456328302000601

Vitaliano, P. P., Scanlan, J. M., Zhang, J., Savage, M. V., Hirsch, I. B., & Siegler, I. C. (2002). A path model of chronic stress, the metabolic syndrome, and coronary heart disease. *Psychosomatic Medicine*, *64*(3), 418–435. https://doi.org/10.1097/00006842-200205000-00006

Wagenmakers, E. J., & Farrell, S. (2004). AIC model selection using Akaike weights. *Psychonomic Bulletin and Review*, *11*(1), 192–196. https://doi.org/10.3758/BF03206482

Walker, A. R., Luque, D., Le Pelley, M. E., & Beesley, T. (2019). The role of uncertainty in attentional and choice exploration. *Psychonomic Bulletin and Review*, *26*(6), 1911–1916. https://doi.org/10.3758/s13423-019-01653-2

Weinberg, A., Ethridge, P., Ait Oumeziane, B., & Foti, D. (2019). Time-frequency analyses in event-related potential methodologies. In *Oxford Handbook of Human EEG Frequency Analysis*. Oxford University Press.

Weinberg, A., Riesel, A., & Proudfit, G. H. (2014). Show me the Money: The impact of actual rewards and losses on the feedback negativity. *Brain and Cognition*, *87*(1), 134–139. https://doi.org/10.1016/j.bandc.2014.03.015

Williams, C. C., Ferguson, T. D., Hassall, C. D., Abimbola, W., & Krigolson, O. E. (2021). The

ERP , frequency , and time – frequency correlates of feedback processing : Insights from a large sample study. *Psychophysiology*, *58*(e13722), 1–26. https://doi.org/10.1111/psyp.13722

Wilson, R. C., & Collins, A. G. E. (2019). Ten simple rules for the computational modeling of behavioral data. *ELife*, *8*(e49547), 1–33.

Wilson, R. C., Geana, A., White, J. M., Ludvig, E. A., & Cohen, J. D. (2014). Humans use directed and random exploration to solve the explore-exploit dilemma. *Journal of Experimental Psychology. General*, *143*(6), 2074–2081. https://doi.org/10.1037/a0038199

Wu, Y., & Zhou, X. (2009). The P300 and reward valence, magnitude, and expectancy in outcome evaluation. *Brain Research*, *1286*, 114–122. https://doi.org/10.1016/j.brainres.2009.06.032

Wüst, S., Federenko, I., Hellhammer, D. H., & Kirschbaum, C. (2000). Genetic factors, perceived chronic stress, and the free cortisol response to awakening. *Psychoneuroendocrinology*, *25*(7), 707–720. https://doi.org/10.1016/S0306-4530(00)00021-4

Xia, L., Gu, R., Zhang, D., & Luo, Y. (2017). Anxious individuals are impulsive decision-makers in the delay discounting task: An ERP study. *Frontiers in Behavioral Neuroscience*, *11*(January), 1–11. https://doi.org/10.3389/fnbeh.2017.00005

Yasui, N., Lim, S., Linke, C., White, A., & White, M. (2019). An Empirical and Conceptual Categorization of Value-based Exploration Methods. *36th International Conference on Machine Learning*, 1–5.

Yeung, N., & Sanfey, A. G. (2004). Independent coding of reward magnitude and valence in the human brain. *Journal of Neuroscience*, *24*(28), 6258–6264.

https://doi.org/10.1523/JNEUROSCI.4537-03.2004

Yu, A. J., & Dayan, P. (2003). Expected and unexpected uncertainty: ACh and NE in the neocortex. *Advances in Neural Information Processing Systems*.

Yu, A. J., & Dayan, P. (2005). Uncertainty, neuromodulation, and attention. *Neuron*, *46*(4), 681–692. https://doi.org/10.1016/j.neuron.2005.04.026

Yu, J. C., Fiore, V. G., Briggs, R. W., Braud, J., Rubia, K., Adinoff, B., & Gu, X. (2020). An insula-driven network computes decision uncertainty and promotes abstinence in chronic cocaine users. *European Journal of Neuroscience*, *52*(12), 4923–4936. https://doi.org/10.1111/ejn.14917

Zajkowski, W. K., Kossut, M., & Wilson, R. C. (2017). A causal role for right frontopolar cortex in directed , but not random , exploration. *ELife*, *6*(e27430), 1–18.

