# Where not what: the role of spatial-motor processing in decision-making

# WHERE NOT WHAT: THE ROLE OF SPATIAL-MOTOR PROCESSING IN DECISION-MAKING

BY

PARKER J. BANKS, B.A.Sc.

A THESIS

SUBMITTED TO THE DEPARTMENT OF PSYCHOLOGY, NEUROSCIENCE, & BEHAVIOUR

AND THE SCHOOL OF GRADUATE STUDIES

OF MCMASTER UNIVERSITY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

© Copyright by Parker J. Banks, June 2021

All Rights Reserved

Doctor of Philosophy (2021) McMaster University

(Psychology, Neuroscience, & Behaviour) Hamilton, Ontario, Canada

| | |
|---|---|
| TITLE: | Where not what: the role of spatial-motor processing in decision-making |
| AUTHOR: | Parker J. Banks<br>B.A.Sc., (Music & Psychology)<br>University of Lethbridge, Lethbridge, Alberta |
| SUPERVISOR: | Dr. Patrick J. Bennett & Dr. Allison B. Sekuler |
| NUMBER OF PAGES: | xvii, 202 |

*To my wife Livia, for the unending support, love, and hard work that allowed me to pursue a PhD. To my sons Sebastian and Donovan, my greatest accomplishments to date. And to anyone who continues to work towards an understanding of the human brain, one of the last great frontiers.*

# Abstract

Decision-making is comprised of an incredibly varied set of behaviours. However, all vertebrates tend to repeat previously rewarding actions and avoid those that have led to loss, behaviours known collectively as the win-stay, lose-shift strategy. This response strategy is supported by the sensorimotor striatum and nucleus accumbens, structures also implicated in spatial processing and the integration of sensory information in order to guide motor action. Therefore, choices may be represented as spatial-motor actions whose value is determined by the rewards and punishments associated with that action. In this dissertation I demonstrate that the *location* of choices relative to previous rewards and punishments, rather than their identities, determines their value. Chapters 2 and 4 demonstrate that the location of rewards and punishments drives future decisions to win-stay or lose-shift towards that location. Even when choices differ in colour or shape, choice value is determined by location, not visual identity. Chapter 3 compares decision-making when two, six, twelve, or eighteen choices are present, finding that the value of a win or loss is not tied to a single location, but is distributed throughout the choice environment. Finally, Chapter 5 provides anatomical support for the spatial-motor basis of choice. Specifically, win-stay responses are associated with greater oscillatory activity than win-shift responses in the motor cortex corresponding to the hand used to make a

choice, whereas lose-shift responses are accompanied by greater activation of frontal systems compared to lose-stay responses. The win-stay and lose-shift behaviours activate structures known to project to different regions of the striatum. Overall, this dissertation provides behavioural evidence that choice location, not visual identity, determines choice value.

# Acknowledgements

Thank you to Patrick Bennett and Allison Sekuler for working as my advisors and for the numerous hours spent revising this dissertation. You gave structure my ideas, allowing me to finish this dissertation. You gave me the resources I needed to accomplish my research, and you otherwise gave me free reign, leaving me to pursue the questions I thought important. If you had been anyone else, I would not have been so free to pursue a job, house, family, and children at the same time as a PhD.

Thank you to Aaron Gruber. It was my undergraduate work with you that provided the inspiration for this thesis, and the continuing collaboration that provided the technical knowledge I needed to ask the right questions.

And thank you to Donna Waxman and Hailey Wright. Without the hundreds of hours you both spent recruiting and running participants, my research would not have been nearly as extensive, or of sufficient quality to build my dissertation on.

# Contents

# List of Figures

# Chapter 1

# General Introduction

Humans are remarkably efficient at decision-making in an unpredictable, and often hostile world [Summerfield and Tsetsos, 2015]. In particular, our tendency to repeat rewarding decisions and avoid those that have led to punishment forms an efficient choice strategy in changing environments. This win-stay, lose-shift strategy was first codified by Thorndike [1898, 1911] in the nineteenth century. Since then, animal models have informed our understanding of the anatomical basis of theses strategies [Brush et al., 1961, Olds, 1962]. However, research on the neural basis of decision-making has only recently been extended to humans, with the advent of high-resolution neuroimaging technologies.

Animal studies have highlighted the importance of the sensorimotor striatum to the win-stay and lose-shift responses, including the putamen, caudate nucleus, and nucleus accumbens [Packard et al., 1989, Packard and White, 1991, McDonald and White, 2013, Skelin et al., 2014, Gruber et al., 2017, Grospe et al., 2018, Thapa and Gruber, 2018]. Each of these regions integrate information from across the limbic, associative, and sensorimotor systems in order to guide selection of spatial-motor

actions [Murray et al., 2011, Northcutt, 2008]. This dual role in value-based decision strategies and spatial-motor action suggests the brain represents choices as spatial-motor actions and not abstract symbols. Consequently, altering the location or motor action required to make a choice should dissociate choice value from its identity, particularly as it relates to the win-stay and lose-shift behaviours.

In this dissertation I provide evidence from behaviour, computational models, and electroencephalography demonstrating that choice value is represented in spatial coordinates. Specifically, I used the Matching Pennies task to measure decision-making performance in a competitive environment, analyzed behavioural data in this task with the *Q-learning* model [Sutton and Barto, 1998], which is known to accurately represent dopamine activity of some neurons in the striatum [Schultz et al., 1997, Glimcher, 2011], and used EEG to investigate the neural correlates of reward-processing and choice.

## 1.1 The Anatomy of Choice

The striatum is divided into three major sub-components that are important for decision-making: the nucleus accumbens (NAc), caudate nucleus (dorsomedial striatum), and putamen (dorsolateral striatum). These regions are necessary components of the cortico-striatal-thalamic loop circuits that interpret sensory information and use it to guide behaviour. These circuits guide motor action [Groenewegen, 2003], drive motivation [Cardinal et al., 2002], support memory for motor sequences [Albouy et al., 2008], spatial processing [van der Meer et al., 2010], reward expectation [Apicella et al., 1992], and goal-directed control of action-outcome associations [Fuccillo, 2016, Yin et al., 2005]. They also govern the win-stay and lose-shift responses. For

example, the lose-shift has been associated with function of the dorsomedial [Grospe et al., 2018, Skelin et al., 2014], ventrolateral [Thapa and Gruber, 2018], and dorsolateral [Packard et al., 1989, Gruber et al., 2017] striatum (i.e., the putamen), the insula [Danckert et al., 2011], dorsal hippocampus [Chen et al., 2012], and anterior cingulate [Paulus et al., 2002a] in a variety of experimental contexts. The win-stay has been found to rely on the nucleus accumbens [Gruber et al., 2017], dorsolateral striatum [Packard and White, 1991, McDonald and White, 2013], fimbria-fornix [Packard et al., 1989], pedunculopontine tegmental nucleus [Syed et al., 2016], lateral habenula [Thapa et al., 2019], anterior cingulate, and inferior prefrontal cortex [Paulus et al., 2002a]. However, the importance of these different regions varies with task modality and demands.

In order to regulate motor action selection each of these three striatal sub-regions, the putamen, caudate nucleus, and nucleus accumbens, receive a variety of cortical and sub-cortical inputs. The precuneus and the somatosensory, motor, orbitofrontal, and insular cortices provide excitatory inputs to the putamen [Malach and Graybiel, 1986, Brasted et al., 1999, Cavanna and Trimble, 2006, Pan et al., 2010]. The nuclues accumbens receives inputs from the hippocampus, profrontal cortex, amygdala, and hypothalamus [Gerfen, 1984, Voorn et al., 2004, Kelley et al., 1982]. Finally, the hippocampus, amygdala, anterior cingulate, precuneus and visual cortex project to the caudate nucleus [Khibnik et al., 2014, Fuccillo, 2016, Cavanna and Trimble, 2006]. A summary of these connections is provided in Figure 1.1.

Because each of these structures integrate different sets of cortical inputs, they represent spatial-motor action in different reference frames that are linked to different decision-making strategies. The putamen receives direct inputs from motor and

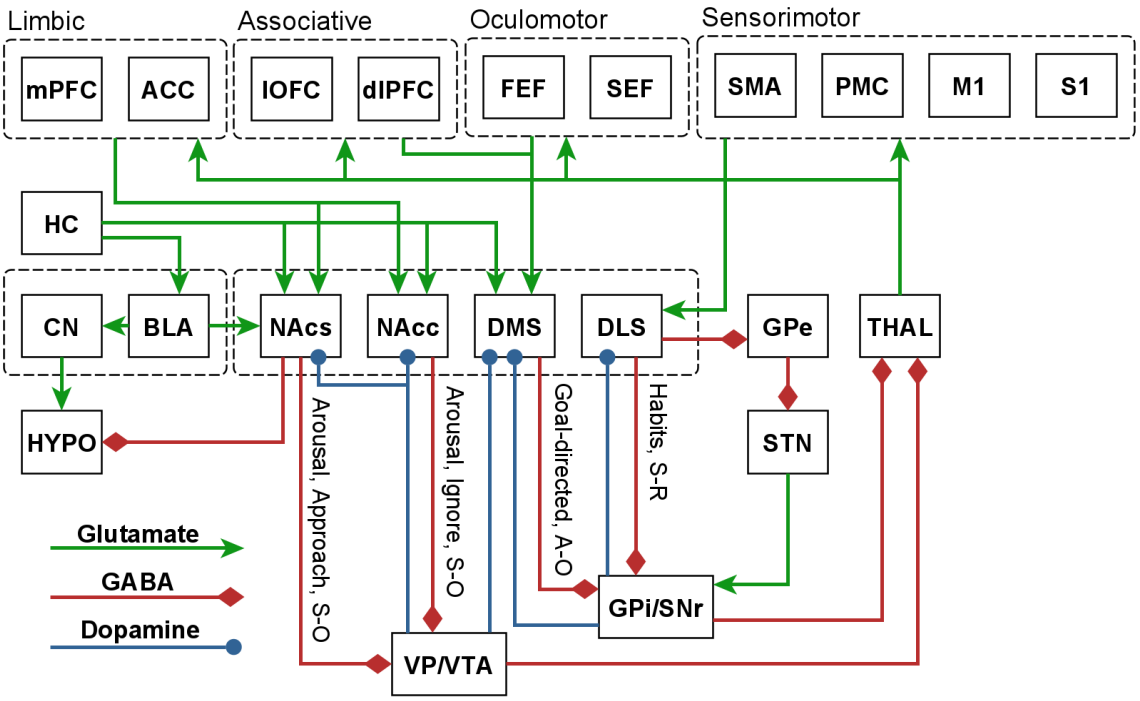


Figure 1.1: Simplified diagram of circuitry comprising the cortico-striatal-thalamic loops. Abbreviations for each neural structure are ACC: anterior cingulate cortex, BLA: basolateral complex of the amygdala, CN: central nucleus of the amygdala, DLS: putamen/dorsolateral striatum, dlPFC: dorsolateral prefrontal cortex, DMS: caudate nucleus/dorsomedial striatum, FEF: frontal eye field, GPe: external globus pallidus, GPi/SNr: internal globus pallidus/substantia nigra pars reticulata, HC: hippocampus, HYPO: hypothalamus, lOFC: lateral orbitofrontal cortex, M1: primary motor cortex, mPFC: medial prefrontal cortex, NAcc: nucleus accumbens core, NAcs: nucleus accumbens shell, PMC: premotor cortex, S1: primary somatosensory cortex, SEF: supplementary eye field, SMA: supplementary motor area, STN: subthalamic nucleus, THAL: thalamus, VP/VTA: ventral pallidum/ventral tegmental area.

somatosensory systems. It is highly associative, exhibiting activity tied to specific actions [Burton et al., 2014]. Consequently, actions are represented in self-referential (egocentric) coordinates [Kesner and DiMattia, 1987, Palencia and Ragozzino, 2005]. Conversely, since the nucleus accumbens receives heavily processed inputs from the

hippocampus and prefrontal regions, it represents actions in a more complex, world-centred (allocentric) reference frame [De Leonibus et al., 2005]. The caudate nucleus, receiving both low-level sensory and higher-order associative inputs, is implicated in both egocentric and allocentric representations of motor actions [Ragozzino et al., 2002, Postle and D'Esposito, 2003, Possin et al., 2017]. Therefore, if motor responses relevant to reward-seeking and loss-adverse behaviour are supported by dissociated striatal circuits, then these actions will be represented in different spatial reference frames.

Human behavioural studies support the anatomical dissociation of the win-stay and lose-shift responses. For example, the memory trace supporting lose-shift responding decays following punishment, while win-stay behaviour is stable over time [Gruber and Thapa, 2016, Ivan et al., 2018]. During childhood development, the ability to suppress habitual lose-shift responding improves with age, while-win stay responding does not [Ivan et al., 2018]. Lose-shift suppression is also uniquely disrupted in adults under cognitive load [Ivan et al., 2018]. Finally, the win-stay and lose-shift are differentially affected by the monetary value of wins and losses, the level of feedback provided following outcomes, and the context in which they are presented [Banks et al., 2018].

However, behavioural measures are insufficient to anatomically distinguish the win-stay and lose-shift. In fact, the anatomical basis for these behaviours may vary depending on the context in which actions are performed. For example, the putamen is associated with habitual stimulus-response learning [Burton et al., 2014] while the nucleus accumbens is implicated in goal directed control of optimal action-outcome response strategies [Fuccillo, 2016, Yin et al., 2005]. Consequently, either one of these

structures may be necessary for the win-stay or lose-shift, depending on whether actions are part of a habitual or goal-directed strategy. Therefore, multiple overlapping networks may compete to drive the win-stay and lose-shift behaviours [Packard and White, 1991].

## 1.2 Cannabis Use and Decision-Making

The effects of cannabis use on decision-making provide another means to investigate the neural basis of the win-stay and lose-shift behaviours. For example, acute cannabis administration greatly reduces lose-shift behaviour, but has little effect on the win-stay [Wong et al., 2017a]. Cannabis use also results in impaired spatial processing and visuospatial memory in humans [Pope et al., 1997, Cha et al., 2007]. These behavioural changes following cannabis use result from altered reward processing and striatal function. Acute $\Delta^9$-tetrahydrocannabinol (THC) administration increases dopamine release throughout the striatum, resulting in attenuated loss aversion [Jentsch et al., 1998, Sakurai-Yamashita et al., 1989]. The lateral striatum is the most strongly affected, due to exhibiting the highest density of dopamine transporters [Coulter et al., 1997] and endocannabinoid receptors [Herkenham et al., 1991]. Consequently, cannabis-induced disruption of dopamine signaling in the lateral striatum may affect spatial memory, motor control, and visospatial learning, while disruption of the medial striatum may affect reward processing and goal-directed learning [Darvas and Palmiter, 2009, 2010].

Long-term cannabis sensitization has additional effects on brain function and behaviour. Chronic drug use results in reduced sensory inputs to the putamen and connectivity between the associative/limbic loops and the NAc/caudate nucleus [Blanco-Hinojo et al., 2017, Lichenstein et al., 2017]. Consequently, chronic cannabis use weakens suppression of the habitual lose-shift response [Wong et al., 2017b]. Chronic cannabis use also causes increased volume and density of the nucleus accumbens [Gilman et al., 2014], reduced volume of the putamen [Yip et al., 2014], striatal hyperactivity during reward processing [Jager et al., 2013, Nestor et al., 2010], and reduced dopamine release [van de Giessen et al., 2017]. Since the legalization of cannabis in Canada, recreational usage has increased [Rotermann, 2019]. Therefore, considering the effects of cannabis use on decision-making should be of great importance to policy-makers.

## 1.3 The Experiments

In the following dissertation I seek to evaluate three hypotheses. First, that the value of a decision is determined by its associated spatial-motor action, not a specific choice or visual stimuli. Second, that the win-stay and lose-shift are processed in different spatial reference frames. Third, that habitual cannabis use alters our ability to inhibit lose-shift responses. To evaluate these hypotheses I conducted four experiments in which participants played the game *Matching Pennies* against a computer opponent in exchange for monetary reward. Importantly, this task required inhibition of the win-stay and lose-shift responses in order to avoid exploitation by the computer opponent. By manipulating the context in which actions are performed, we can discover the cognitive processes responsible for the win-stay and lose-shift behaviours.

The experiments in Chapter 2 demonstrated that, in right-handed subjects, suppression of win-stay and lose-shift responses were not specific to the hand used to make decisions. However, response tendencies were specific to choice location and the position of the hand relative to that location. Win-stay responding was much more likely towards choices ipsilateral to (on the same side as) the hand being used, while lose-shift responding was more likely for choices contralateral to (on the opposite side as) the hand being used. This spatial bias in lose-shift responding was much stronger than that present for win-stay behaviour, suggesting spatial representations of choice are particularly important to the lose-shift. However, the extent of spatial bias in win-stay responses were heavily influenced by the hand being used, suggesting it is represented in egocentric, self-referential coordinates. Cannabis use was also associated with sexually dimorphic changes in lose-shift responding. Female cannabis users exhibited increased lose-shift responses while male cannabis users were better able to suppress the lose-shift.

The experiments in Chapter 3 further investigated the spatial properties of the lose-shift and win-stay by comparing choice behaviour in two-choice tasks against those when six, twelve, or eighteen choices are present. This manipulation made it possible to examine whether choice value is tied to a fixed choice or distributed throughout space, and whether manipulating the number of choices present affects the spatial distribution of win-stay and lose-shift behaviour. I found that as the number of choices increase, participants switched from using independent win-stay/lose-shift decision strategies to a single win-shift/lose-shift strategy. I also demonstrate that while the win-stay behaviour consists of a single action, the lose-shift is comprised of two sub-behaviours: "foraging", where nearby choices are explored, and "avoidance",

where individuals shift to a new region of their environment. Finally, the sexually dimorphic effects of cannabis use on lose-shift responding were replicated.

Although the experiments in Chapters 2 and 3 show that spatial location determines choice value, this result was only demonstrated for visually identical choices that remained in fixed positions throughout an experiment. Therefore, in Chapter 4 I investigated competition between visual and spatial choice cues for visually distinct choices that moved spatial locations between trials. I found that even when choices were visually distinct spatial position, rather than visual choice cues, drove win-stay and lose-shift behaviour. Aversion of losing choices was retained when they moved spatial locations, so long as all choices maintained the same spatial arrangement relative to one another. Consequently, the lose-shift is calculated in allocentric spatial coordinates. Conversely, any change in choice position altered win-stay tendencies, indicating it is calculated in egocentric coordinates. We also replicated the sexually dimorphic effects of cannabis use on lose-shift behaviour, finding that lose-shifting increased in female cannabis users and decreased in males.

In Chapter 5 I used electroencephalography (EEG) to record the neural representation of choices during reward processing and execution of win-stay and lose-shift responses. I found that win-stay responding was associated with greater activation of visuospatial (precuneus), somatosensory, and motor systems known to support egocentric processing of space. Moreover, during motor action win-stay specific activity (relative to the win-shift) was present in the motor cortex directly corresponding (contralateral) to the hand being used. Therefore, win-stay related choice value is directly represented in the motor cortex in egocentric coordinates. Activity during motor execution of lose-shift responses was localized within frontal, rather than motor,

structures associated with allocentric spatial processing. Finally, cannabis use was not associated with sexually dimorphic changes in lose-shift responding. Instead, both males and females were better able to suppress the lose-shift, a sign of acute, rather than chronic, cannabis use. This discrepancy in cannabis-associated lose-shifting is likely due to the study in Chapter 5 being conducted directly following legalization of recreational cannabis use in Canada, when first-time cannabis users were prevalent.