Zhang, L., Wang, K., Zhu, C., Yu, F., & Chen, X. (2015). Trait anxiety has effect on decision making under ambiguity but not decision making under risk. *PLoS ONE*, *10*(5), 1–21. https://doi.org/10.1371/journal.pone.0127189

Zhang, X., Li, P., Chen, J., & Li, H. (2020). Acute stress impairs reward positivity effect in probabilistic learning. *Psychophysiology*, *57*(4), 1–11. https://doi.org/10.1111/psyp.13531

Zimmer, F., Scheibe, K., Stock, M., & Stock, W. G. (2019). Fake news in social media: Bad algorithms or biased users? *Journal of Information Science Theory and Practice*, *7*(2), 40–53. https://doi.org/10.1633/JISTaP.2019.7.2.4

Appendix A

Trier Inventory for the Assessment of Chronic Stress

On the following pages you will find descriptions of situations and experiences. Please answer how often each event has happened to you in the past 3 months, using a rating of never, rarely, sometimes, often or very often.

Choose the ratings as follows:

0	= Never	(I have not experienced it in the past 3 months)
1	= Rarely	(I have rarely experienced it in the past 3 months)
2	= Sometimes	(I have sometimes experienced it in the past 3 months)
3	= Often	(I have experienced it often in the past 3 months)
4	= Very Often	(I experienced it very often in the past 3 months)

Please answer all items in order, without skipping any. Some statements may sound similar to others, but please answer them anyway. There is no need to rush. Take your time and think carefully about each answer

1. I have to postpone much needed rest
   (0) Never (3) Often
   (1) Rarely (4) Very Often
   (2) Sometimes
2. I receive too little appreciation for my accomplishments
   (0) Never (3) Often
   (1) Rarely (4) Very Often
   (2) Sometimes
3. I make too many mistakes because of what I have to do demands too much of me
   (0) Never (3) Often
   (1) Rarely (4) Very Often
   (2) Sometimes
4. I do not have enough time to perform my daily tasks
   (0) Never (3) Often
   (1) Rarely (4) Very Often
   (2) Sometimes
5. I must perform tasks that seem nonsensical to me
   (0) Never (3) Often
   (1) Rarely (4) Very Often
   (2) Sometimes
6. I have differences of opinion that lead to tension
   (0) Never (3) Often
   (1) Rarely (4) Very Often
   (2) Sometimes
7. I have work to do that involves carrying a lot of responsibility for other people
   (0) Never (3) Often
   (1) Rarely (4) Very Often
   (2) Sometimes
8. Situations in which I must make an effort to win other's trust
   (0) Never (3) Often
   (1) Rarely (4) Very Often
   (2) Sometimes
9. I worry that something unpleasant may happen
   (0) Never (3) Often
   (1) Rarely (4) Very Often

(2) Sometimes

10. My daily tasks are not interesting
    (0) Never (3) Often
    (1) Rarely (4) Very Often
    (2) Sometimes
11. Times when I am lonely
    (0) Never (3) Often
    (1) Rarely (4) Very Often
    (2) Sometimes
12. Situations when I must take pains to have a good relationship with others
    (0) Never (3) Often
    (1) Rarely (4) Very Often
    (2) Sometimes
13. I have to perform tasks that I don't enjoy
    (0) Never (3) Often
    (1) Rarely (4) Very Often
    (2) Sometimes
14. I have tasks to perform during which I am being critically observed
    (0) Never (3) Often
    (1) Rarely (4) Very Often
    (2) Sometimes
15. I have conflicts with others because they have different goals
    (0) Never (3) Often
    (1) Rarely (4) Very Often
    (2) Sometimes
16. Times when I cannot suppress worrisome thoughts
    (0) Never (3) Often
    (1) Rarely (4) Very Often
    (2) Sometimes
17. Times when so many appointments accumulate that I can barely get caught up
    (0) Never (3) Often
    (1) Rarely (4) Very Often
    (2) Sometimes
18. I try in vain to gain recognition for my good work
    (0) Never (3) Often
    (1) Rarely (4) Very Often
    (2) Sometimes
19. I spend a lot of time dealing with other peoples' problems
    (0) Never (3) Often
    (1) Rarely (4) Very Often
    (2) Sometimes
20. I perform my tasks inadequately, despite trying my best
    (0) Never (3) Often
    (1) Rarely (4) Very Often
    (2) Sometimes
21. Times when none of my tasks seem meaningful to me
    (0) Never (3) Often
    (1) Rarely (4) Very Often
    (2) Sometimes
22. I have work to do that must not disappoint others
    (0) Never (3) Often
    (1) Rarely (4) Very Often
    (2) Sometimes
23. I have to try and make a good impression on people

(0) Never (3) Often
(1) Rarely (4) Very Often
(2) Sometimes

24. Times when I can no longer cope with the demands of my work
(0) Never (3) Often
(1) Rarely (4) Very Often
(2) Sometimes
25. Times when my worries overwhelm me
(0) Never (3) Often
(1) Rarely (4) Very Often
(2) Sometimes
26. I have conflicts with others because I do not act the way they expect me to
(0) Never (3) Often
(1) Rarely (4) Very Often
(2) Sometimes
27. Times when I must work under strict deadlines
(0) Never (3) Often
(1) Rarely (4) Very Often
(2) Sometimes
28. I have to deal with other peoples problems too much
(0) Never (3) Often
(1) Rarely (4) Very Often
(2) Sometimes
29. Times when I do not have the opportunity to share my thoughts and feelings with others
(0) Never (3) Often
(1) Rarely (4) Very Often
(2) Sometimes
30. Situations in which it depends entirely on me if a relationship with another person develops satisfactorily
(0) Never (3) Often
(1) Rarely (4) Very Often
(2) Sometimes
31. Although I do my best, my work is not appreciated
(0) Never (3) Often
(1) Rarely (4) Very Often
(2) Sometimes
32. I have tasks to fulfill that pressure me to prove myself
(0) Never (3) Often
(1) Rarely (4) Very Often
(2) Sometimes
33. I have conflicts with others because they meddle too much in my affairs
(0) Never (3) Often
(1) Rarely (4) Very Often
(2) Sometimes
34. Times when I am isolated from other people
(0) Never (3) Often
(1) Rarely (4) Very Often
(2) Sometimes
35. Times when I am not able to perform as well as expected
(0) Never (3) Often
(1) Rarely (4) Very Often
(2) Sometimes
36. Times when I worry a lot and cannot stop
(0) Never (3) Often
(1) Rarely (4) Very Often

(2) Sometimes

37. I object to duties that I must fulfill
    (0) Never (3) Often
    (1) Rarely (4) Very Often
    (2) Sometimes
38. Times when I have too many duties to fulfill
    (0) Never (3) Often
    (1) Rarely (4) Very Often
    (2) Sometimes
39. I must frequently care for the well-being of others
    (0) Never (3) Often
    (1) Rarely (4) Very Often
    (2) Sometimes
40. Situations in which I must make an effort to please others
    (0) Never (3) Often
    (1) Rarely (4) Very Often
    (2) Sometimes
41. Times when I have nothing meaningful to do
    (0) Never (3) Often
    (1) Rarely (4) Very Often
    (2) Sometimes
42. Times when I have too little contact with other people
    (0) Never (3) Often
    (1) Rarely (4) Very Often
    (2) Sometimes
43. People have high expectations for the tasks that I must fulfill
    (0) Never (3) Often
    (1) Rarely (4) Very Often
    (2) Sometimes
44. Times that my work overwhelms me
    (0) Never (3) Often
    (1) Rarely (4) Very Often
    (2) Sometimes
45. I have arguments with people that lead to long-lasting conflicts
    (0) Never (3) Often
    (1) Rarely (4) Very Often
    (2) Sometimes
46. I am not adequately rewarded for my efforts
    (0) Never (3) Often
    (1) Rarely (4) Very Often
    (2) Sometimes
47. I worry that I will not be able to fulfill my tasks
    (0) Never (3) Often
    (1) Rarely (4) Very Often
    (2) Sometimes
48. I must do work that does not take advantage of my abilities
    (0) Never (3) Often
    (1) Rarely (4) Very Often
    (2) Sometimes
49. Situations in which the well-being of others depends on how well I work
    (0) Never (3) Often
    (1) Rarely (4) Very Often
    (2) Sometimes
50. I have too many tasks to perform

(0) Never (3) Often
(1) Rarely (4) Very Often
(2) Sometimes

51. Times when I miss having contact with others
(0) Never (3) Often
(1) Rarely (4) Very Often
(2) Sometimes
52. I have unnecessary conflict with others
(0) Never (3) Often
(1) Rarely (4) Very Often
(2) Sometimes
53. Times when I have no tasks that make me happy
(0) Never (3) Often
(1) Rarely (4) Very Often
(2) Sometimes
54. I experience having too much to do
(0) Never (3) Often
(1) Rarely (4) Very Often
(2) Sometimes
55. Although I try, I do not fulfill my duties as I should
(0) Never (3) Often
(1) Rarely (4) Very Often
(2) Sometimes
56. Times when I have no friends to do things with
(0) Never (3) Often
(1) Rarely (4) Very Often
(2) Sometimes
57. Times when my responsibilities for others becomes a burden to me
(0) Never (3) Often
(1) Rarely (4) Very Often
(2) Sometimes