In conclusion, my dissertation provides evidence from behaviour, electrophysiology, and computational modeling that choice value is supported by spatial-motor processing. In particular, I demonstrate that spatial and motor processing interact to determine choice value (Chapter 2). Even when a conflict between choice location and visual identity is present, the location of a choice drives win-stay and lose-shift behaviour (Chapter 4). Moreover, the effects of wins or losses are not tied to a specific choice, but distributed throughout space (Chapter 3). However, the win-stay is supported by cortical circuits implicated in egocentric representations of space while the lose-shift is calculated in allocentric coordinates (Chapter 5).

# Chapter 2

# Spatial-Motor Function Supports the Win-Stay and Lose-Shift Responses

## 2.1 Introduction

Among vertebrates, the most prevalent habitual decision strategies are to repeat previously rewarding choices and avoid those that have led to failure. These win-stay and lose-shift strategies are thought to be an innate response of the sensorimotor striatum. However, the primary function of the striatum is to integrate sensory information to drive motor action [Murray et al., 2011, Northcutt, 2008]. Though its structure has remained largely unchanged over the last 530 million years, as the mammalian brain evolved the striatum became increasingly connected with visuospatial, tactile, motor, and motivational systems throughout the cortex and midbrain [Reiner et al., 1998]. Consequently, the"simple" task of motor action selection includes sensory processing,

memory, motivation, sensorimotor habits, and goal-directed assessment of action-outcome associations [Floresco et al., 2008, Gruber and McDonald, 2012, Reig and Silberberg, 2014].

In order to regulate the number of processes needed for motor action selection, the striatum is divided into a number of sub-regions. The putamen or dorsolateral striatum (DLS) receives direct sensory inputs from the somatosensory, motor, orbitofrontal, and insular cortices [Brasted et al., 1999, Pan et al., 2010, Malach and Graybiel, 1986]. Spatial-motor actions are coded with respect to the body, or in egocentric spatial coordinates [Kesner and DiMattia, 1987, Palencia and Ragozzino, 2005]. Consequently, decision-related activity in the DLS is highly associative [Burton et al., 2014], in the sense of being tied to specific spatial-motor actions, environmental contexts, and expected outcomes [Burton et al., 2015, Skelin et al., 2014]. The nuclues accumbens (NAc) or ventral striatum receives inputs from the hippocampus, profrontal cortex, amygdala, and hypothalamus [Gerfen, 1984, Voorn et al., 2004, Kelley et al., 1982]. These heavily processed inputs represent actions in allocentric (i.e., world-centered) and egocentric coordinates during spatial-motor learning [De Leonibus et al., 2005]. However, the NAc is primarily associated with goal-directed control of decisions, with activity representing motivation towards more highly-valued goals [Burton et al., 2015]. The caudate nucleus or dorsomedial striatum (DMS), which receives inputs from the hippocampus, amygdala, anterior cingulate (ACC), and visual cortex [Khibnik et al., 2014, Fuccillo, 2016], also drives goal directed action. Like the DLS, the DMS also exhibits associative properties, coding both reward value and motor action-outcome associations [Yin et al., 2005, Burton

et al., 2015]. Consequently, it has been implicated in decision-making in both egocentric and allocentric frames of reference [Ragozzino et al., 2002, Postle and D'Esposito, 2003, Possin et al., 2017].

Due to its role in associative learning between stimuli, actions, and rewards, the sensorimotor striatum supports habitual decision strategies. In particular, it is related to repetition of previously rewarded actions (win-stay) and avoidance of those that have led to punishment (lose-shift). During competitive decision-making tasks lose-shift responding has been found to rely on the putamen (DLS) while the win-stay does not [Gruber et al., 2017, Skelin et al., 2014, Danckert et al., 2011]. However, lose-shifting may also rely on the medial striatum [Skelin et al., 2014, Clarke et al., 2008] including the caudate and NAc. The medial striatum is known to support alternation between response strategies, consistent with its association with the lose-shift [Ragozzino et al., 2002, Ragozzino, 2007, McDonald et al., 2008]. Conversely, Thapa and Gruber [2018] demonstrated that lesions to the ventrolateral striatum, but not the DMS, influence lose-shift behaviour. Win-stay responding has been found to rely on the dorsal striatum [McDonald and White, 2013], DMS [Packard et al., 1989] and NAc core [Gruber et al., 2017] in different experimental contexts. However, Thapa and Gruber [2018] assert the DMS and VLS are not necessary for win-stay behaviour. This debate suggests that win-stay and lose-shift responses do not depend on distinct, discrete loci, but instead are supported by multiple overlapping networks [Packard and White, 1991] that are variably affected by task context and demands. Regardless, these results indicate the sensorimotor striatum is necessary for both behaviours.

Given that motor action selection, spatial processing, and reward-based response strategies rely on the sensorimotor striatum, the value of choices may be influenced

in part by their associated spatial-motor actions, rather than their abstract identity. Therefore, we hypothesized win-stay and lose-shift responding is influenced by the hand used to make choices (i.e., motor action) and the location of choices relative to the observer. We tested this hypothesis while participants were engaged in a competitive decision-making task between two choices located to their left and right. Crucially, participants completed this task with their right and left hands over two separate sessions. We found overall win-stay and lose-shift responding was unaffected by the hand used, but that behaviour was strongly influenced by the location of choices relative to the individual. Specifically, lose-shifting was the dominant response strategy when avoiding choices located opposite to the hand being used (e.g., left hand & right choice) whereas win-stay responding dominated when choices were on the same side as the hand. This spatial bias in lose-shift responding was present for both the dominant and non-dominant hands, but a spatial bias in win-stay behaviour only manifested during use of the dominant hand. This last result suggests the win-stay may be calculated in egocentric spatial coordinates (i.e., with reference to the self) supported by motor representations of choice, whereas allocentric spatial processing (i.e., world centered) is more relevant to the lose-shift. In both cases, the value of choices and their associated rewards rely on their location in space and the movement required to reach that space.

The current experiment also examined the effect of cannabis use on choice behaviour. Recreational cannabis use influences spatial memory, motor control, reward

processing, and goal-directed learning [Cha et al., 2007, van Hell et al., 2010], particularly in females [Pope et al., 1997]. Acute $\Delta^9$- tetrahydrocannabinol (THC) administration increases dopamine release in the DLS, while the ventromedial striatum remains unaffected [Sakurai-Yamashita et al., 1989]. Behaviourally, dopamine signalling in the DLS is necessary for normal spatial memory, motor control, and visospatial learning, while reward processing and goal-directed learning rely on dopamine function in the medial striatum [Darvas and Palmiter, 2009, 2010]. Consequently, THC and amphetamine use cause large changes in lose-shift behaviour in rats and humans, though win-stay responding is only weakly affected [Wong et al., 2017b,a, Paulus et al., 2002b]. Therefore, we also hypothesized that cannabis use may influence the reliance of lose-shift behaviour on spatial-motor processing. We found that cannabis use does not affect the relationship between lose-shifting and choice location. Lose-shift behaviour is elevated in female cannabis users, who are less able to inhibit sensorimotor responding via executive systems. However, male cannabis users exhibit no differences in choice behaviour, suggesting the effects of recreational cannabis use on decision-making are sexually dimorphic.

## 2.2 Methods

### 2.2.1 Behavioural Task

In order to assess the tendency towards habitual win-stay and lose-shift responding, participants played the competitive game "Matching Pennies" against a computer opponent. This task requires participants to suppress habitual response patterns that could be exploited by the computer opponent. The task display consisted of two

distinct targets (blue circles) presented on the left and right sides of a 15″ touchscreen monitor (Fig. 2.1). On each trial, the computer would predict the participant's choice, based on patterns in their past decisions. Participants would select either target using their left or right hand. If the computer failed to predict the selected target, reward feedback would follow, indicating "You Win" for 1.5 s paired with a high frequency auditory tone. If the computer prediction was correct, "You Lose" was presented on the screen paired with a low frequency tone.

To minimize the number of wins gained by each subject, the computer used four types of algorithms to detect patterns in (*i*) participants' choices; (*ii*) switching from one choice to another; (*iii*) choices paired with rewards (e.g., left choice after a loss); and (*iv*) switching paired with rewards (e.g., swapping choices after a loss). These four algorithms were used to examine the most recent choices made (e.g., shifting from left to right after a loss) and find other instances of that choice pattern in the participant's past choice history. The choice that most frequently followed these past instances was selected as the prediction for the future choice. Each algorithm considered choice patterns 1-6 trials in length, resulting in 24 total strategies used by the computer. On each trial, the best performing strategy (computed over all previous trials in the session) was used to predict participants' choices. In addition, if all strategies failed to beat the participant on $\geq 50\%$ of trials, the computer would select choices randomly. The optimal choice strategy among participants is to suppress all habitual choice strategies, including the win-stay and lose-shift. Consequently, this task indicates to what extent participants lose-shift, win-stay, and exhibit cognitive flexibility in competitive situations.

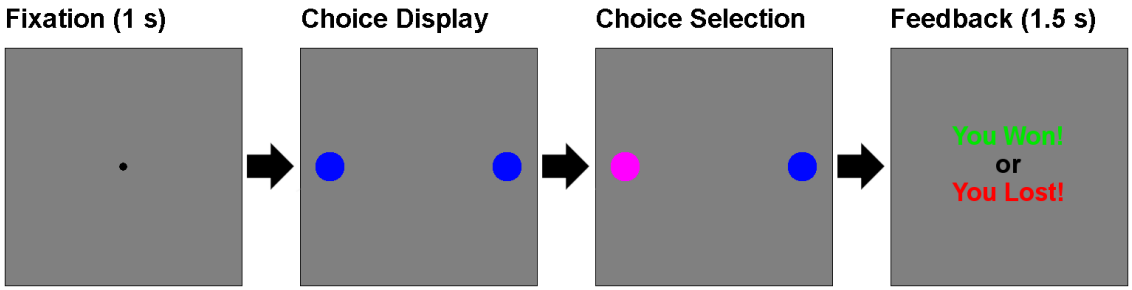


Figure 2.1: Time course of the behavioural task.

### 2.2.2 Procedure

All procedures and experimental tasks were approved by the McMaster University Research Ethics Board. Eighty-nine right handed subjects (42 males, mean age = 19.9, $SD = 2.4$ years) from McMaster University participated in the study in exchange for payment. After providing informed consent, subjects participated in two, six hundred trial sessions matching pennies over two consecutive days. On each day they were instructed to use either their right or left hand to complete the experiment. To ensure that only one hand was used to complete each session, subjects were required to keep their other hand on a sensor that would pause the experiment whenever their hand was removed. There were four groups of subjects who used their right hand on both days (N = 23), their left hand on both days (N = 22), switched from their left to right hand between days (N = 22), or switched from their right to left (N = 22). Only right-handed subjects were included in the study, as assessed by the Edinburgh handedness inventory [Oldfield et al., 1971].

Prior to the task, participants were informed they would receive 3¢ for each winning trial, nothing for losing trials, and that their total winnings would be rounded up to the nearest $5 upon experiment completion. While they were informed they

were competing against a computer opponent, no guidance as to optimal decision-making strategies was given. After task completion, participants were screened with the South Oaks Gambling Screen, the alcohol, smoking and substance involvement screening test (ASSIST) v3.0, the Adult ADHD Self-Report Scale v1.1, and a demographic questionnaire. Habitual cannabis users were defined as those meeting the criteria for brief or intensive treatment (score $> 3$) on the ASSIST cannabis subtest. Twelve males and seven females met the criteria for cannabis use requiring intervention.

### 2.2.3 Analysis

For each participant, decision times and the number of wins, lose-shift responses, win-stay responses were averaged over five blocks of 120 trials during each experimental session. Decision times were measured as the duration between the end of the 1.5 seconds of reward feedback and the following response. They were log-transformed and averaged after removing RTs $<$50 ms or $>$30 s. The prevalence of mixed-response strategizing, or behavioural flexibility, was assessed for each session day as binary response entropy ($H$) calculated from independent sequences of 4-trials as:

$$H = \sum_{i=1}^{k} P_i + log_2 P_i \tag{2.1}$$

where $P_i$ is the probability of each choice sequence, and $k$ is the total number of sequences possible (i.e., 16). For example, a participant that only alternated choices would have a entropy of zero bits, while one exhibiting random responses would have a maximum entropy of four bits.

To extract underlying decision-processes from choice behaviour we used the *Q-learning with forgetting* [Barraclough et al., 2004] reinforcement learning model to examine the effects of hand used on learning, reward sensitivity, and choice stochasticity. In this model, the probability of selecting one of the two choices ($C_i$ and $C_j$) on a given trial ($t$) are calculated according to the softmax equation [Sutton and Barto, 2018]:

$$P(C_t = i|Q_i, Q_j) = \frac{exp(\beta \times Q_{i(t)})}{exp(\beta \times Q_{i(t)}) + exp(\beta \times Q_{j(t)})} \tag{2.2}$$

where $Q_i$ and $Q_j$ are the interval values each subject assigns to choices $i$ and $j$. $\beta$ refers to the inverse temperature that balances the opposing tendencies to exploit known action-reward associations and to explore more of the state/action space. As such, larger values of $\beta$ indicate a greater tendency to choose the most highly valued action rather than explore new actions. The values of each choice are updated from rewards ($R$) according to the following rules:

$$Q_i(t) = \begin{cases} Q_{i(t-1)} \times (1-\alpha) + \alpha\kappa_1, & \text{if } C_{t-1} = i, R_{t-1} = 1 \\ Q_{i(t-1)} \times (1-\alpha) - \alpha\kappa_2, & \text{if } C_{t-1} = i, R_{t-1} = 0 \\ Q_{i(t-1)} \times (1-\alpha), & \text{if } C_{t-1} \neq i \end{cases} \tag{2.3}$$

where $\alpha$ is the learning and forgetting rates for the selected and non-selected actions. Larger values of $\alpha$ result in individual wins and losses having a greater impact on changes in $Q_i$ and $Q_j$. $\kappa_1$ is the strength of reinforcement from reward and $\kappa_2$ is the strength of aversion from failing to receive a reward. As $\kappa_2$ is subtracted from the $Q_i$ and $Q_j$, positive value of $\kappa_2$ indicate greater loss aversion. These three parameters

were treated as stochastic variables that follow a random walk process and therefore they were free to vary throughout the experiment. Conversely, $\beta$ was treated as a deterministic variable that remained fixed throughout the experiment. Bayesian priors for $Q_i$, $Q_j$, $\alpha$, and $\beta$ were defined as normal distributions with a $\mu$ of 0 and $\sigma$ of 1. The logistic ($\frac{1}{1+e^{-x}}$) and exponential functions were applied to priors for $\alpha$ and $\beta$, resulting in a $\mu$ of 0.5 and 1 respectively. These parameters were fit for each subject using the VBA toolbox [Daunizeau et al., 2014]. Prior to statistical analysis, these measures were also averaged over trial blocks.

To determine how win-stay and lose-shift responding influenced RL parameters, and how these responses differed as a function of hand used, we performed a Volterra decomposition of $\kappa_1$, and $\kappa_2$ values for each trial onto four basis functions ($u$): previous choice, outcome, win-stay, and lose-shift responding, according to Eq. 2.4:

$$x_t = \omega^0 + \sum_{\tau} \omega^1_{\tau} u_{t-\tau} + \sum_{\tau_1} \sum_{\tau_2} \omega^2_{\tau_1,\tau_2} u_{t-\tau_1} u_{t-\tau_2} + ... \tag{2.4}$$

Volterra modelling expresses the input response characteristics of non-linear systems as Volterra weights [Boyd et al., 1984]. At each trial $t$ the Volterra weight $x$ of a given parameter is estimated from inputs $u$ over trials $t$ to a lag of $\tau$ (set to 32 trials) using a series of Volterra kernels $\omega$. The first kernel $\omega^1$ represents the linear transformation of lagged input basis functions into the output, $\omega^2$ represents the effect of past inputs being dependent on other earlier inputs, and so on. These weights provide a measure of how a participant's valuation of each choice changes from baseline in response to past choices and outcomes. The benefit of Volterra modelling over an analysis of raw prediction error is that the effect of current and past inputs on hidden state responses can be estimated. Inputs were also orthogonalized so that the

effect of one input (e.g., win-stay) is computed independently of all other inputs (e.g., outcome). For example, wins and losses were represented as a +1/-1 vector (i.e., $A_0$), six hundred trials long. The effect of win-stay responses (i.e., $A_1$) orthogonal to that of other outcomes were calculated by projecting $A_1$ into the null space of $A_0$ using the Moore–Penrose pseudoinverse (i.e., $A_0 \times A_0^{+}(A_0' \times A_0) \times A_0'$). To control for trial order effects, we also detrended inputs prior to decomposition using linear, quadratic, and cubic polynomials.

These measures (i.e., behavioural, RL parameters, & Volterra weights) were analyzed with repeated-measure, mixed-effects models that included trial blocks (i.e., 1-5) or trial lag (for Volterra decomposition) as a covariate, while maintaining the assumption of equal slopes between groups. Because some subjects used only one hand over both sessions, a maximal random-effects structure [Barr et al., 2013] could not be fit to the data. However, random intercepts, effects of session day, treatment group, and random slopes of trial block were fit in R using the lme4 package [Bates et al., 2014] using the Nelder-Mead optimizer. Degrees of freedom and $p$-values were calculated using the Welch-Satterthwaite equation and type-III sums of squares. The effects of cannabis use and sex were assessed via planned t-tests conducted on behavioural measures averaged across each session, rather than across trial blocks.

## 2.3 Results

### 2.3.1 Choice Behaviour is not Specific to the Hand Used

Each participant performed 1200 trials of matching pennies over two days, resulting in a dataset of 106,800 trials. Given the sensorimotor striatum drives both motor action

and win-stay/lose-shift responding, we hypothesized that choice behaviour is specific to the hand used. Therefore, two-way mixed-effects models tested the effects of hand used, session day, and their interaction on block-averaged rates of win-stay responses, lose-shift responses, task performance, and decision times. The fixed effect of trial block on each day was included as a covariate, with the assumption of equal slopes. Each model also included the random effects of subject, hand nested within subject, and trial block. In other words, the fixed effects of hand used, session day, and block were tested after controlling for random variability between subjects, the hand used by each subject, and over trials. As response entropy was averaged over each day, a $2 \times 2$ ANOVA tested the effects of hand used and session day on behaviour.

We also tested to what extent optimal win-stay behaviour, lose-shift behaviour, and response entropy each contributed to task performance using multiple regression. Both the linear and quadratic effects of win-stay and lose-shift behaviour were included, as response rates above or below 50% are sub-optimal. However, only the linear effects of response entropy were considered. The relative importance of each behaviour to win-rate was calculated as partial eta-squared ($\eta_P^2$; $SS_{effect}/(SS_{effect} + SS_{error})$).