Positive and Negative Affect Schedule

This scale consists of a number of words that describe different feelings and emotions. Read each item and then mark the appropriate answer in the space next to that word. Indicate to what extent you feel this right now, that is, at the present moment. Use the following scale to record your answers

1. Interested.
   (1) Very slightly or not at all (4) Quite a bit
   (2) A Little (5) Extremely
   (3) Moderately
2. Distressed
   (1) Very slightly or not at all (4) Quite a bit
   (2) A Little (5) Extremely
   (3) Moderately
3. Excited
   (1) Very slightly or not at all (4) Quite a bit
   (2) A Little (5) Extremely
   (3) Moderately
4. Upset
   (1) Very slightly or not at all (4) Quite a bit
   (2) A Little (5) Extremely
   (3) Moderately
5. Strong
   (1) Very slightly or not at all (4) Quite a bit
   (2) A Little (5) Extremely
   (3) Moderately
6. Guilty
   (1) Very slightly or not at all (4) Quite a bit
   (2) A Little (5) Extremely
   (3) Moderately
7. Scared
   (1) Very slightly or not at all (4) Quite a bit
   (2) A Little (5) Extremely
   (3) Moderately
8. Hostile
   (1) Very slightly or not at all (4) Quite a bit
   (2) A Little (5) Extremely
   (3) Moderately
9. Enthusiastic
   (1) Very slightly or not at all (4) Quite a bit
   (2) A Little (5) Extremely
   (3) Moderately
10. Proud
   (1) Very slightly or not at all (4) Quite a bit
   (2) A Little (5) Extremely
   (3) Moderately
11. Irritable
   (1) Very slightly or not at all (4) Quite a bit
   (2) A Little (5) Extremely
   (3) Moderately
12. Alert
   (1) Very slightly or not at all (4) Quite a bit
   (2) A Little (5) Extremely
   (3) Moderately

13. Ashamed
    (1) Very slightly or not at all (4) Quite a bit
    (2) A Little (5) Extremely
    (3) Moderately
14. Inspired
    (1) Very slightly or not at all (4) Quite a bit
    (2) A Little (5) Extremely
    (3) Moderately
15. Nervous
    (1) Very slightly or not at all (4) Quite a bit
    (2) A Little (5) Extremely
    (3) Moderately
16. Determined
    (1) Very slightly or not at all (4) Quite a bit
    (2) A Little (5) Extremely
    (3) Moderately
17. Attentive
    (1) Very slightly or not at all (4) Quite a bit
    (2) A Little (5) Extremely
    (3) Moderately
18. Jittery
    (1) Very slightly or not at all (4) Quite a bit
    (2) A Little (5) Extremely
    (3) Moderately
19. Active
    (1) Very slightly or not at all (4) Quite a bit
    (2) A Little (5) Extremely
    (3) Moderately
20. Afraid
    (1) Very slightly or not at all (4) Quite a bit
    (2) A Little (5) Extremely
    (3) Moderately

State Trait Anxiety Inventory – State Scale

A number of statements which people have used to describe themselves are given below. Read each statement and then select the appropriate option to the statement to indicate how you feel *Right* now, that is, at *This Moment*. There are no right or wrong answers. Do not spend too much time on any one statement but give the answer which seems to describe your present feelings best.