As seen in Figure 2.2.A, variability in lose-shift behaviour was not significantly associated with the hand used, trial blocks, sessions, or the hand $\times$ session interaction ($p > .215$ in all cases). Consequently, suppression of the lose-shift response remains relatively stable during long periods of decision-making. Conversely, suppression of the win-stay response improved over trial blocks [$F_{1,292.45} = 13.272$, $p < .001$, $SD\upsilon_0 = .010$]. However, as seen in Figure 2.2.B, this improvement does not transfer between sessions, regardless of the hand used ($p > .404$ in all cases). Despite the improvement

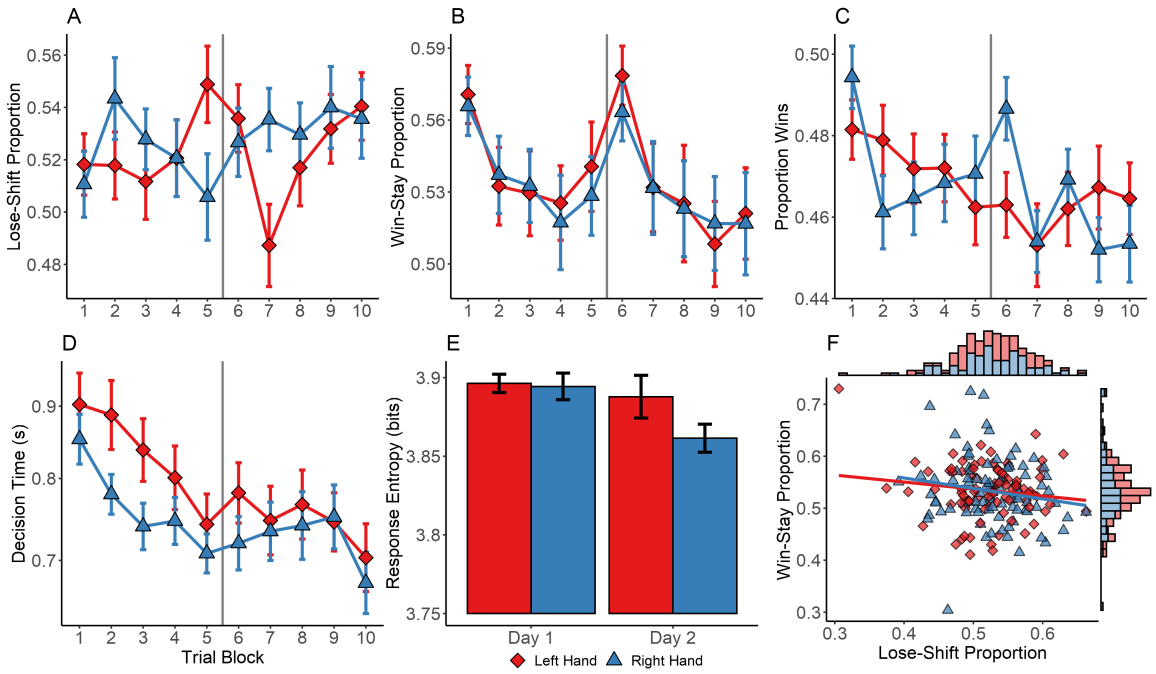


Figure 2.2: Effects of hand used and trial block on proportion of lose-shift responses (A), win-stay responses (B), wins received (C), and log decision times (D). Grey vertical line indicates break between trial days. E: effect of hand used and trial day on response entropy. F: correlation between win-stay and lose-shift responding.

in win-stay suppression and the stability of lose-shift suppression, task performance (win-rate) declined between sessions [$F_{1,87.41} = 7.516$, $p = .007$, $SDv_0 = .008$] and trial blocks within each session [$F_{1,260.24} = 7.040$, $p = .008$, $SDv_0 = .004$], as seen in Figure 2.2.C. However, the effect of hand used on task performance was not significant ($p = .527$). Therefore, lose-shift and win-stay behaviour is not specific to the hand used, but transfers across motor modalities. Furthermore, lose-shift suppression remains static over trial blocks, and is not subject to the learning effects within the time course of the experiment, while win-stay suppression does improve with time. As seen in Figure 2.2.D, decision times also decreased over trial blocks [$F_{1,97.09} = 23.023$, $p < .001$, $SDv_0 = .048$] and sessions [$F_{1,88.85} = 10.966$, $p < .001$, $SDv_0 = .143$] but

did not significantly differ between the hands [$F_{1,118.69} = 3.385$, $p = .068$].

Response entropy was assessed via a 2 (left/right hand) $\times$ 2 (session 1/2) ANOVA. As seen in Figure 2.2.E, entropy was reduced on day 2, as indicated by a significant main effect of session [$F_{1,174} = 4.728$, $p = .031$]. However, the effect of hand used and the hand $\times$ day interaction were not significant ($p > .138$ in both cases). Therefore, the long and repetitive nature of the task is associated with a decline in mixed-response strategizing.

Across participants, rates of win-stay behaviour [$F_{2,172} = 11.670$, $p <= .001$, $\eta_P^2 = .119$], and lose-shift behaviour [$F_{2,172} = 10.659$, $p <= .001$, $\eta_P^2 = .110$] equally contributed to task performance. However, after accounting for the effects of the win-stay and lose-shift, response entropy (mixed-strategy responding) was the greatest determinant of task performance [$F_{1,172} = 30.701$, $p <= .001$, $\eta_P^2 = .151$]. Despite the importance of win-stay and lose-shift suppression to task performance, both behaviours were uncorrelated regardless of the hand used ($p > .119$ in both cases). Consequently, the win-stay and lose-shift are dissociated strategies supported by separate cognitive processes (Fig. 2.2.F). However, this dissociation breaks down with practice, boredom, or fatigue. For either hand the win-stay and lose-shift were uncorrelated during day 1 [r(87)=-.040, p=.711], but exhibited a weak, negative correlation during day 2 of the experiment [r(87)=-.154, p=.040]. Consequently, as sessions progressed a default switching strategy becomes more prominent. In no case did sex or cannabis use influence the relationship between the lose-shift and win-stay behaviours ($p > .241$ in all cases).

### 2.3.2 Spatial Location Drives Choice Value

We have demonstrated that unlike pure motor skills, habitual decision strategies are not affected by use of the dominant or non-dominant hand. Instead, the win-stay and lose-shift may be influenced by the spatial location of choices. Therefore, we investigated how the choice position relative to the body influenced win-stay and lose-shift behaviour. The effects of choice location, hand, and session were tested via 2 (left/right choice) $\times$ 2 (left/right hand) $\times$ 2 (day 1/2) mixed-effects models. Changes in behaviour over trials blocks were also controlled for while maintaining the assumption of equal slopes. The statistical model included random effects of subject, session nested within subject, previous choice within subject, and trial block. To aid in interpretation, lose-shift behaviour was recoded as lose-stay (1 - lose-shift) before analysis.

| | Lose-Shift | | | | Win-Stay | | | |
|---|---|---|---|---|---|---|---|---|
| Variable | $\beta$ | $SE(\beta)$ | $p\beta$ | $SD\upsilon_0$ | $\beta$ | $SE(\beta)$ | $p\beta$ | $SD\upsilon_0$ |
| Intercept | .536 | .007 | <.001 | .015 | .562 | .008 | <.001 | .000 |
| Block | .002 | .002 | .442 | .009 | -.010 | .003 | <.001 | .011 |
| Day (1) | -.005 | .003 | .104 | .021 | .001 | .004 | .777 | .030 |
| Hand (L.H.) | -.004 | .004 | .307 | — | .001 | .004 | .791 | — |
| Choice (L.C.) | .010 | .004 | .008 | .029 | -.013 | .003 | <.001 | .014 |
| Hand $\times$ Choice | -.035 | .003 | <.001 | — | .021 | .003 | <.001 | — |
| Hand $\times$ Day | -.001 | .004 | .739 | — | -.000 | .004 | .920 | — |
| Choice $\times$ Day | -.005 | .003 | .054 | — | .003 | .003 | .303 | — |

Table 2.1: Results of mixed effects models analyzing the relationship between choice location and hand used on win-stay and lose-shift responses. Coefficients represent difference of reference parameter (Day 1, left hand, and left choice) from grand mean.

Accounting for choice location and hand used considerably improved model fit, relative to a model that only considered hand used [$\chi^2(4) = 167.78$, $p < .001$, conditional $r^2 = .217$]. As seen in Figure 2.3.A and Table 2.1 lose-stay (and consequently

lose-shift) responding was significantly affected by previous choice location [$F_{1,88.62} = 7.455$, $p = .008$, $SD\upsilon_0 = .029$] and a previous choice × hand interaction [$F_{1,517.97} = 127.253$, $p < .001$]. No other main effects or interactions were significant ($p > .054$).

A model accounting for choice location also fit significantly better than one that only considered hand used [$\chi^2(4) = 59.206$, $p < .001$, conditional $r^2 = .149$]. Win-stay responding was also affected by previous choice location [$F_{1,97.92} = 15.410$, $p < .001$, $SD\upsilon_0 = .014$] and a hand × location interaction [$F_{1,391.99} = 39.395$, $p < .001$] after controlling for the effect of trial block [$F_{1,392.70} = 17.483$, $p <= .001$, $SD\upsilon_0 = .011$]. No other effects were significant ($p > .303$). As seen in Figure 2.3.B, win-stay behaviour is more prominent when the previously winning choice is ipsilateral to (i.e., on the same side as) the hand being used. Similarly, lose-shifting dominates when the previously losing choice is contralateral to the hand. This spatial bias is much greater for the lose-stay ($M = 7.44\%$, $SE = .79\%$) than the win-stay ($M = 5.10\%$, $SE = .80\%$), indicating spatial processing is more important to lose-shift responding.

Though a bias towards the ipsilateral choice was present when using either hand, it appeared to be stronger when using the dominant hand. Therefore, from the 44 subjects who alternated hands between sessions, we tested whether they exhibited a stronger bias when using their right hand. Preference for the dominant hand was measured by comparing the difference in win-stay or lose-shift rates for the ipsilateral v. contralateral choices when using the left and right hands. The differences between left and right-hand ipsilateral biases were assessed using one-way, mixed-effects models that estimated the effect of hand used while also controlling for trial block. The model incorporated random intercepts and slopes over trial blocks.

As seen in Figure 2.3.C and Table 2.2 an ipsilateral bias in lose-stay responding

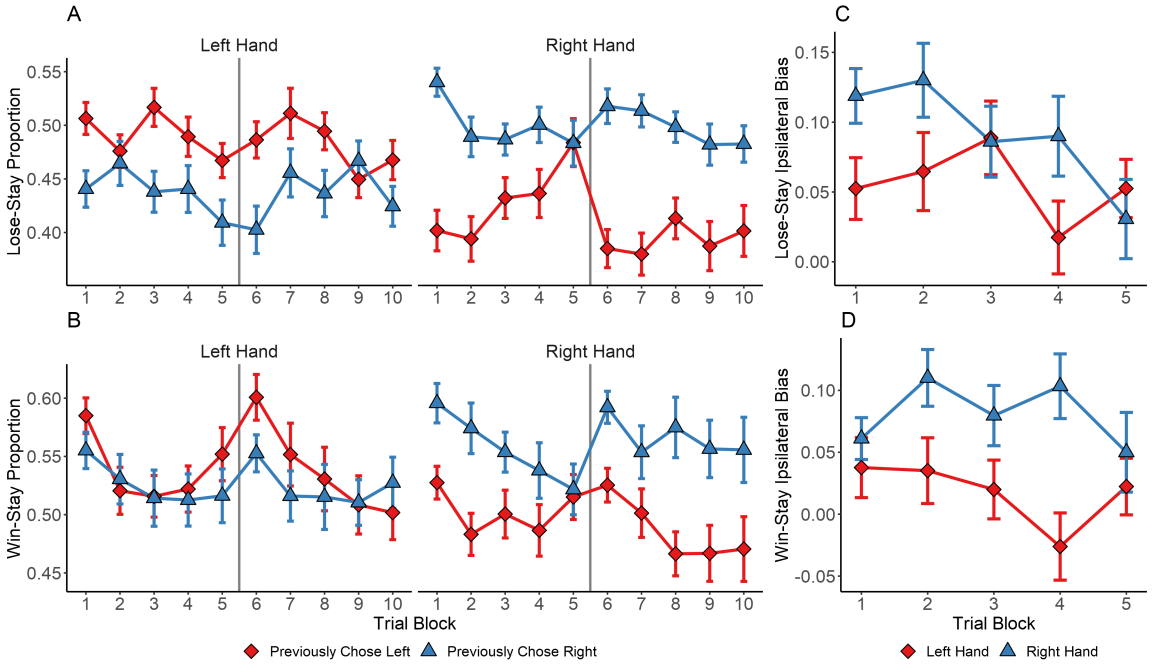


Figure 2.3: A-B: effects of hand used and choice location on win-stay and lose-stay (1 - lose-shift) behaviour. Grey vertical line indicates break between trial days. C-D: bias towards ipsilateral choices (relative to contralateral) during win-stay and lose-shift responding with the left and right hands.

| | Lose-Stay Bias | | | | Win-Stay Bias | | | |
|---|---|---|---|---|---|---|---|---|
| Variable | $\beta$ | $SE(\beta)$ | $p\beta$ | $SD\upsilon_0$ | $\beta$ | $SE(\beta)$ | $p\beta$ | $SD\upsilon_0$ |
| Intercept | .113 | .019 | <.001 | .044 | .067 | .018 | <.001 | .034 |
| Block | -.013 | .005 | .014 | .009 | -.006 | .005 | .274 | .011 |
| Hand (L.H.) | -.017 | .007 | .027 | — | -.029 | .007 | <.001 | — |

Table 2.2: Results of mixed effects models analyzing the effect of hand used on ipsilateral biases in win-stay and lose-shift responses.

was present in both hands, but significantly greater for the dominant hand [$F_{1,403.37} = 4.909$, $p = .027$]. A significant effect of trial block was also present [$F_{1,175.45} = 6.122$, $p = .014$, $SD\upsilon_0 = .009$] as spatial biases were attenuated with practice. However, the ipsilateral bias in win-stay responding was only present for the dominant hand [$F_{1,404.90} = 15.900$, $p < .001$], as seen in figure 2.3.D. The effect of trial block was

not significant ($p = .274$), indicating spatial bias in win-stay responding remained constant over time. That spatial processing of the win-stay was influenced by motor action ($M = 5.75\%$, $SE = 1.58$) much more than the lose-stay ($M = 3.13\%$, $SE = 1.46$) indicates the win-stay may be calculated in egocentric spatial coordinates.

### 2.3.3 Computational Results

We have demonstrated that the dominant and non-dominant hands do not differ in overall usage of sensorimotor and mixed-response strategies. However, response modality does affect dependence on spatial choice cues during win-stay responding (Fig. 2.3). Response modality may also influence underlying reward processing and learning that drive behaviour. Therefore, we used computational modeling and Volterra decomposition to analyze how choice valuation and learning processes changed in response to feedback and different choices. We used the *Q-learning with forgetting* (FQ) model to derive learning rates ($\alpha$), inverse temperature ($\beta$), reward strength ($\kappa_1$), and punishment strength ($\kappa_2$) from participants. This model was compared against the *Q-learning* (Q) model of Sutton and Barto [2018] and *Q-learning with differential forgetting* (DFQ) of Ito and Doya [2009]. The Q model contains only the $\alpha$ and $\beta$ parameters, while DFQ adds a $\alpha_2$ term to account for forgetting processes.

Accuracy of these three models were compared using the negative log-likelihood:

$$\text{Negative log-likelihood} = -\frac{1}{n} \times \sum_{i=1}^{n} log(P(i)) \tag{2.5}$$

The FQ model had a significantly lower negative log-likelihood ($M = .038$, $SD = .020$) than both the Q [$M = .661$, $SD = .098$, $t(177) = 81.597$, $p < .001$] and DFQ

models [$M = .048$, $SD = .025$, $t(177) = 4.149$, $p < .001$], indicating that it provided the best fit to the data. Model fit did not depend on experimental session, use of the dominant/non-dominant hand, participant sex, or cannabis use ($p > .118$ in all cases), indicating the model was equally applicable to all groups investigated.

Mixed-effects models tested whether $\alpha$, $\kappa_1$, and $\kappa_2$ varied between sessions, hands, the session $\times$ hand interaction, or trials. As in previous analyses, the assumption of equal slopes in RL parameters over trials was maintained. To aid in model convergence trial numbers were rescaled to have a variance of 1. The random effects of subject, session nested within subject, and random slopes over trial blocks were also controlled for.

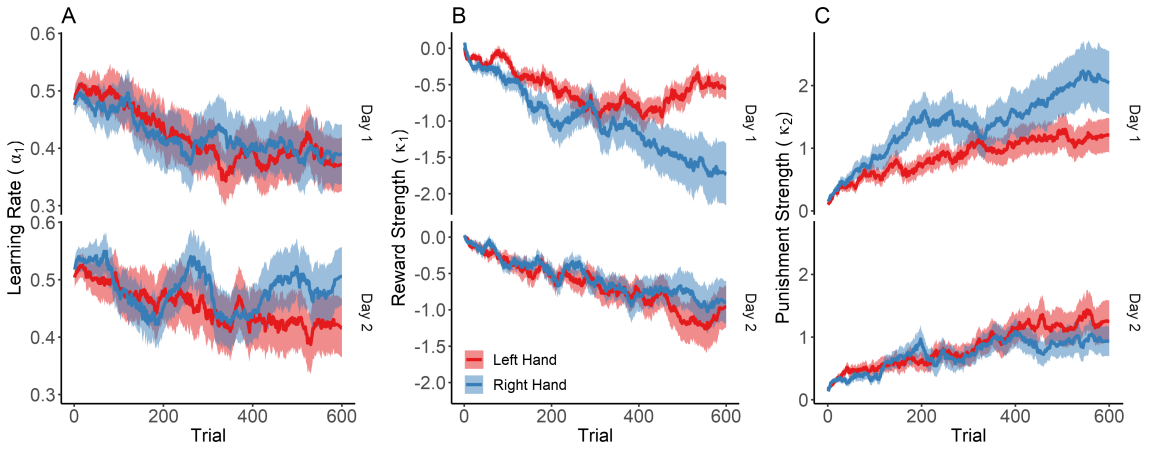


Figure 2.4: Change in learning rate (A, $\alpha$), reward value (B, $\kappa_1$) and punishment strength (C, $\kappa_2$) over session days and trials during use of the left and right hands. Shaded area indicated the standard error of the mean.

As seen in Figure 2.4, use of the dominant hand caused reward strength ($\kappa_1$) to decline more rapidly over trials and punishment strength ($\kappa_2$) to increase. The effects of hand, session, and the hand $\times$ session interaction on learning rates ($\alpha$), $\kappa_1$, and $\kappa_2$ were not statistically significant ($p > .052$ in all cases). Learning rates [$F_{1,89.00}$

= 13.957, $p < .001$, $SD\upsilon_0 = .065$] and reward strength [$F_{1,89.00} = 34.877$, $p < .001$, $SD\upsilon_0 = .452$] significantly declined over trials. This pattern of $\alpha$ decay over time is similar to that previously found in human behaviour [Nassar et al., 2010, McGuire et al., 2014]. Punishment strength also increased [$F_{1,89.00} = 32.832$, $p < .001$, $SD\upsilon_0 = .505$]. While the average value assigned to wins and losses did not differ with the hand used, participants may exhibit trial-to-trial variations in $\alpha$, $\kappa_1$, and $kappa_2$ that depend on spatial-motor action.

### 2.3.4 Volterra Decomposition

We used Volterra decomposition to investigate the immediate and future effects choice location, reward, win-stay, and lose-shift responding have on hidden states. This method accounts for the effects choices and feedback have on future changes in $\alpha$, $\kappa_1$, and $\kappa_2$ over the following $n \in (1, 32)$ trials. The change each of these parameters exhibit in response to an event (e.g., choosing the right v. left option, winning v. losing) are referred to as Volterra weights.

$2 \times 2$ mixed-effects models tested whether Volterra weights for $\alpha$, $\kappa_1$, and $\kappa_2$ varied with hand used (left/right) or session (day 1/2). As Volterra weights were calculated over the thirty-two trials following an event, log-transformed trial lag was included as a factor while maintaining the assumption of equal slopes. The random effects of subject, session nested within subject, and trial lag were also controlled for.

The results of each mixed-effect model are provided in Tables 2.3-2.5. As seen in Table 2.3, learning rates changed in response to rewards. This change significantly differed between the hands [$F_{1,156.69} = 4.001$, $p = .047$] as winning with the left hand increased $\alpha$ (relative to losing) while $\alpha$ decreased with the right. A significant effect

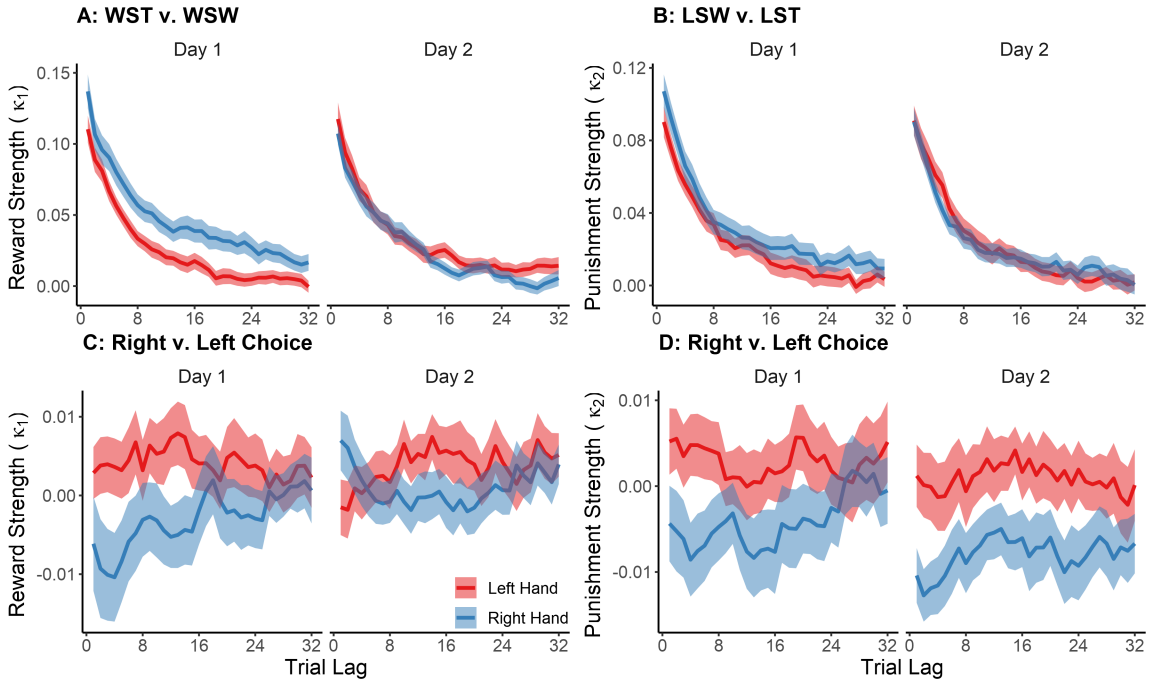


Figure 2.5: A: change in reward value ($\kappa_1$) associated with win-stay responses (relative to win-shift) made with the left and right hands. B: change in punishment strength ($\kappa_2$) associated with lose-shift responses. C-D: change in reward and punishment strength associated with selecting the right (relative to the left) choice location.

of trial lag was also present [$F_{1,88.79} = 9.164$, $p = .003$, $SD\upsilon_0 = .008$], as learning rates declined following feedback. Consequently, choice value was more heavily influenced by recent rewards and punishments during use of the non-dominant hand. Learning rates also decreased following win-stay responses before settling back to baseline [$F_{1,88.85} = 26.418$, $p < .001$, $SD\upsilon_0 = .019$]. Lose-shifting increased $\alpha$ which then declined over subsequent trials [$F_{1,88.88} = 37.631$, $p < .001$, $SD\upsilon_0 = .017$]. However, the change in $\alpha$ following win-stay or lose-shift responses did not vary between hands or sessions ($p > .185$).

Win-stay responses also influenced reward strength. As seen in Figure 2.5.A and Table 2.4, $\kappa_1$ spiked following a win-stay and declined back to baseline over subsequent

| | $\alpha$: Right v. Left | | | | $\alpha$: Win v. Loss | | | |
|---|---|---|---|---|---|---|---|---|
| Variable | $\beta$ | $SE(\beta)$ | $p\beta$ | $SD\upsilon_0$ | $\beta$ | $SE(\beta)$ | $p\beta$ | $SD\upsilon_0$ |
| Intercept | .000 | .004 | .910 | .027 | .003 | .003 | .325 | .021 |
| $ln$(Lag) | .000 | .001 | .938 | .011 | -.003 | .001 | .003 | .008 |
| Day | -.003 | .002 | .276 | .031 | -.002 | .002 | .251 | .028 |
| Hand | -.001 | .003 | .637 | — | .005 | .002 | .047 | — |
| Day×Hand | .002 | .003 | .481 | — | .004 | .002 | .127 | — |
| | $\alpha$: WST v. WSW | | | | $\alpha$: LSW v. LST | | | |
| Intercept | -.030 | .006 | <.001 | .049 | .026 | .006 | <.001 | .043 |
| $ln$(Lag) | .011 | .002 | <.001 | .019 | -.012 | .002 | <.001 | .017 |
| Day | -.002 | .003 | .410 | .035 | -.004 | .003 | .186 | .043 |
| Hand | .002 | .003 | .651 | — | .001 | .004 | .735 | — |
| Day×Hand | -.000 | .003 | .905 | — | -.002 | .004 | .673 | — |

Table 2.3: Results of mixed effects models analyzing changes in learning rates ($\alpha$) in response to choice location, outcome, win-stay, and lose-shift responses.

trials [$F_{1,88.71} = 438.48$, $p < .001$, $SD\upsilon_0 = .014$]. This spike was greatest during use of the right hand, particularly on day 1. Therefore, a significant hand × session interaction was present [$F_{1,123.80} = 8.561$, $p = .004$]. $\kappa_1$ was not influenced by choice location or lose-shift responses ($p > .173$). However, following a winning outcome the value of future rewards ($\kappa_1$) decreased before returning to baseline over subsequent trials [$F_{1,89.49} = 13.965$, $p < .001$, $SD\upsilon_0 = .006$].

As seen in Figure 2.5.B and Table 2.5, lose-shift responding is also tied to a large spike in punishment strength that decays over subsequent trials [$F_{1,88.80} = 314.632$, $p < .001$, $SD\upsilon_0 = .014$]. However, this spike is not influenced by the hand used ($p > .223$). Instead, changes in $\kappa_2$ are linked to choice location and this relationship varies with the hand used [$F_{1,142.37} = 9.218$, $p = .003$]. As seen in figure 2.5.D, loss aversion is less severe for punished choices ipsilateral to the hand being used. Win-stay behaviour is also linked to a spike in punishment strength that subsequently

| | $\kappa_1$: Right v. Left | | | | $\kappa_1$: Win v. Loss | | | |
|---|---|---|---|---|---|---|---|---|
| Variable | $\beta$ | $SE(\beta)$ | $p\beta$ | $SD\upsilon_0$ | $\beta$ | $SE(\beta)$ | $p\beta$ | $SD\upsilon_0$ |
| Intercept | -.001 | .003 | .712 | .022 | -.008 | .002 | <.001 | .009 |
| $ln$(Lag) | .001 | .001 | .298 | .008 | .003 | .001 | <.001 | .006 |
| Day | -.001 | .001 | .495 | .017 | -.000 | .002 | .940 | .024 |
| Hand | .002 | .002 | .214 | — | .002 | .002 | .368 | — |
| Day×Hand | .001 | .002 | .378 | — | -.001 | .002 | .455 | — |
| | $\kappa_1$: WST v. WSW | | | | $\kappa_1$: LSW v. LST | | | |
| Intercept | .114 | .005 | <.001 | .043 | .004 | .003 | .173 | .023 |
| $ln$(Lag) | -.032 | .002 | <.001 | .014 | .000 | .001 | .878 | .010 |
| Day | .003 | .002 | .193 | .032 | .000 | .002 | .857 | .026 |
| Hand | -.002 | .003 | .569 | — | -.000 | .002 | .856 | — |
| Day×Hand | -.009 | .003 | .004 | — | -.002 | .002 | .399 | — |

Table 2.4: Results of mixed effects models analyzing changes in reward valuation ($\kappa_1$) in response to choice location, outcome, win-stay, and lose-shift responses.

decays [$F_{1,88.88} = 19.147$, $p < .001$, $SD\upsilon_0 = .010$]. It also does not vary with the hand used.

| | $\kappa_2$: Right v. Left | | | | $\kappa_2$: Win v. Loss | | | |
|---|---|---|---|---|---|---|---|---|
| Variable | $\beta$ | $SE(\beta)$ | $p\beta$ | $SD\upsilon_0$ | $\beta$ | $SE(\beta)$ | $p\beta$ | $SD\upsilon_0$ |
| Intercept | -.004 | .002 | .097 | .016 | -.002 | .002 | .322 | .011 |
| $ln$(Lag) | .001 | .001 | .427 | .008 | .000 | .001 | .741 | .006 |
| Day | .001 | .001 | .283 | .017 | -.002 | .002 | .199 | .020 |
| Hand | .005 | .002 | .003 | — | -.003 | .002 | .099 | — |
| Day×Hand | -.000 | .002 | .994 | — | -.000 | .002 | .993 | — |
| | $\kappa_2$: WST v. WSW | | | | $\kappa_2$: LSW v. LST | | | |
| Intercept | .021 | .004 | <.001 | .028 | .093 | .005 | <.001 | .040 |
| $ln$(Lag) | -.005 | .001 | <.001 | .030 | -.027 | .002 | <.001 | .014 |
| Day | .002 | .002 | .429 | .030 | .002 | .002 | .293 | .028 |
| Hand | .000 | .003 | .882 | — | -.003 | .003 | 223 | — |
| Day×Hand | -.001 | .003 | .698 | — | -.003 | .003 | 346 | — |

Table 2.5: Results of mixed effects models analyzing changes in punishment aversion ($\kappa_2$) in response to choice location, outcome, win-stay, and lose-shift responses.

### 2.3.5 Sexually dimorphic effects of cannabis use

Given the importance of cannabis use to decision-making, we also sought to reveal how cannabis use and biological sex influenced choice behaviour. Measures of win-stay and lose-shift responding, wins, response times, and response entropy were averaged across each 600-trial session. For each of these measures, a three-way ANOVA assessed the effects of sex (male/female), session (day 1/2), recreational cannabis use status (controls/habitual users), and all two-way interactions. To account for unequal numbers of cannabis users and controls, each ANOVA employed Type-III sums of squares and the sum-to-zero constraint. Degrees of freedom were estimated via Satterthwaite's method.

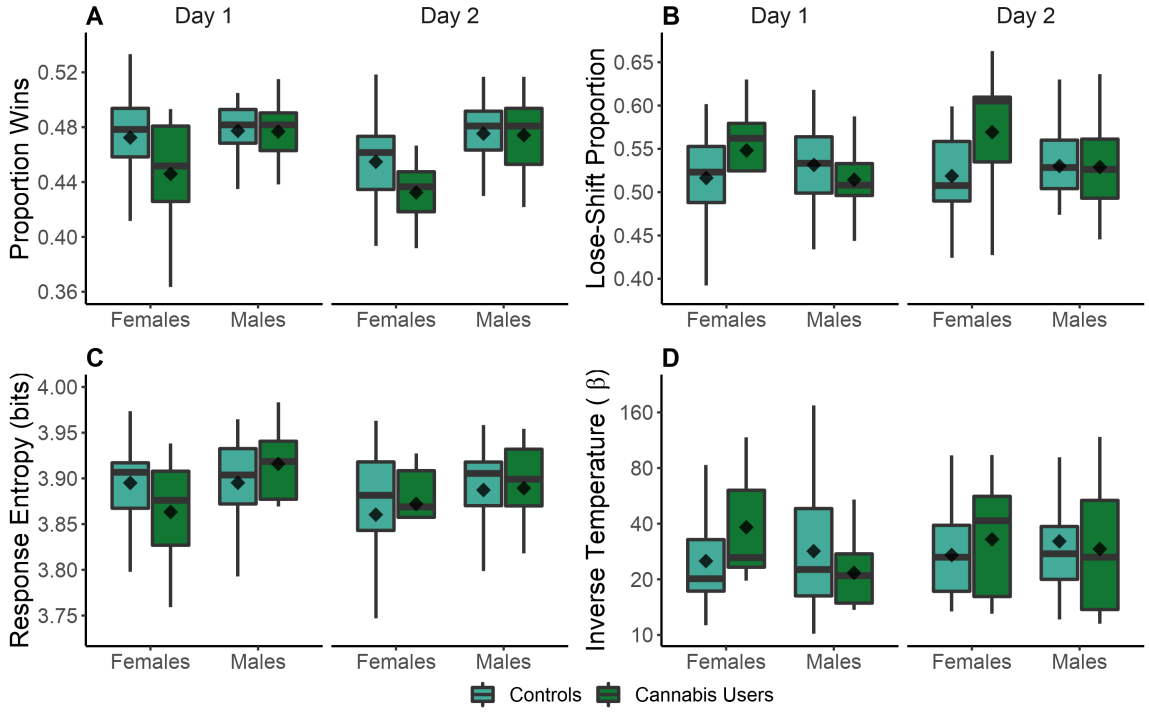


Figure 2.6: Effects of sex and cannabis use on proportion wins (A), lose-shift proportion (B), response entropy (C), and inverse temperature ($\beta$, D).

Task performance significantly decreased between sessions [$F_{1,89} = 4.359$, $p = .040$]. Females also exhibited significantly worse task performance than men [$F_{1,89} = 10.996$, $p = .001$]. As seen in Figure 2.6.A, this effect was driven primarily by a reduced win rate among female cannabis users. However, the effects of cannabis use [$F_{1,89} = 3.049$, $p = .084$] and the sex $\times$ cannabis use interaction [$F_{1,89} = 2.666$, $p = .106$] fell short of significance. Female cannabis users also lose-shifted more than controls. As seen in Figure 2.6.B there was a significant sex $\times$ cannabis use interaction [$F_{1,89} = 4.671$, $p = .033$]. However, no other effects were significant ($p > .170$). Win-stay behaviour ($p > .112$), response entropy ($p > .068$, Fig. 2.6.C), and log response times ($p > .094$) were also unaffected by sex or cannabis use. Overall, habitual cannabis use is associated with weakened executive control over sensorimotor decision-processes in females, as evidenced by their elevated lose-shifting. These results are consistent with Wong et al. [2017b] who reported increased lose-shifting with chronic amphetamine use and those of Cha et al. [2007] reporting impaired spatial learning in females exposed to cannabis.

### 2.3.6 Female cannabis users exhibit greater sensitivity in learning rates and reward valuation

Since recreational cannabis use is associated with altered sensorimotor responding in females, it may also influence the learning processes and feedback valuation that drive behaviour. As in Section 2.3.3 we fit the *Q-learning with forgetting* (FQ) model to subject responses. Mixed-effects models tested whether sex, cannabis use, session, and all two-way interactions influenced $\alpha$, $\kappa_1$, and $kappa_2$ while collapsing over hand used. Change in RL parameters over trials were controlled for as a covariate while

maintaining the assumption of equal slopes. The random effects of subject, session nested within subject, and trials were also included. The effects of sex, cannabis use, and session on inverse temperature ($\beta$) from each session were also tested using a three-way, mixed-effects ANOVA that accounted for random intercepts.

In all cases cannabis use and sex were not significantly associated with differences in $\alpha$, $\kappa_1$, or $\kappa_2$ ($p > .157$). However, $\beta$ varied with a significant sex $\times$ cannabis interaction [$F_{1,40.90} = 4.295$, $p = .046$]. As seen in Figure 2.6.D, female cannabis users exhibited greater response entropy relative to controls, while $\beta$ was reduced in male users. Therefore, female cannabis users were more likely to exploit known choice-outcome associations rather than engage in exploratory or random decisions.

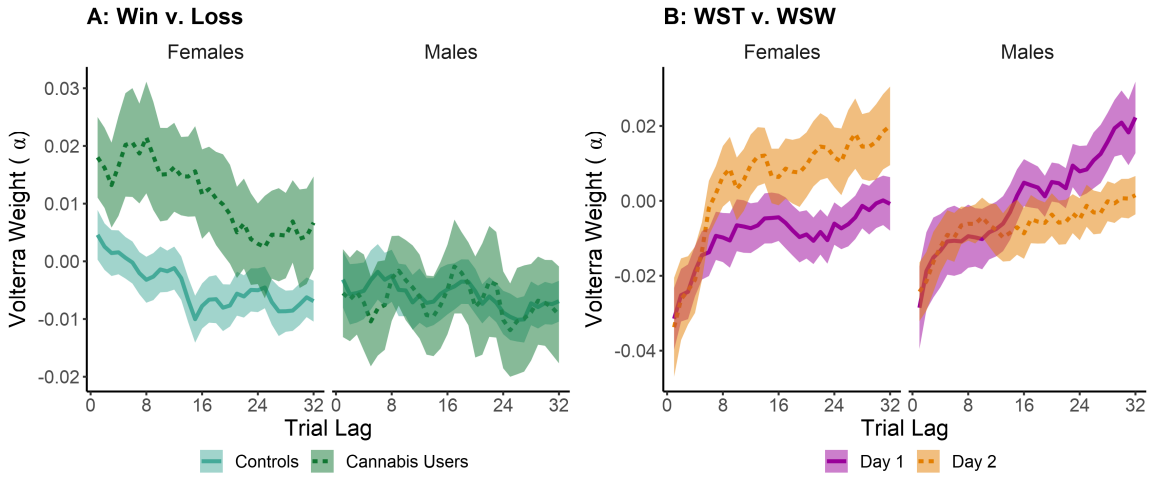


Figure 2.7: A: change in learning rate ($\alpha$) in response to wins (relative to losses), for male & female cannabis users and controls. B: change in learning rate following win-stay responses. The shaded area denotes the standard error of the mean.