1. I feel calm
   (1) Not at all (3) Moderately So
   (2) Somewhat (4) Very Much So
2. I feel secure
   (1) Not at all (3) Moderately So
   (2) Somewhat (4) Very Much So
3. I am tense
   (1) Not at all (3) Moderately So
   (2) Somewhat (4) Very Much So
4. I feel strained
   (1) Not at all (3) Moderately So
   (2) Somewhat (4) Very Much So
5. I feel at ease
   (1) Not at all (3) Moderately So
   (2) Somewhat (4) Very Much So
6. I feel upset
   (1) Not at all (3) Moderately So
   (2) Somewhat (4) Very Much So
7. I am presently worrying over possible misfortunes
   (1) Not at all (3) Moderately So
   (2) Somewhat (4) Very Much So
8. I feel satisfied
   (1) Not at all (3) Moderately So
   (2) Somewhat (4) Very Much So
9. I feel frightened
   (1) Not at all (3) Moderately So
   (2) Somewhat (4) Very Much So
10. I feel comfortable
   (1) Not at all (3) Moderately So
   (2) Somewhat (4) Very Much So
11. I feel self-confident
   (1) Not at all (3) Moderately So
   (2) Somewhat (4) Very Much So
12. I feel nervous
   (1) Not at all (3) Moderately So
   (2) Somewhat (4) Very Much So
13. I am jittery
   (1) Not at all (3) Moderately So
   (2) Somewhat (4) Very Much So
14. I feel indecisive
   (1) Not at all (3) Moderately So
   (2) Somewhat (4) Very Much So
15. I am relaxed
   (1) Not at all (3) Moderately So
   (2) Somewhat (4) Very Much So
16. I feel content
   (1) Not at all (3) Moderately So

(2) Somewhat (4) Very Much So

17. I am worried
    (1) Not at all (3) Moderately So
    (2) Somewhat (4) Very Much So
18. I feel confused
    (1) Not at all (3) Moderately So
    (2) Somewhat (4) Very Much So
19. I feel steady
    (1) Not at all (3) Moderately So
    (2) Somewhat (4) Very Much So
20. I feel pleasant
    (1) Not at all (3) Moderately So
    (2) Somewhat (4) Very Much So

State-Trait Anxiety Inventory – Trait Scale

A number of statements which people have used to describe themselves are given below. Read each statement and then select the appropriate number to the statement to indicates how your Generally feel. There are no right or wrong answer. Do not spend too much time on any one statement but give the answer which seems to describe how you generally feel.

1. I feel pleasant
   (1) Not at all (3) Moderately So
   (2) Somewhat (4) Very Much So
2. I feel nervous and restless
   (1) Not at all (3) Moderately So
   (2) Somewhat (4) Very Much So
3. I feel satisfied with myself
   (1) Not at all (3) Moderately So
   (2) Somewhat (4) Very Much So
4. I wish I could be as happy as others seem to be
   (1) Not at all (3) Moderately So
   (2) Somewhat (4) Very Much So
5. I feel like a failure
   (1) Not at all (3) Moderately So
   (2) Somewhat (4) Very Much So
6. I feel rested
   (1) Not at all (3) Moderately So
   (2) Somewhat (4) Very Much So
7. I am “calm, cool, and collected”
   (1) Not at all (3) Moderately So
   (2) Somewhat (4) Very Much So
8. I feel that difficulties are piling up so that I cannot overcome them
   (1) Not at all (3) Moderately So
   (2) Somewhat (4) Very Much So
9. I worry too much over something that does not really matter
   (1) Not at all (3) Moderately So
   (2) Somewhat (4) Very Much So
10. I am happy
   (1) Not at all (3) Moderately So
   (2) Somewhat (4) Very Much So
11. I have disturbing thoughts
   (1) Not at all (3) Moderately So
   (2) Somewhat (4) Very Much So
12. I lack self-confidence
   (1) Not at all (3) Moderately So
   (2) Somewhat (4) Very Much So
13. I feel secure
   (1) Not at all (3) Moderately So
   (2) Somewhat (4) Very Much So
14. I make decisions easily
   (1) Not at all (3) Moderately So
   (2) Somewhat (4) Very Much So
15. I feel inadequate
   (1) Not at all (3) Moderately So
   (2) Somewhat (4) Very Much So
16. I am content
   (1) Not at all (3) Moderately So

(2) Somewhat (4) Very Much So

17. Some unimportant thought runs through my mind and it bothers me
    (1) Not at all (3) Moderately So
    (2) Somewhat (4) Very Much So
18. I take disappointments too keenly and I can't put them out of my mind
    (1) Not at all (3) Moderately So
    (2) Somewhat (4) Very Much So
19. I am a steady person
    (1) Not at all (3) Moderately So
    (2) Somewhat (4) Very Much So
20. I get in a state of tension or turmoil as I think over my recent concerns and interests
    (1) Not at all (3) Moderately So
    (2) Somewhat (4) Very Much So

Appendix B

Trier Social Stress Test – Stress Version

Participant is led to interview room where two experimenters who the participant has not yet met are present and are seated at a table.