While cannabis use did not influence average outcome valuation and learning rates, it may affect our response to choices and feedback. Volterra decomposition was used to investigate how $\alpha$, $\kappa_1$, and $\kappa_2$ changed in concert with choices, feedback, win-stay, and lose-shift responding, and how this change differed between cannabis

users and controls. We used three-way, mixed-effects models to account for effects of sex, cannabis use, session, and all two-way interactions on changes in reinforcement learning parameters, while also controlling for trial lag. The statistical model also included random effects of subject, day within subject, and log-transformed trial lag.

We found no effects of sex or cannabis use on changes in $\kappa_1$ or $\kappa_2$ induced by winning outcomes, win-stay, or lose-shift responding ($p > .190$). The change in learning rates following winning outcomes was significantly influenced by sex [$F_{1,87.73} = 5.529$, $p = .021$], session day [$F_{1,89.00} = 4.351$, $p = .040$], and trial lag [$F_{1,88.80} = 9.165$, $p = .003$]. Females exhibit a greater increase in $\alpha$ following rewards (Fig 2.7.A). As with task performance, lose-shift behaviour, and $\beta$ this difference is primarily driven by female cannabis users. Therefore, females (particularly cannabis users) are more influenced by proximal outcomes following wins.

Females also demonstrated different volterra weights over sessions. As seen in Figure 2.7.B, learning rates ($\alpha$) declined sharply after win-stay responses and returned to baseline over subsequent trials [$F_{1,88.85} = 26.416$, $p <= .001$, $SD\upsilon_0 = .019$]. On day two female learning rates recovered much faster, as indicated by a significant sex $\times$ session interaction [$F_{1,89.00} = 3.988$, $p = .049$]. However, the effects of cannabis, the cannabis $\times$ sex, and cannabis $\times$ session interactions were not significant ($p > .802$).

## 2.4 General Discussion & Conclusions

In the present study we investigated the decision-making processes that drive reward seeking behaviour and loss aversion. The sensorimotor striatum, including the putamen, caudate nucleus, and nucleus accumbens have been hypothesized to support win-stay and lose-shift responding in different experimental contexts [Skelin et al.,

2014, Gruber et al., 2017, Thapa and Gruber, 2018, Grospe et al., 2018]. However, these striatal circuits also are associated with motor control, action selection, and egocentric and allocentric processing of space. Given the multiple roles the striatum has in economic decision-making strategies, motor action selection, and spatial processing, we hypothesized that the lose-shift and win-stay may be calculated with respect to the hand used and choice location.

We demonstrated that in humans, the win-stay and lose-shift behaviours comprise distinct decision strategies supported by separate cognitive processes. While suppression of both these behaviours contribute equally to performance at the Matching Pennies task, they remain uncorrelated across subjects. Moreover, lose-shift suppression remained stable over the experimental session, while win-stay responding improved with practice. However, this improvement did not persist between experimental sessions, but was specific to the session in which it was learned.

Unlike skills such as typing, our data suggest that choice value is not solely driven by motor processes. Win-stay and lose-shift responding, response entropy, learning rates, reward strength, and punishment aversion were all unaffected by use of the dominant or non-dominant hands (Figs. 2.2 and 2.4). Instead, we found that these sensorimotor response strategies are the joint product of motor, visuospatial, and reward systems. Lose-shifting was strongly dependent on choice-location, regardless of the hand being used. For example, during use of the right (dominant) hand, participants would lose-shift away from the left choice but tend to lose-stay after selecting the right choice. Conversely, participants were more likely to win-stay following choices made ipsilateral to the hand being used (e.g., right hand and right choice). Consequently, the value of choices may not be calculated relative to their visual identity,

but according to their position in space.

However, the spatial bias in win-stay responding was only present during use of the dominant hand. When choices were made with the left hand, participants were equally likely to win-stay, regardless of where choices were located relative to the body. Consequently, motor action has a strong modulatory effect on win-stay responding. The putamen (DLS) is known to support win-stay responding [Packard and White, 1991, McDonald and White, 2013] and receives inputs from the somatosensory and motor cortices [Brasted et al., 1999, Pan et al., 2010, Malach and Graybiel, 1986]. Moreover, our finding that one's reference frame (i.e., hand used) affects win-stay behaviour supports previous findings that the putamen supports egocentric (i.e., self-referential) processing of space [Kesner and DiMattia, 1987, Palencia and Ragozzino, 2005]. In right-handed individuals, connectivity between the supplementary motor area and putamen corresponding to the right hand is much greater than that corresponding to the left hand [Pool et al., 2014]. Hand preference is also associated with asymmetries in dopamine activity between the left and right putamen [de la Fuente-Fernández et al., 2000] and caudate nucleus [Evenden and Robbins, 1984] another structure associated with egocentric spatial processing [Ragozzino et al., 2002, Postle and D'Esposito, 2003, Possin et al., 2017]. Consequently, asymmetries in spatial win-stay biases are reflected in the asymmetric connectivity between the putamen and motor cortex.

Lose-shift responding was not strongly influenced by use of the dominant hand, suggesting it is not calculated in egocentric spatial coordinates. Instead it may be supported by the medial striatum and nucleus accumbens [Skelin et al., 2014,

Clarke et al., 2008] regions also associated with alternation between response strategies [Ragozzino et al., 2002, Ragozzino, 2007, McDonald et al., 2008]. The nucleus accumbens is not strongly connected to the motor cortex, but receives inputs from the hippocampus, profrontal cortex, amygdala, and hypothalamus Gerfen [1984], Voorn et al. [2004], Kelley et al. [1982]. Consequently, it represents space in an allocentric (world-centered) frame of reference De Leonibus et al. [2005].

However, our findings that the win-stay may be supported by egocentric spatial processing in the putamen conflicts with several anatomical studies in rats. Gruber et al. [2017] demonstrated that lesions to the nucleus accumbens core, rather than the putamen, reduce rates of win-stay responding. Unlike the present results, the win-stay and lose-shift exhibit a strong negative correlation in rats, suggesting they may be part of a different decision strategy than that used by humans [Gruber and Thapa, 2016, Gruber et al., 2017]. In rats, both the win-stay and lose-shift are driven by habitual, sensorimotor response strategies supported by the lateral striatum [Skelin et al., 2014]. However, humans may also choose to actively lose-shift as part of a conscious goal-directed strategy to outsmart the computer opponent, a strategy unavailable to rats. Goal-directed control of behaviour is mediated by the nucleus accumbens [Burton et al., 2015] while habitual, associative responses are supported by the putamen [Burton et al., 2014]. Consequently, the context in which win-stay or lose-shift responses are made determines the neural circuitry it is supported by [McDonald et al., 2008]. Given that people dedicate greater attention and cognitive resources to avoiding losses, relative to reward-seeking behaviour, it follows that lose-shift suppression is supported by goal-directed systems in the nucleus accumbens [Sokol-Hessner et al., 2013, Lejarraga and Hertwig, 2017].

Using the *Q-learning with forgetting* model, we further explored the relationship between choice valuation, spatial location, and motor action. We demonstrated that the value assigned to wins ($\kappa_1$) and losses ($\kappa_2$) significantly increased following a win-stay or lose-shift response. However, the relationship between $\kappa_1$ and win-stay responding was much greater during use of the dominant hand. Consequently, spatial biases in win-stay responding and the value of rewards that drives this behaviour are both modulated by motor action. Punishment aversion ($\kappa_2$), while responsive to the lose-shift behaviour, was not modulated by motor action. However, alternating between choice locations did influence changes in $\kappa_2$ (after accounting lose-shift/$\kappa_2$ relationship), indicating that the value assigned to punishments are specific to spatial locations. Reward strength ($\kappa_1$) was not significantly associated with changes in location, indicating spatial processing is more important to punishment aversion.

Finally, we demonstrated a sexually dimorphic relationship between habitual cannabis use and decision-making. In females, self-reported cannabis use was associated with decreased task performance due to increased rates of lose-shift behaviour, decreased mixed-strategy responding, and a greater tendency to exploit known choice-outcome associations rather than engage in random decisions. However, male cannabis users exhibited no change in task performance and moderately better suppression of lose-shift responding. Female cannabis users were also more affected by recent outcomes that controls, as evidenced by increased learning rates in response to rewards. This sexually dimorphic effect of cannabis use has been previously found in rats and humans. Females are more susceptible to drug tolerance and sensitization than are males [Wakley et al., 2014, Robinson, 1988], and are more likely to exhibit mood or anxiety disorders co-morbid with cannabis use [Zilberman et al.,

2003]. The greater effect of cannabis use on female decision-making may be due to the presence of estrogen, which enhances striatal dopamine release and alters prefrontal function in response to psychoactive drugs [Becker, 1999]. Male and female rats high in estrogen exhibit prefrontal dysfunction in response to dopamine-enhancing drugs [Shansky et al., 2004, Febo et al., 2005, Sárvári et al., 2014]. Therefore, the heightened susceptibility of the PFC may explain why only females with ASSIST scores show elevated lose-shift responding. However, replication of these results and in-depth anatomical studies are needed to confirm this hypothesis.

Together, our findings indicate some of the cognitive processes that underlie the win-stay and lose-shift behaviours and help clarify the debate as to their anatomical origins. Both the win-stay are lose-shift behaviours are strongly influenced by choice location, indicating that choice value may be determined by spatial location rather than visual identity. Spatial biases in win-stay responding were also influenced by the hand used to make decisions, while lose-shift behaviour was not. Consequently, the win-stay may be processed in self-referential (egocentric) spatial coordinates. Finally, recreational cannabis use was associated with a sexually dimorphic increase of lose-shift responding in females, but decreased responding in males. These results highlight the real-world effects cannabis use have on human decision-making.

# Chapter 3

# The Spatial Distribution of Choice Value

## 3.1 Introduction

In Chapter 2 we demonstrated that the win-stay and lose-shift strategies are influenced by choice location and motor action. Therefore, choice value may be unrelated to visual choice identity and instead driven by its position relative to previous choices and the observer. If so, the likelihood of shifting between choices may follow a spatial tuning function that is solely driven by the continuous distance between choices. Alternatively, people may express biases towards certain spatial locations, but not ones that follow a predictable distribution through space.

In addition, most studies of the win-stay and lose-shift strategies have only considered decision-making between two choices. Under these conditions, the win-stay and lose-shift comprise two independent behavioural strategies, driven by separate sub-regions within the striatum. However, we often have to make decisions between

any number of options. The striatum is associated with the the maintenance of working memory for locations and choices. For example, the dorsomedial striatum encodes learned spatial-motor sequences in rats [Akhlaghpour et al., 2016]. In humans, striatal dopamine synthesis predicts verbal working memory capacity [Cools et al., 2008]. Reutskaja et al. [2018] demonstrated that the putamen and caudate nucleus are the most active when deciding between twelve choices, relative to six or twenty-four. As the number of choices increases, the benefits of having more options available are offset by the greater working memory load. Under this greater load people may abandon the independent win-stay and lose-shift behaviours for a less memory-intensive strategy.

Furthermore, the spatial distribution of the win-stay and lose-shift behaviours may be influenced by the number of choices present. For example, increasing the number of choices present within an area decreases the distance between adjacent choices. If decisions are calculated in egocentric coordinates this increase will not influence the spatial distribution of decisions. However, if allocentric processing drives behaviour then shifts will be calculated relative to the number of choices shifted, rather than their absolute positions. Consequently, the distance shifted between choices will decrease as their number increases.

Therefore, we investigated how decision-making changes in competitive environments when multiple choices are available, ranging from 2 to 18. These choices were arranged in a ring, allowing us to represent the win-stay and lose-shift responses as a distribution of angles on a circle. We report here that decision-making between two choices is characterized by independent win-stay and lose-shift strategies. However, as the number of choices increase, subjects adopt a single win-shift/lose-shift strategy.

Both behaviours also exhibit a greater degree of temporal decay as choices increase, similar to that reported by Gruber and Thapa [2016] during lose-shift responding.

We also demonstrate that sensorimotor decision strategies are not tied to discrete choices, but are distributed as a continuous function of space. The pattern of behaviour exhibited highlights multiple sub-strategies that constitute the lose-shift response. Following losses, subjects completely avoid both their previous choice and those a moderate distance away. Instead, they either select choices directly adjacent to the previous loss (foraging) or shift as far away as possible (complete avoidance). This spatial tuning-function persists regardless of the number of choices present. However, as the number of choices increase and subjects adopt a win-shift/lose-shift strategy, this function becomes more sharply tuned.

Finally, we provide evidence to support our hypothesis that the win-stay and lose-shift are driven by egocentric and allocentric processing respectively. Following wins, subjects tend to move the same angle regardless of how many choices are present. Consequently, choice value following wins is not influenced by the positions of choices relative to one another. However, as the number of choices increase the angle moved following losses decreases, indicating loss aversion is calculated in allocentric coordinates. Overall, these data (57,600 trials from 48 participants) demonstrate that while both the win-stay and lose-shift are driven by spatial processing, they comprise distinct strategies that vary in their reliance on egocentric and allocentric coordinate systems.

## 3.2 Methods

### 3.2.1 Behavioural Task

Participants played the game "Matching Pennies", which was described in section 2.2 of Chapter 2. This task consisted of six, twelve, or eighteen choices arranged in a circle on a 15″ touchscreen monitor (Fig. 3.1). On each trial the subject selected one of the available choices and the computer opponent predicted the choice made using the twenty-four history matching algorithms mentioned in Chapter 2. A successful prediction by the computer resulted in a losing trial, which was indicated by the presentation of a 250 Hz auditory tone and the text "You Lose" for 1.5 s. An nsuccessful prediction resulted in a winning trial, which was indicated by the presentation of a 600 Hz tone and "You Win" for 1.5 s. The optimal strategy in this game is for the participant to respond randomly across trials.

The number of available choices varied between experimental groups. To equate the probability of winning across groups, on each trial the computer would predict which *half* of the available choices the participant was most likely to select. For example, in the twelve-choice condition the computer would predict the six choices the subject was most likely to make on each trial, whereas in the six-choice condition the computer predicted the three most likely choices. Consequently, a subject who responded randomly would have a 50% chance of winning on each trial regardless of the number of choices available.

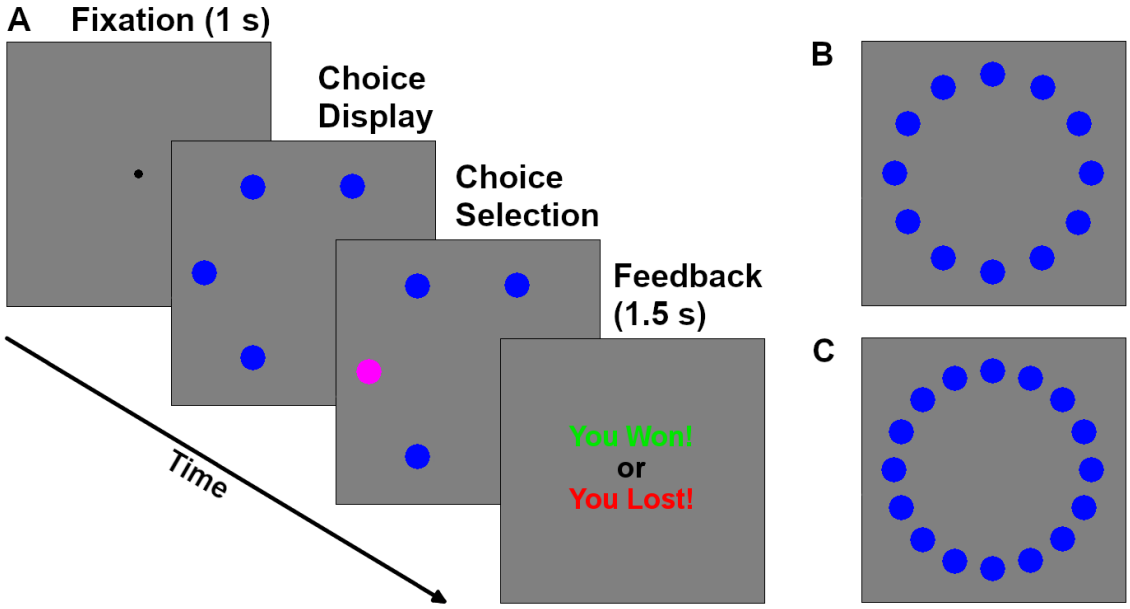


Figure 3.1: Timeline of trials in the matching pennies game and task display in the six (A), twelve (B), and eighteen (C) choice conditions.

### 3.2.2 Procedure

All procedures and experimental tasks were approved by the McMaster University Research Ethics Board. Forty-eight right handed subjects (17 males, mean age = 19.94, $SD = 2.44$) from McMaster University participated in the study in exchange for payment (10$ per hour). After providing informed consent, participants played two sessions of 600 trials of Matching Pennies on two consecutive days. On each day, the task display consisted of either six, twelve, or eighteen choices. These conditions were counterbalanced so that each subject experienced two of the possible three conditions over both days. Each experimental session was comprised of four blocks of 150 trials. On each block the choice circle had a radius of 2.11″, 2.66″, 3.20″, or 3.75″. The order of these four conditions was randomized during each session. However, the position of choices remained fixed during each experimental block. Subjects used their right

hand to make choices throughout each session, keeping their left hand affixed to a pressure sensor on the left side of the task display. Removing their hand would pause the experiment between trials until it was returned to the sensor.

Participants were informed that they would win nothing each time the computer predicted their choice and 3¢ each time it could not, rounded up to the nearest $5 upon completion of the experiment. Although participants were informed they were playing a competitive game against a computer opponent, they were given no guidance regarding optimal decision-making strategies. After task completion, participants completed the South Oaks Gambling Screen, the alcohol, smoking and substance involvement screening test (ASSIST v3.0), Adult ADHD Self-Report Scale v1.1, and an additional demographic questionnaire. Fourteen subjects (5 male, 9 female) met the ASSIST criteria for habitual cannabis use, and twenty-two (8 male, 14 female) met the criteria for habitual use of any recreational drug.

### 3.2.3 Analysis

Forty-eight subjects completed a total of 57,600 trials of "Matching Pennies" against a computer opponent. This data was combined with that of twenty-three subjects from Chapter 2 who used their right hand to select between two choice options on both days of the experiment. For each participant we analyzed the number of wins experienced, win-stay responses, lose-shift responses, and decision times. Shift responses were defined as selecting two different choices on sequential trials, regardless of the angle between them. Decision times were measured as the time to make a response following presentation of the choice selection screen. They were normalized using the log transform and averaged after removing 1895 RTs $< 50$ ms or $> 30$ s.

In order to analyze how the number of choices influenced decision-making, wins, win-stays, lose-shifts, and log decision-times were averaged across each 600 trial session. The linear and quadratic effects of log-transformed choice number on each behavioural measure was assessed via linear regression. We also investigated whether sex and cannabis use influenced task behaviour using 2×2 ANCOVAs that controlled for choice number as a covariate. In each case we tested for the effects of sex, cannabis use, linear and quadractic effects of choice number, and all two-way interactions. Each model employed type-III sums of squares and the sum-to-zero constraint.

To further explore whether sensorimotor response strategies change as the number of choices increase, we investigated the relationship between decision times and behaviour. Generalized binomial, mixed-effects models analyzed whether win-stay or lose-shift tendencies changed as log decision-times or choice numbers increased. Each model was fit in R using the lme4 package [Bates et al., 2014], employed the logit-link function. We assessed linear, quadratic, and cubic effects of decision time, a linear term for choice number, and all 2-way interactions. A maximal random-effects structure was also included to control for the random intercepts and linear, quadratic, and cubic decision-time slopes for each subject. Degrees of freedom and $p$-values in all mixed-effects models were calculated via Satterthwaite's method. Linear, quadratic, and cubic trends were assessed using orthogonal polynomials calculated via the *poly* function in R [R Core Team, 2019].

Finally, we tested whether the spatial distribution of choices can be represented by a relatively simple tuning function. Response proportions towards each choice location (i.e., egocentric bias) and the angle shifted from one trial to the next (i.e., allocentric bias) were analyzed using mixed-effects multiple regression. As choices

were arranged on a continuous circle, choice positions and shifting behaviour were represented as sin and cos transformed angles ($\theta$). For examples, choices located at -90°, 0°, +90°, and 180° correspond to those on the left, top, right, and bottom of the choice array respectively. Similarly, shifting 90° clockwise or counterclockwise were represented as angles of -90° and +90°. Consequently, in measuring egocentric response biases, $\sin(\theta)$ and $\cos(\theta)$ represent right/left and top/bottom asymmetries in choice preferences respectively. In measuring the angle shifted between each trial, $\sin(\theta)$ and $\cos(\theta)$ represent biases clockwise/counterclockwise or towards/away from the previously selected choice. Therefore, we also refer to $\cos(\theta)$ and $\sin(\theta)$ as vertical and horizontal choice position.

For each model, the effects of $\sin(\theta)$, $\cos(\theta)$, previous choice outcome (win v. loss), number of choices (6, 12, or 18), and all two-way interactions were assessed. Due to the different numbers of choices available in each condition, response proportions were multiplied by the total number of choices available. In this way, random responding would result in an average normalized proportion of 1 in every choice condition. As response proportions always averaged to 1 and were arranged on a continuous circle, the intercept for every subject and condition was also always 1. Therefore, a mixed-effects model accounting for the random effects of subject and condition was deemed unnecessary. Instead, ordinary polynomial regression was used. Due to the large number of parameters already modelled, the effects of sex and cannabis use on each spatial tuning function were not considered.

In each analysis we also considered whether the physical radius of the choice array (2.11″, 2.66″, 3.20″, or 3.75″ in each block) influenced the spatial-tuning function, win-stay responding, lose-shift responding, or task performance. In every case display

radius had no effect on behaviour ($p > .249$). Therefore, all behavioural measures were averaged across trial blocks, statistical analyses collapsed across radius, and the effects of display radius were not considered further.

Although computational modeling and measures of response entropy were central to Chapter 2, they were not included in this study. The number of free parameters required and the difficulty in comparing between the 6-18 choice conditions made this analysis prohibitive. For example, in the 18-choice condition, 104976 four-choice sequences are possible.

## 3.3 Results

### 3.3.1 Subjects adopt a win-shift/lose-shift strategy as choices increase

We first investigated whether sensorimotor response strategies change as the number of choices increase. Second-order polynomial regression was used to evaluate the effect of choice number on task performance (i.e., the proportion of wins), lose-shift responding, win-stay responding, and decision times. The results of each model are provided in Table 3.1.

As seen in Figure 3.2.A & B, lose-shift responding increased [$F_{2,139} = 953.47$, $p < .001$] and win-stay responding decreased [$F_{2,139} = 472.61$, $p < .001$] when more choices were present, despite providing no advantage to performance. Moreover, after accounting for the effects of choice numbers on task performance [$F_{2,136} = 9.861$, $p < .001$], increased lose-shifting was negatively associated with performance [$\beta$ = -.240, $F_{2,136} = 10.431$, $p < .001$] and win-stay behaviour was unassociated ($p = .114$).

| Variable | $\beta$ | $SE(\beta)$ | $t$ | $p\beta$ | $\beta$ | $SE(\beta)$ | $t$ | $p\beta$ |
|---|---|---|---|---|---|---|---|---|
| | Lose-Shift, $R^2$ = .932, N = 142 | | | | Win-Stay $R^2$ = .872 | | | |
| Intercept | .794 | .004 | 188.03 | < .001 | .248 | .006 | 38.371 | < .001 |
| $ln$(Choice) | 2.096 | .050 | 41.66 | < .001 | -2.308 | .077 | -30.018 | < .001 |
| $ln$(Choice)$^2$ | -.658 | .050 | -13.08 | < .001 | .511 | .077 | 6.646 | < .001 |
| | % Wins $R^2$ = .166 | | | | $ln$ Decision Time $R^2$ = .333 | | | |
| Intercept | .457 | .003 | 142.197 | < .001 | -.226 | .031 | -7.366 | < .001 |
| $ln$(Choice) | .010 | .038 | .253 | .801 | 2.999 | .366 | 8.195 | < .001 |
| $ln$(Choice)$^2$ | .201 | .038 | 5.256 | < .001 | -.550 | .366 | -1.501 | .136 |

Table 3.1: Results of models analyzing the relationship between number of choices (log transformed) and win-stay/lose-shift responding.

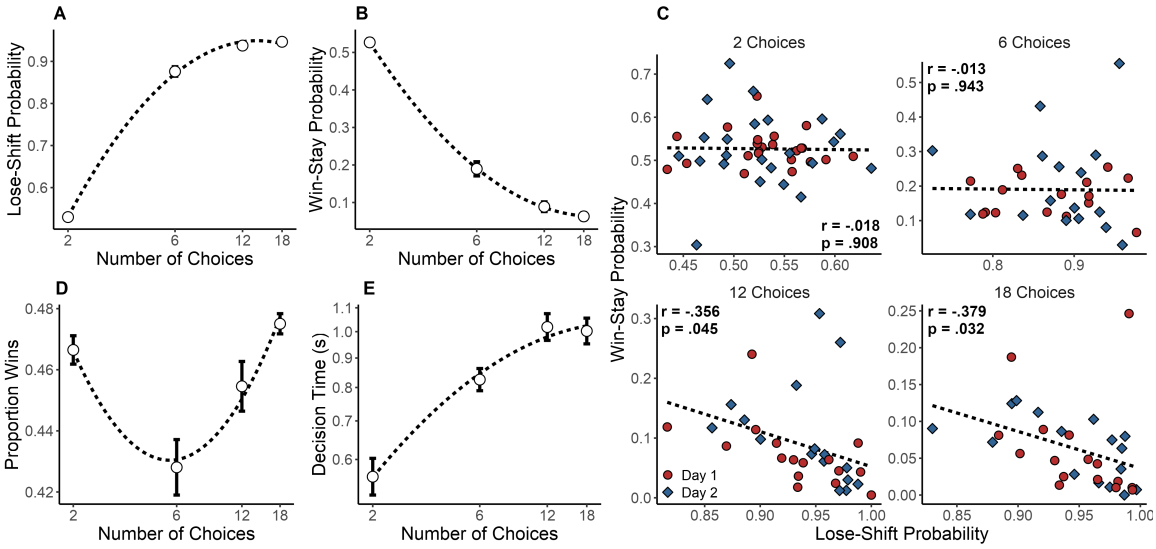


Figure 3.2: Effects of the number of choices present on lose-shift (A) and win-stay responding (B), decision times (D), and proportion of wins (E). C: correlations between lose-shift and win-stay responding in the two, six, twelve, and eighteen choice conditions.

Instead of being driven by task demands, the change in behaviour was due to the adoption of a single win-shift/lose-shift response strategy. As seen in Figure 3.2.C, the correlation between win-stay and lose-shift behaviour increases with the number of choices. When two choices are available lose-shift and win-stay tendencies are

uncorrelated [$r(44)$ = -.018, $p$ = .908], which is consistent with the idea that lose-shift and win-stay behaviour are controlled by independent cognitive processes. The lack of correlation is maintained for six choices [$r(30)$ = -.013, $p$ = .943], but breaks down when twelve [$r(30)$ = -.356, $p$ = .045] or eighteen [$r(30)$ = -.379, $p$ = .032] choices are present.

This change in strategy was not associated with a monotonic change in task performance. Instead, as the number of choices increased from two to six, task performance decreased before steadily increasing again (Fig. 3.2.D). This significant quadratic relationship [$F_{2,139}$ = 13.844, $p < .001$] may be due to the observed transition between the uncorrelated win-stay/lose-shift and correlated win-shift/lose-shift strategies. We hypothesize that participants in the low and high choice conditions are equally effective at using different cognitive processes in each context. However, the six-choice condition lies on a balance point between these two strategies and participants were unable to use a single choice strategy optimized for that condition. This hypothesis is supported by the fact that lose-shift behaviour and performance were only correlated in the 6-choice condition [$r(30)$ = -.358, $p$ = .044], indicating that behaviour was more predictable. Finally, decision times exhibited an asymptotic trajectory that increased with the number of choices before levelling out (Fig. 3.2.E).

### 3.3.2 The memory trace supporting win-shift responding decays over time

We have demonstrated that as the number of choices available increase, a single process begins to govern both the win-stay and lose-shift behaviours. Gruber and Thapa [2016] and Ivan et al. [2018] have demonstrated that when win-stay and lose-shift

responses are uncorrelated they exhibit different temporal properties. Specifically, during the time following a reward or punishment, the tendency to lose-shift decays while win-stay behaviour does not change. Therefore, further consideration of decision times may indicate whether a single strategy governs decision-making in the high-choice condition. In particular, we tested whether the temporal characteristics of win-stay and lose-shift behaviour becomes more similar with increasing choice number.

To explore how the relationship between decision times and wst/lsw behaviour changes as choice number increases, we used mixed-effects binomial regression models with a logit-link function. The fixed effects of log-transformed choice number, linear, quadratic, and cubic effects of decision time, and all 2-way interactions were estimated. Random intercepts and linear, quadratic, and cubic decision time slopes for each subject were also included, in order to control for variability between subjects.

| | Lose-Shift, N = 57147 | | | | Win-Stay, N = 49140 | | | |
|---|---|---|---|---|---|---|---|---|
| Variable | $\beta$ | $SE(\beta)$ | $p\beta$ | $SD\upsilon_0$ | $\beta$ | $SE(\beta)$ | $p\beta$ | $SD\upsilon_0$ |
| Intercept | -.480 | .079 | <.001 | .494 | .925 | .069 | <.001 | .430 |
| RT | .150 | .080 | .061 | .393 | .004 | .076 | .962 | .361 |
| $RT^2$ | -.048 | .027 | .080 | .092 | -.027 | .029 | .357 | .080 |
| $RT^3$ | -.032 | .016 | .044 | .058 | .020 | .014 | .155 | .027 |
| $ln$(Choice) | 1.315 | .048 | <.001 | — | -1.355 | .042 | <.001 | — |
| RT × $ln$(Choice) | .081 | .055 | .139 | — | -.129 | .051 | .012 | — |
| $RT^2$ × $ln$(Choice) | -.059 | .019 | .002 | — | .067 | .020 | <.001 | — |
| $RT^3$ × $ln$(Choice) | -.012 | .012 | .256 | — | .026 | .009 | .006 | |

Table 3.2: Results of mixed effects models analyzing the relationship between response times and log-transformed number of choices on the win-stay and lose-shift responses.

The results of our analysis are provided in Table 3.2. Lose-shift responding varied as a function of choice number [$\chi^2(1) = 766.275$, $p < .001$], cubic trends in decision

time [$\chi^2(1) = 4.053$, $p = .044$], and a choice × quadratic time interaction [$\chi^2(1) = 9.976$, $p = .001$]. No other effects or interactions were present. As seen in Figure 3.3 lose-shift responding exhibits a strong time decay beginning between 3-10 seconds after reward presentation. Therefore, the temporal properties of the process supporting lose-shift responding do not change, regardless of how many choices are present. However, as the number of available choices decrease, both rates of lose-shifting and of temporal decay increase. While this time frame for decay may seem excessively long, it is similar to that reported in humans in controlled settings [Ivan et al., 2018]. Moreover, our analysis of responses following long decision times was based on sufficient data, as 4397 lose-shift decision times exceeded 3 seconds and 511 exceeded 10 seconds.

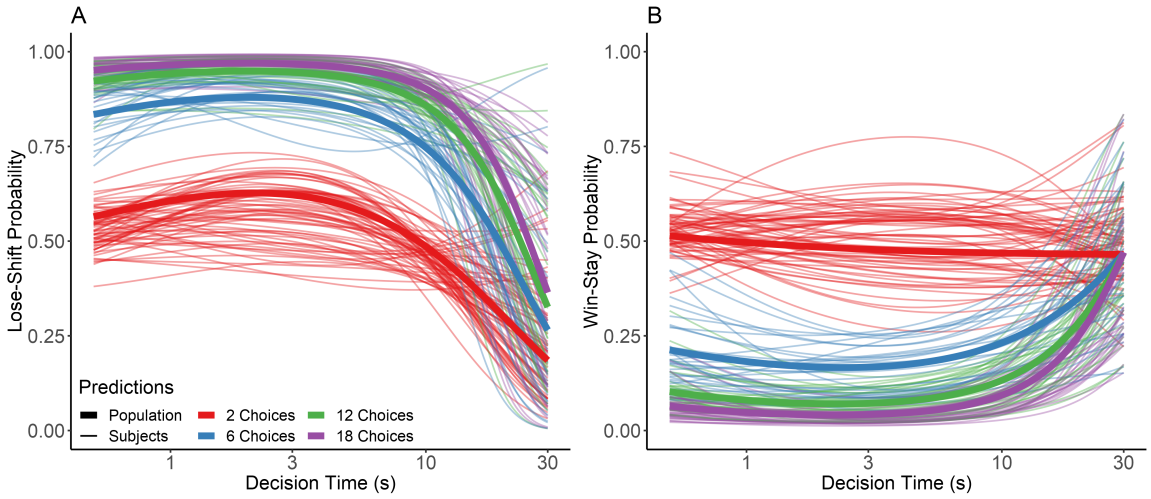


Figure 3.3: Changes in rates of lose-shift (A) and win-stay (B) responding as a function of the amount of time preceding a decision and the number of choices available. The line of best fit is provided for each population (fixed effects) and subject (fixed and random effects).

Win-stay responding exhibited significant effects of choice number [$\chi^2(1) = 1074.548$, $p < .001$], choice × linear [$\chi^2(1) = 6.370$, $p = .012$], choice × quadratic [$\chi^2(1) =$

11.458, $p < .001$], and choice × cubic [$\chi^2(1) = 7.688$, $p = .006$] decision time interactions. As reported by Ivan et al. [2018], we found that win-stay behaviour does not change over time in the two-choice condition. As the number of choices increase, participants adopt a strong win-shift tendency (Fig. 3.3.B). This win-shift responding decays over time, with the rate of decay increasing in the high-choice conditions. Therefore, the temporal properties of the win-stay mirror those of the lose-shift. These results suggest that when more than two choices are present, decision-making following both wins and losses is governed by a single, shifting-based strategy.

### 3.3.3 Male cannabis users suppress lose-shift responding

In Chapter 2 cannabis use had a sexually dimorphic effect on sensorimotor response strategies. Female cannabis users lose-shifted more than controls, while lose-shifting was reduced in male cannabis users. Therefore, we analyzed whether decision-making varied with cannabis use when subjects employed the win-shift/lose-shift strategy in the high-choice condition. The effects of sex, cannabis use, log-transformed choice number, and all two-way interactions on task performance, win-stay responding, lose-shift responding, and decision times were tested with 2 × 2 ANCOVAs.

Lose-shift responding exhibited a significant effect of sex [$F_{1,132} = 5.293$, $p = .023$] and a sex × cannabis use interaction [$F_{1,132} = 4.274$, $p = .041$] after controlling for choice number [$F_{1,132} = 713.634$, $p < .001$]. As seen in Figure 3.4.A male cannabis users were better able to suppress lose-shift responding as the number of choices increased. Female cannabis users lose-shifted more than controls when two-choices were present, but they did not differ in other conditions. All other effects and interactions were not significant ($p > .216$ in each case).

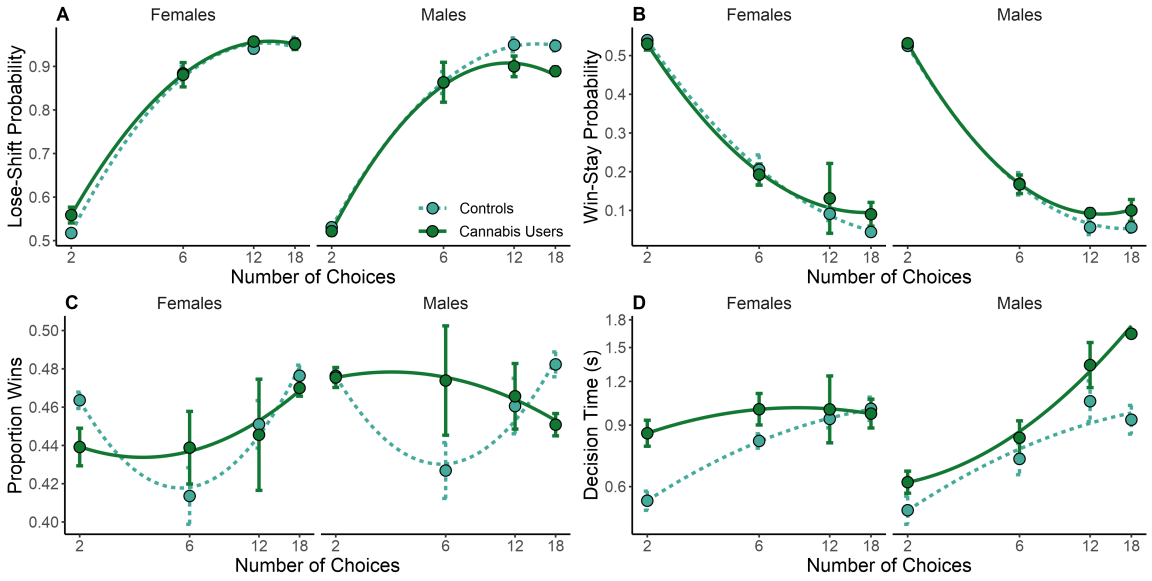


Figure 3.4: Effects of cannabis use and sex on lose-shift responding (A), win-stay responding (B), task performance (C), and log decision times (D).

Win-stay responding was not influenced by sex, cannabis use, or any interactions ($p > .229$) after controlling for the effects of choice number [$F_{2,132} = 346.213$, $p < .001$]. However, task performance did vary between the sexes [$F_{1,132} = 4.521$, $p = .035$]. As seen in Figure 3.4.C, males exhibited better task performance than females in almost all conditions. While the effects of choice number were significant [$F_{2,132} = 6.379$, $p = .002$] no other effects or interactions were ($p > .091$). Finally, both male and female cannabis users took longer to make decisions than controls (Fig. 3.4.D). Consequently, there were significant effects of sex [$F_{1,132} = 6.526$, $p = .012$] and choice number [$F_{2,132} = 27.083$, $p < .001$] on log decision-times. The effects of sex and all other interactions were not significant ($p > .422$).