Experimenter: *"Your task in this part of the experiment is the following: please imagine that you have applied for your ideal job and have been invited for an interview. You must now convince the panel members why you are the ideal candidate by giving a talk. You will have three minutes to prepare a talk to convince the panel, before having five minutes to present the talk. Please note that you will be recorded by a camera for subsequent voice and behavioural analysis.*

This is the "selection panel" (introduce panel). This selection panel has been trained to monitor your behaviour and will take notes during your talk. You should try to leave the best possible impression and assume the role of the applicant for the duration of the talk as best as you can. The panel will reserve the right to ask follow-up questions in case of uncertainties to receive all necessary information from you. Following your talk, you will be given a second task by the panel, which will only be explained to you by the panel.

You may take some notes now, which you must not use during your talk. Do you have any questions?

You now have three minutes to prepare your speech. There is a pen and piece of paper on the table for your use."

After three minutes has elapsed, the Chair begins video recording and hits start on the timer and says:

> **Chair:** *"Please begin your talk. You may not use the notes you have made."*

If participant stops before 5 minutes, wait for about 20 seconds and the chair then says:

> **Chair:** *"You still have time left. Please continue".*

If participant does not continue, start asking questions after 10 seconds:

Questions to ask the research participant during the "job interview"

1. *What are your personal strengths?*
2. *What are your major weaknesses?*
3. *Why do you think you are especially well-qualified for this task?*
4. *Why do you think you are better qualified then the other applicants?*
5. *You just mentioned your qualities in respect to…, what do you in particular think about…?*
6. *You just spoke about…, what exactly do you then think about…?*
7. *What kind of leading qualities do you have?*
8. *What do you think about teamwork?*
9. *Where do you see your position in a team?*
10. *What can you constructively add to a team?*
11. *What do your employees appreciate about you most?*
12. *Would you be willing to work on the weekends if this be deemed necessary?*
13. *What kind of qualities to you expect from your co-workers?*
14. *Under what circumstances would you be willing to compensate for the mistakes your co-workers make?*
15. *What do your family/friends especially appreciate about you?*
16. *Please complete the following sentence: "I am the best at/in…"*

When the five minutes is up, let them finish their sentence, and then say:

> **Chair:** *"Stop, the interview is now over. We now want you to solve a calculation task. This task is unrelated to the job interview. Please subtract the number 17 continually, starting at 2023. Please do this as quickly as possible. Should you make a mistake, we will point out your mistake and you must start over again. Do you have any questions?"*

Chair should mark down errors on the paper. If the participant looks for whether they are correct or not, simply nod for correct answers. Should the participant miscalculate say:

> **Chair:** *"Stop - mistake. Start over at 2023 please."*

2023	1683	1343	1003	663	323
2006	1666	1326	986	646	306
1989	1649	1309	969	629	289
1972	1632	1292	952	612	272
1955	1615	1275	935	595	255
1938	1598	1258	918	578	238
1921	1581	1241	901	561	221
1904	1564	1224	884	544	204

1887	1547	1207	867	527	187
1870	1530	1190	850	510	170
1853	1513	1173	833	493	153
1836	1496	1156	816	476	136
1819	1479	1139	799	459	119
1802	1462	1122	782	442	102
1785	1445	1105	765	425	85
1768	1428	1088	748	408	68
1751	1411	1071	731	391	51
1734	1394	1054	714	374	34
1717	1377	1037	697	357	17
1700	1360	1020	680	340	0

After five minutes has elapsed:

> **Chair:** *"Thank you for your time. Please leave the room where the experimenters will be waiting. You will then complete the rest of the experimental session and will be given a full debrief at the end of the experiment."*

Trier Social Stress Test – Placebo Version

Participant is led by the experimenter to an empty room where they are instructed:

> **Experimenter:** "*You are now going to have three minutes to think about a talk about a recent movie, novel, or holiday. After the three minutes is up, you will have to talk aloud for 5 minutes about the topic of interest. You will not be giving the talk to anyone and you will not be recorded. Just take the time to think about what you would like to say about the topic of interest."*

The experimenter leaves the room and returns after three minutes has elapsed:

> **Experimenter:** "*You will now have five minutes to start your talk. Again, there is no recording, and no one is listening."*

The experimenter leaves the room and returns after five minutes has elapsed:

> **Experimenter:** *"Okay please start counting up in steps of 15, starting at 0. This will last for five minutes."*

Experimenter then leaves the room and returns after five minutes has elapsed:

> **Experimenter:** *"What number did you get to?"*