Overall, the current experiments and those in Chapter 2 suggest the effects of cannabis on decision-making are sexually dimorphic: female cannabis users are less able to suppress habitual lose-shift behaviour while male cannabis users exhibit greater

suppression.

### 3.3.4 Choice valuation follows a spatial tuning function

Finally, we analyzed whether choice value is calculated in egocentric spatial coordinates, allocentric coordinates, or is calculated according to choice identity, irrespective of spatial location. If value is calculated in egocentric coordinates, participants should prefer specific locations relative to the hand being used. If allocentric processing dominates, patterns should emerge in the angle shifted from one choice to the next. In both cases, choice preferences should follow a spatial tuning function similar to that of head direction cells in the entorhinal cortex. For example, surround inhibition has also been observed in the motor system. During movement, the basal ganglia inhibits competing motor actions that may interfere with execution of the desired response [Mink, 1996, Sohn and Hallett, 2004]. Therefore, we hypothesize this inhibition will extend to locations associated with specific motor actions.

To test whether participants preferred specific choices we used polynomial regression fit to preferences for each choice. Response proportions were multiplied by the total number of choices available, ensuring they would average to one in each condition. As this transform ensured responses had an intercept of one for each participant and choice condition, a mixed-effects model that controlled for random intercepts was not required.

In fitting egocentric response biases, we found that a model accounting for the effects of $\cos(\theta)$, $\sin(\theta)$, $\sin(\theta)^2$, and all two-way interactions between $\cos(\theta)$, $\sin(\theta)$, and the number of available choices provided the best fit of scaled response proportions. This model fit significantly better than one limited to the linear effects of choice angle

[$F_{3,2295} = 15.531$, $p < .001$], and was no worse than a model accounting for $\sin(\theta)^3$ [$F_{3,2292} = .231$, $p = .875$]. Moreover, our model fit equally well as one accounting for previous outcome on choice preferences [$F_{3,2292} = .320$, $p = .811$]. Consequently, egocentric biases towards certain choice locations was not influenced by reward or punishment. Instead they were associated with affordances of the task environment in relation to the body.

| | Angle Moved, N = 2304, $R^2 = .109$ | | | |
|---|---|---|---|---|
| Variable | $\beta$ | $SE(\beta)$ | $t$ | $p\beta$ |
| Intercept | 1.000 | .007 | 149.317 | <.001 |
| Vertical | .207 | 1.114 | .186 | .853 |
| Horizontal | 4.249 | 1.056 | 4.022 | <.001 |
| Horizontal$^2$ | .292 | 1.056 | .277 | .782 |
| Vertical × Horizontal | 181.727 | 21.822 | 8.328 | <.001 |
| Vertical × Horizontal$^2$ | 71.296 | 23.904 | 2.983 | .003 |
| Vertical × Choice | -.210 | .072 | -2.919 | .004 |
| Horizontal × Choice | -.264 | .072 | -3.679 | <.001 |
| Horizontal$^2$× Choice | .115 | .072 | 1.603 | .109 |

Table 3.3: Results of multiple regression on egocentric spatial preferences. Vertical and horizontal effects refer to $\cos\theta$ and $\sin\theta$ respectively.

The results of our final model are provided in Table 3.3. There were significant effects of $\sin(\theta)$ [$F_{2,2295} = 8.128$, $p < .001$] and the cos × sin [$F_{2,2295} = 39.124$, $p < .001$], cos × choice number [$F_{1,2295} = 8.522$, $p = .003$], and sin × choice number interactions [$F_{2,2295} = 8.054$, $p < .001$]. Figure 3.5.A depicts preferences for each choice as a function of its position within the task display (Fig. 3.1). Choices located at -90°, 0°, +90°, and 180° correspond to those on the left, top, right, and bottom of the choice array respectively. For each choice, scaled response proportions refer to the proportion of times that choice was selected multiplied by the number of total choices available (6, 12, or 18).

As seen in Figure 3.5.A, subjects preferred choices located on a diagonal from the upper-right to lower-left of the task display and avoided those located in the upper-left or lower-right. This spatial pattern coincides with the natural movement of the right arm within the task domain. At the start of each trial subjects were asked to place their right hand on the table, directly in front of the task display, which was oriented vertically towards the participant. Natural extension of the elbow would cause the hand to move from the lower-left to upper-right of the display, causing the observed spatial pattern in choice preferences. Therefore, a sizable proportion of choice behaviour ($R^2 = .109$) can be accounted for by the physical affordances of the task. As choice numbers increased this egocentric spatial bias shifted down and left, towards the starting position of the hand. More choices are associated with a greater working-memory load and effort required during decision-making [Bett et al., 2012, Frey et al., 2015]. Therefore, the tendency to move the arm less when more choices are present may be due to reduced motivation.

We also analysed allocentric response biases as the angle shifted from one choice to the next on subsequent trials. The best fitting model accounted for the effects of $\cos(\theta)$, $\cos(\theta)^2$, $\cos(\theta)^3$, the $\cos(\theta)\times$ previous outcome (win/loss) interactions, and the $\cos(\theta)\times$ choice number interactions. This model fit better than that those only accounting for quadratic [$F_{3,2294} = 5.125$, $p = .002$] or linear [$F_{6,2294} = 10.822$, $p < .001$] trends in angle cosine. While a quartic model fit better than the cubic one [$F_{3,2291} = 3.308$, $p$ .019], it was not considered in order to limit model complexity. Finally, the inclusion of $\sin(\theta)$ (i.e., left/right bias) did not improve model fit [$F_{6,2288} = .448$, $p$ .847], indicating allocentric processing was only associated with the distance from the previously selected target.

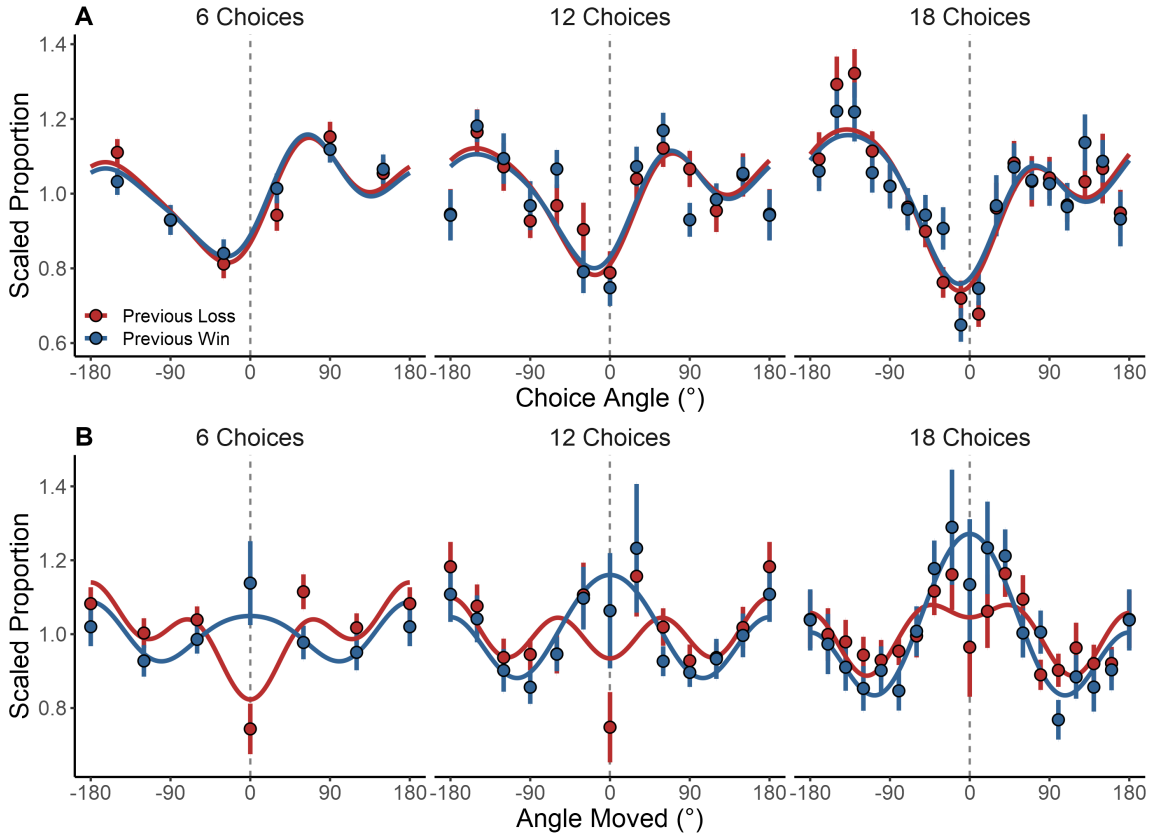


Figure 3.5: A: mean proportion of times each choice was selected following wins and losses, as a function of its position within the choice display (i.e., angle from the top of the screen). Choice angles of 0° and 180° indicate choices located at the top and bottom of the display. B: mean proportion of times participants moved a given angle following wins and losses. Angles of 0°, -90°, or +90° indicate stay responses, shifting 90° counterclockwise, or 90° clockwise. Response proportions were scaled by multiplying against the total number of choices available (6, 12, or 18).

Model parameters are provided in Table 3.4. There were significant effects of $\cos(\theta)$ [$F_{3,2294} = 6.897$, $p < .001$], the $\cos(\theta)\times$ choice [$F_{3,2294} = 7.777$, $p < .001$], and $\cos(\theta)\times$ outcome [$F_{3,2294} = 10.004$, $p < .001$] interactions.

Figure 3.5.B depicts response preferences as a function of the angle between two consecutive choices. For example, a high proportion of 0° responses would indicate that stay behaviour was prevalent while a preference for +90° would indicate a

| Angle Moved, N = 2304, $R^2$ = .047 | | | | |
|---|---|---|---|---|
| Variable | $\beta$ | $SE(\beta)$ | $t$ | $p\beta$ |
| Intercept | 1.000 | .009 | 113.735 | <.001 |
| Vertical | -5.587 | 1.450 | -3.854 | <.001 |
| Vertical$^2$ | -1.807 | 1.450 | -1.247 | .213 |
| Vertical$^3$ | -2.565 | 1.239 | -2.071 | .039 |
| Vertical × Choice | 0.410 | .094 | 4.345 | <.001 |
| Vertical$^2$× Choice | 0.198 | .094 | 2.093 | .036 |
| Vertical$^3$× Choice | .023 | .084 | .269 | .788 |
| Vertical × Outcome | 3.118 | .844 | 3.694 | <.001 |
| Vertical$^2$× Outcome | 2.938 | .844 | 3.481 | .001 |
| Vertical$^3$× Outcome | 1.740 | .844 | 2.061 | .039 |

Table 3.4: Results of multiple regression on angle moved between choices. Vertical effects of choice location refer to $\cos\theta$.

tendency to shift 90° clockwise from the previous choice. As seen in Figure 3.5.B, lose-shifting was comprised of several sub-behaviours. Following a loss, participants avoided the previously selected choice. Instead they shifted a small distance towards choices adjacent to the previously selected option. This behaviour is reminiscent of foraging wherein an animal explores its immediate surroundings and has been demonstrated in both rats and pigeons while exploring a radial arm maze Olton et al. [1977], Bond et al. [1981]. Alternatively, participants shifted as far away (180°) from the previous loss as possible; a behaviour we refer to as complete avoidance. In both cases, subjects avoided choices a moderate distance (90°) away from the previous loss.

Following winning outcomes, subjects frequently repeated their previously selected action (i.e., moving 0°). As with losses, they also avoided choices a moderate distance away and shifted towards those 180° from their previous choice. Therefore, the win-stay and lose-shift primarily differ in the choice to repeat the same action or forage around its immediate neighbourhood.

As the number of choices increased from six to eighteen, subjects were more likely to lose-stay, relative to chance responding (i.e., a scaled proportion of 1). Rates of win-stay behaviour also did not change relative to chance. However, because the total number of choices increased, overall win-stay and lose-stay behaviour declined. Instead, subjects were more likely to explore choices immediately adjacent to their previous action, regardless whether they won or lost. They also became more avoidant of choices a moderate distance from their previous action. Finally, rates of complete avoidance (i.e., 180° rotations) did not consistently vary between conditions. Overall, we find that choice valuation following wins and losses are accompanied by unique spatial-tuning functions. As seen in Figure 3.5.B, choice valuation following wins exhibits center-surround suppression, similar to that seen in the visual system. As choice numbers increase, these tuning functions sharpen, driven by greater exploration of nearby choices and avoidance of moderately distant ones.

## 3.4 Discussion & Conclusions

While we often think of the decisions we make as abstract concepts driven by emotion or logic, they are very much rooted in the physical world. In Chapter 2 we demonstrated that the value of choices, the win-stay, and lose-shift responses are largely determined by their location relative to the body. However, the exact relationship between choice location and value remained unknown. Are our preferences for certain locations arbitrary, or do they follow a systematic spatial-tuning function. Moreover, are choice preferences the same regardless of environmental context or does our decision-strategy change depending on the number of choices available? To answer these questions, participants played the game Matching Pennies with six,

twelve, or eighteen choices available. We report here that win-stay and lose-shift behaviour follow a spatial tuning function governed primarily by the distance between choices and the number of choices available. Moreover, as choice numbers increase humans adopt a win-shift/lose-shift decision strategy. We also report that the effects of cannabis use on choice behaviour are sexually dimorphic. Male cannabis users exhibit improved lose-shift suppression while females do not.

Most investigations of decision-making have demonstrated that the win-stay and lose-shift are independent decision strategies reliant on different neural circuits [Chapter 2; Gruber and Thapa, 2016]. However, the anatomical basis for these behaviours remains a subject of debate. The putamen, caudate nuclues, and nucleus accumbens have all been shown to support win-stay and lose-shift responding in different experimental contexts [Packard et al., 1989, Ragozzino, 2007, Clarke et al., 2008, McDonald and White, 2013, Skelin et al., 2014, Gruber et al., 2017]. This discrepancy may be due to the fact that our decision-strategies change as choice numbers increase. For example, the nucleus accumbens is necessary for win-stay responding when two choices are present [Gruber et al., 2017] while the dorsal and dorsomedial striatum are necessary in 8-choice tasks [Packard et al., 1989, McDonald and White, 2013]. Therefore, we investigated how win-stay and lose-shift behaviours change as the number of choice options increases.

We demonstrated that, as choice numbers increase, our independent win-stay and lose-shift strategies give way to a single win-shift/lose-shift strategy. Consequently, the win-stay and lose-shift behaviours become increasingly more correlated across participants. However, adoption of a single, shifting-based strategy does not negatively impact task performance; wins in the eighteen-choice was equal to that when

two choices are present. Instead, participants exhibited sub-optimal performance during the transition from a win-stay/lose-shift to win-shift/lose-shift strategy. This quadratic relationship between demonstrated in previous studies of choice preference. Reutskaja and Hogarth [2009] found that as more choices are available to us, are desire for more options is offset by the additional effort needed to consider more choices.

Striatal activity also exhibits a quadratic relationship with the number of choices available. The putamen, caudate nucleus, and anterior cingulate are maximally active when presented with twelve choices, relative to six or twenty-four [Reutskaja et al., 2018]. These regions are also associated with maintenance of working memory for locations and choices [Akhlaghpour et al., 2016, Cools et al., 2008]. Therefore, the striatum modifies decision strategies in response to set size [Kim et al., 2014], in addition to supporting the win-stay and lose-shift responses.

The temporal properties of the win-stay and lose-shift behaviours provides further evidence that a single win-shift/lose-shift strategy dominates when many choices are present. In the time following losses, the tendency to lose-shift decays over time while stay behaviour following wins does not [Gruber and Thapa, 2016, Ivan et al., 2018]. Therefore, different memory traces supported by separate cognitive systems are responsible for both behaviours. However, as choice numbers increase win-stay responses exhibit temporal decay identical to that of the lose-shift. Consequently, when multiple choices are present a single decision-process is utilized.

Finally, we demonstrated that win-stay and lose-shift responses are determined primarily by the distance between subsequent choices, and that choice preferences exhibit properties similar to spatial tuning functions. For example, following a win subjects were highly likely to repeat their previous choice or shift long distances

to a new region of their environment. However, they avoided locations a moderate distance away from their previous choice. Losses were also followed by large shifts between choice locations. However, subjects were also likely to forage around their immediate environment (i.e., select an immediately adjacent choice). This tendency to "forage" around nearby choices has also been demonstrated in rats, pigeons, and even fish searching for food in a radial arm maze [Olton et al., 1977, Bond et al., 1981, Roitblat et al., 1982]. Consequently, alternation between adjacent locations is a strategy universally applied by many animals while foraging. Moreover, as the number of choices increased, this tendency to forage grew much stronger following both wins and losses. However, animals do not exhibit the tendency for 180° rotations that humans do.

The spatial distribution of lose-shift responses indicate shifting is not comprised of a single behaviour, but two. These different shifting-based strategies also highlight our hypothesis that the win-stay and lose-shift are calculated in egocentric and allocentric coordinates respectively. The win-stay, by its nature, is a relatively simple behaviour. It only requires is the repetition of an action in the same spatial frame of reference. Consequently, circuits necessary for allocentric spatial processing, such as the hippocampus, anterior cingulate, and orbitofrontal cortex [Kelley et al., 1982, Gerfen, 1984, Voorn et al., 2004, De Leonibus et al., 2005] are unnecessary. Conversely, shift responses can be calculated relative to one's body, one's environment, or using a number of strategies. In each situation the resulting motor action needed to perform a shift response will differ. It follows then that lose-shift responding requires requires a more complex, allocentric frame of reference supported by the nucleus accumbens and associative cortex.

We also investigated whether cannabis use was associated with sexually dimorphic changes in win-stay and lose-shift responding. As in Chapter 2, male cannabis users were better able to suppress sensorimotor lose-shift responding, regardless of the number of choices present. However, female cannabis users were not significantly worse at suppressing the lose-shift response relative to controls. Regardless, we have replicated our previous findings that the effects of cannabis use on behaviour are sexually dimorphic.

In sum, we find that we find that the greatest determinant of choice value, and the decision to win-stay or lose-shift, is the position of a choice. These spatial tuning functions may stem from several sub-strategies we use in dealing with losses and wins, such as exploitation of known outcomes, exploration of our immediate neighbourhood, and complete avoidance of losses. These results support previous anatomical findings that the win-stay and lose-shift are supported by the sensorimotor striatum, which in turn receives inputs from motor, somatosensory, and visuospatial circuits. However, in the current task choices were visually identical and only differentiable on the basis of spatial location. Consequently, spatial processing may not drive win-stay and lose-shift behaviour when choices are visually distinct. Therefore, in the next chapter we will explore whether choice position or visual identity are more important in determining choice value.

# Chapter 4

# Spatial Location, Not Visual Identity, Determines Choice Value

## 4.1 Introduction

Adaptive decision-making is governed by the interaction of several brain circuits, each of which has unique aspects that are advantageous under particular circumstances. For instance, a classic distinction has been made between goal-directed control systems, involving the prefrontal cortex and medial striatum, and habitual control systems comprised of the sensorimotor cortex and lateral striatum [Balleine and O'Doherty, 2010, Gruber and McDonald, 2012]. The goal-directed system appears to implement executive functions, such as working memory and strategic planning [Fuster, 1989, Passingham and Wise, 2012]. Here we report a new dissociation among executive and sensorimotor systems governing choice, which allows us to quantify their interaction while accounting for important confounding factors such as decision time and learning.

When rewards are uncertain, the most pervasive strategy in animals and humans is to repeat choices that have previously led to reward (win-stay), and to shift away from choices following reward omission [lose-shift; see Thorndike, 1911, Kamil and Hunter, 1970, Worthy et al., 2013]. Although the win-stay and lose-shift are complementary behaviours, they are anatomically disassociated among goal-directed and sensorimotor systems. Lesions to the rodent lateral striatum (LS), which is homologous to the human putamen and essential for sensorimotor control [Parent and Hazrati, 1995], disrupt lose-shift responding but not win-stay behaviour [Skelin et al., 2014, Gruber et al., 2017, Thapa and Gruber, 2018]. A similar shifting deficit has been observed in humans with damage to putamen or insula [Danckert et al., 2011]. Conversely, lesions of the rodent ventromedial striatum (VS), a key structure in goal-directed control that recieves inputs from prefrontal cortex [Voorn et al., 2004], disrupts win-stay but not lose-shift responding [Gruber et al., 2017]. Several other behavioural features in rodents and humans support this anatomical disassociation. The win-stay and lose-shift exhibit different temporal dynamics [Gruber and Thapa, 2016], developmental trajectories [Ivan et al., 2018], and responses to reward feedback [Banks et al., 2018]. Moreover, lose-shift responding (but not win-stay) drastically increases in adult humans under cognitive load, and in young children [Ivan et al., 2018]. These data suggest that executive function can override lose-shift responding, which can be characterized as a reflexive response by the sensorimotor striatum. This hypothesis is consistent with a long history of research indicating that executive function can suppress reflexive or habitual motor responses [Chamberlain and Sahakian, 2007].

The LS/putamen receives prominent inputs from both the somatosensory and motor cortices [Brasted et al., 1999], and encodes the motor aspects of decision-making

[Burton et al., 2015]. Consequently, decisions and their associated motor actions are represented in egocentric (body-centred) spatial coordinates [Kesner and DiMattia, 1987, Palencia and Ragozzino, 2005]. The dorsolateral caudate, which receives inputs from the dorsolateral PFC, is also necessary for egocentric spatial processing [Possin et al., 2017]. The VS/nucleus accumbens encodes the value of choices and can engage a wide range of spatial-motor actions when executing a single decision involving abstract representations [Burton et al., 2015, Mashhoori et al., 2018]. It encodes responses in both egocentric and allocentric (world-centred) spatial coordinates, likely involving its prominent inputs from the hippocampus and prefrontal cortex (PFC) [De Leonibus et al., 2005, Voorn et al., 2004, Possin et al., 2017]. These data suggest that the control of actions by sensorimotor systems will be restricted to an egocentric framework heavily dependent on the position of targets, whereas the control by executive systems have the capacity to use abstract features of targets. This is supported by the dissociated effects of cannabis on neural structures and performance on spatial versus non-spatial tasks, as described next.

The recreational use of psychoactive substances has complex short-term and long-term effects on the brain, some of which dissociate. $\Delta^9$- tetrahydrocannabinol (THC) administration increases dopamine release in the LS, while the VS remains unaffected [Sakurai-Yamashita et al., 1989]. Behaviourally, dopamine signalling in the dorsolateral striatum is necessary for normal spatial memory, motor control, and visospatial learning, while reward processing and goal-directed learning rely on dopamine signalling in the medial striatum [Darvas and Palmiter, 2009, 2010]. This provides an explanation for the observation that THC reduces spatial processing and visuospatial memory in humans [Cha et al., 2007], particularly in females [Pope et al., 1997].

Recreational drugs also differentially influence win-stay and lose-shift responding. THC and amphetamine cause large changes in lose-shift behaviour in rats and humans, while the win stay is only weakly affected [Wong et al., 2017b,a, Paulus et al., 2002b].

Because the LS is necessary for lose-shift responding, and it processes information in egocentric coordinates, we hypothesized that lose-shift responses are calculated according to the position of a target relative to the participant, rather than other visual features of target identity. We further hypothesized that frequent cannabis use will disrupt the normal positional-dependence of lose-shift responses and the ability of executive systems to govern sensorimotor control. We tested these hypotheses while human participants were engaged in a competitive decision-making task between two choices. Crucially, the choices were visually distinct and changed their spatial configuration unpredictably between each trial. We found that following a loss, lose-shift behaviour was robustly associated with a choice's previous location, rather than its visual identity. The win-stay was only weakly associated with previous choice position, and this association was eliminated by global changes in target position. Although female cannabis users exhibited reduced task performance and increased lose-shift responding, their reliance on spatial information did not differ from controls. Male cannabis users, however, did exhibit a reduced reliance on spatial information. These data support the dissociation of choice among systems with different spatial propensities, and reveal a sexual dimorphism of recreational cannabis use and the function of these systems.

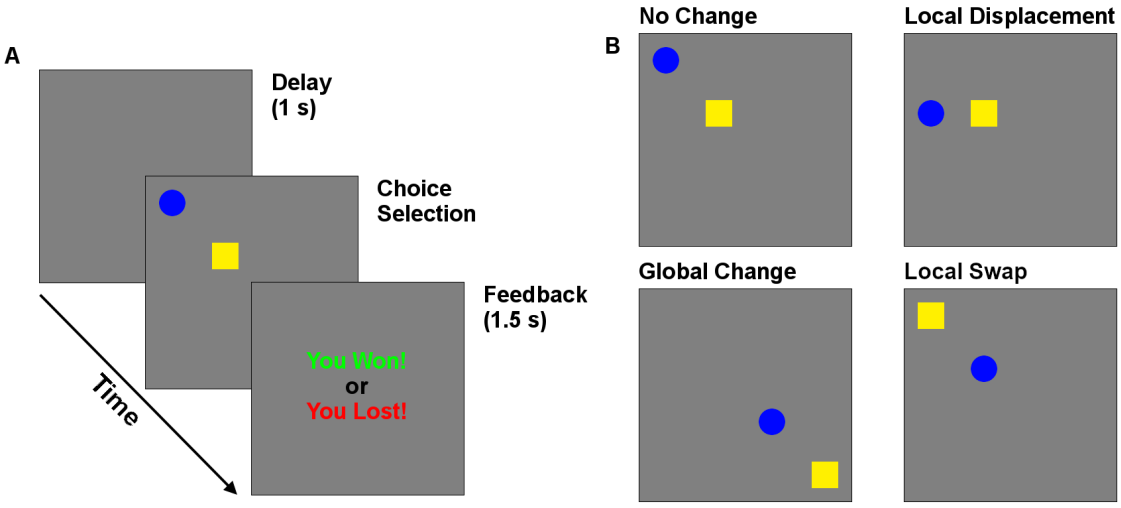


Figure 4.1: Behavioural task. A: timeline of trials in the matching pennies game. B: possible reconfiguration of targets between trial, which could undergo both local (swap & displace) and global changes in position.

## 4.2 Methods

### 4.2.1 Behavioural Task

During the experiment, participants played a competitive game colloquially called "Matching Pennies" against a computer opponent. The task display consisted of two distinct targets (a blue circle and yellow square) presented on a 15″ touchscreen monitor (Figure 4.1). On each trial, participants could choose one target by touching it on the screen. They would then receive visual feedback indicating "You Win" or "You Lose" for 1.5 s. On each trial, the computer algorithm attempted to predict which target would be selected. If the participant selected this target, the trial was a loss. Otherwise it was a win. The algorithm attempted to minimize the number of wins for participants. The optimal strategy for the participants is to be unpredictable

in choice, in which case they win on 50% of trials. Because the win-stay and lose-shift are predictable, subjects should learn to suppress these responses as the session progresses. This task provides measures of lose-shift, win-stay, and cognitive flexibility as participants adapted their choices to the computer opponent.

The computer used four types of algorithms to detect patterns in (*i*) participants' choices, (*ii*) switching from one choice to another, (*iii*) choices paired with rewards (e.g., blue square after a loss), and (*iv*) switching paired with rewards (e.g., swapping choices after a loss). Specifically, on each trial the computer examined a subject's recent choice and reward history (e.g., shifting from the blue target to the yellow after a loss). The choice that most accurately predicted the subject's past choice history was selected as the prediction of the present choice. Patterns of choices 1-6 trials in length were considered, resulting in 24 total prediction strategies. On each trial, the best performing strategy (computed over all previous trials in the session) was used to predict participants' choices. If all strategies failed to beat the participant on $\geq$ 50% of past trials, the computer would select choices randomly.

The effect of cue position was investigated by moving the location of one or both cues from one trial to the next. The changes came in two types - local and global. The screen was divided into four equal quadrants, each of which contained an invisible 2×2 grid in its center. Local changes occurred within each grid, while global changes involved shifting among quadrants. Three local manipulations are of particular interest to investigate the importance of position and cue identity. The *Control* condition refers to cue positions remaining stationary between trials. The *Swap* condition occurs when the targets swap positions (a local change). Finally, the *Displace* condition occurs when the previously selected choice moves to a previously empty position in

the same 2×2 choice grid, while the other target remains in its previous position. Global changes occurred independently of each of these 3 local changes, for a total of 6 possible changes in target positions across subsequent trials, which were selected randomly on each trial. This manipulation allowed us to determine how the position of targets relative to each other, and to the participants, affected choice following wins and losses. In particular, this design allowed us to test if participants avoid the screen position of a target following a loss, as expected by the egocentric processing framework of LS or instead avoid the target regardless of position.

### 4.2.2 Procedure

All procedures and experimental tasks were approved by the McMaster University Research Ethics Board. One hundred six undergraduates (53 males, mean age = 19.40, $SD = 2.74$) from McMaster University participated in the study in exchange for payment. After providing informed consent, participants played 600 trials of the task. They were informed that they would win nothing each time the computer predicted their choice and 3¢ each time it could not, and that their total winnings would be rounded up to the nearest $5 upon completion of the experiment. Participants were given no guidance as to optimal decision-making strategies.

After completing the Matching Pennies task, participants completed the South Oaks Gambling Screen, the alcohol, smoking and substance involvement screening test (ASSIST) v3.0, Adult ADHD Self-Report Scale v1.1, and an additional demographic questionnaire. Habitual cannabis users were defined as those meeting the criteria for brief or intensive treatment (i.e., a score $> 3$) on the ASSIST cannabis subtest. Total drug use was also recorded as the ASSIST score summed across all drug subtypes.

Males had a mean ASSIST score of 19.11, with 32.08, 39.62, and 60.38% meeting the criteria for alcohol, cannabis, or any recreational drug use requiring intervention. Females had a mean ASSIST of 16.66, with 28.30, 30.19, and 45.28% meeting alcohol, cannabis, or general drug use criteria (See Table 4.1).

| Sex | Group | N | Age | Cannabis | ASSIST |
|---|---|---|---|---|---|
| Female | Control | 37 | 19.2±2.0 | 0.3±.9 | 4.1±6.2 |
| | Cannabis | 16 | 18.9±1.1 | 12.8±8.6 | 40.0±25.4 |
| Male | Control | 32 | 19.3±2.6 | 0.7±1.2 | 8.6±8.4 |
| | Cannabis | 21 | 20.29±4.3 | 12.8±8.0 | 35.1±21.3 |

Table 4.1: Demographic data and ASSIST questionnaire scores (± SEM) for the studied groups.

### 4.2.3 Analysis

Participants' responses were analyzed for the proportions of lose-shift and win-stay responses averaged over five, one hundred twenty trial blocks, and conditioned on the type of cue shift relative to the previous trial's target positions. As a measure of behavioural flexibility, the binary response entropy ($H$) for each participant was calculated from 4-trial choice sequences as:

$$H = \sum_{i=1}^{k} P_i + log_2 P_i \tag{4.1}$$

where $P_i$ is the probability of each choice sequence, and $k$ is the total number of sequences possible (i.e., 16). For example, a participant that exhibited the choice pattern "circle-square-circle-square" to the exclusion of all other patterns, would have an entropy of 0 bits, while a participant responding randomly would have an entropy of 4 bits. Response entropy and task performance were averaged over the experimental

session for each participant. Decision times were measured as the time to make a response following presentation of the choice selection screen. They were normalized using the inverse transform (1/RT) and averaged after removing 131 erroneous RTs of <3 ms. The inverse transform was used to normalize RTs because it produced more normalized (Gaussian) distributions than did the log or square-root transforms.

The differences among marginal means of derived quantities (decision time, lose-shift, etc.) were tested by analysis of variance (for categorical factors) or co-variance (for continuous factors) using repeated-measure, mixed-effects models. Each model utilized a maximal random-effects structure and was fit in R using the lme4 package [Bates et al., 2014]. A maximal model ensured that variations in effects between participants, and between trial blocks within each participant, were properly controlled [Barr et al., 2013]. Degrees of freedom and $p$-values were calculated using the Welch-Satterthwaite equation and Type-III sums of squares. The effects of local changes in position were assessed with planned paired t-tests comparing the effects of spatial swaps & displacement relative to the no-change condition.

We also used the *Q-learning with forgetting* [Barraclough et al., 2004] reinforcement learning model to examine the effects of cannabis use, local changes, and global changes on reward sensitivity, choice stochasticity, and learning rates. In this model the probability of selecting one of the two choices ($C$) on a given trial ($t$) is calculated with the softmax equation [Sutton and Barto, 2018]:

$$P(C_t = i|Q_i, Q_j) = \frac{\exp(\beta \times Q_{i(t)})}{\exp(\beta \times Q_{i(t)}) + \exp(\beta \times Q_{j(t)})}, \tag{4.2}$$

where $Q_i$ and $Q_j$ are the value each subject assigns to choices $i$ and $j$. $\beta$ refers to the

inverse temperature that balances the opposing tendencies to exploit known action-reward associations versus exploring more of the state/action space. As such, larger values of $\beta$ indicate a greater tendency to choose the most highly valued action. The values of each choice are updated from rewards ($R$) according to the following rules:

$$Q_i(t) = \begin{cases} Q_{i(t-1)} \times (1-\alpha) + \alpha\kappa_1, & \text{if } C_{t-1} = i, R_{t-1} = 1 \\ Q_{i(t-1)} \times (1-\alpha) - \alpha\kappa_2, & \text{if } C_{t-1} = i, R_{t-1} = 0 \\ Q_{i(t-1)} \times (1-\alpha), & \text{if } C_{t-1} \neq i \end{cases} \tag{4.3}$$

where $\alpha$ is the learning and forgetting rates for the chosen and unchosen action, $\kappa_1$ is the strength of reinforcement from reward, and $\kappa_2$ is the strength of aversion from failing to receive a reward. These three parameters were treated as stochastic variables that follow a random walk process. As such, they were free to vary throughout the experiment. Conversely, $\beta$ was treated as a deterministic variable that remained fixed throughout the experiment. These parameters were fit for each subject using the VBA toolbox [Daunizeau et al., 2014].

To determine how local swaps, displacement, and global changes influenced RL parameters (i.e., hidden state values), we performed a Volterra decomposition of $\alpha$, $\kappa_1$, and $\kappa_2$ values for each trial onto previous choices, outcomes, local displacements, swaps, and global changes (relative to no change). The Volterra weights of these five basis functions ($u$) were calculated according to Eq. 4.4:

$$x_t = \omega^0 + \sum_{\tau} \omega^1_{\tau} u_{t-\tau} + \sum_{\tau_1} \sum_{\tau_2} \omega^2_{\tau_1,\tau_2} u_{t-\tau_1} u_{t-\tau_2} + ... \tag{4.4}$$

Volterra modelling allows for observation of input response characteristics of non-linear systems as Volterra weights [Boyd et al., 1984]. At each trial $t$ the Volterra weight $x$ of a given parameter is estimated from inputs $u$ over trials $t$ to a lag of $\tau$ (set to 32 trials) using a series of Volterra kernels $\omega$. The first kernel $\omega^1$ represents the linear transformation of lagged input basis functions into the output, $\omega^2$ represents the effect of past inputs being dependent on other earlier inputs, and so on. These weights provide a measure of how subjects' valuation of each choice changes from baseline in response to past choices and outcomes. The benefit of Volterra modelling over analysis of raw prediction error is that the effect of current and past inputs on hidden state responses can be estimated. Inputs were also orthogonalized so that the effect of one input (e.g., local swaps) is computed independently of all other inputs (e.g., global changes). To control for trial order effects, we also detrended inputs prior to decomposition using a cubic polynomial.

## 4.3 Results

### 4.3.1 Relationship between choice behaviour & cannabis use

Each of the 106 included participants performed 600 trials of the task, for a total of 63,600 trials in the dataset. We first sought to reveal how recreational cannabis use and biological sex affected overall performance on the task. We compared the effects of sex (male, female) and habitual cannabis use on the proportion of wins with a 2×2 factorial ANOVA. The results are shown in Figure 4.2.A. The ANOVA revealed a significant main effect of cannabis use [$F_{1,102} = 4.772$, $p = .032$] and a significant sex × cannabis use interaction [$F_{1,102} = 6.540$, $p = .012$]. The main effect of sex was not

significant [$F_{1,102} = 1.226$, $p = .271$]. The interaction was significant because cannabis use was associated with decreased task performance in females [$t(51)$ = -3.123, $p$ = .003, $d$ = -.934][A,1] but not in males [$t(51)$ = .285, $p$ = .777, $d$ = .080].

A similar interaction between sex and cannabis use was present for response entropy and decision times (Figs. 4.2.B & C). A 2 × 2 ANOVA on response entropy (Fig. 4.2.B) found that the main effects of sex [$F_{1,102} = 3.043$, $p = .084$], cannabis use [$F_{1,102} = 2.949$, $p$ = .089], and the sex × cannabis use interaction [$F_{1,102} = 3.762$, $p$ = .055] on response entropy fell just short of significance. However, female cannabis users did exhibit significantly lower response entropy relative to controls [$t(51)$= -2.201, $p$ = .032, $d$ = -.658][B] whereas males did not [$t(51)$= .194, $p$ = .847, $d$ = .054]. A 2 × 2 ANOVA on decision times found non-significant main effects of sex and cannabis use ($p > .112$ in both cases), but a significant sex × cannabis use interaction [$F_{1,102} = 7.701$, $p = .007$]. Cannabis use was associated with decreased decision times in females [$t(51)$ = -3.024, $p$ = .004, $d$ = -.905][C], while those of men were again unaffected ($p = .399$).

As expected, task performance was positively correlated with response entropy [$r(104) = .699$, $p < .001$], negatively correlated with mean lose-shift tendencies [$r(104)$ = -.605, $p < .001$], and not correlated with win-stay responding [$r$ (104) = -.095, $p$ = .333]. Therefore, frequent cannabis use in females was associated with increased lose-shift responding and decreased response times. These behavioural features are consistent with the hypothesis that decisions were strongly influenced by sensorimotor control of the decision process. Moreover, the tendency for increasingly stereotyped response sequences in females that frequently used cannabis further suggests a reduction in cognitive flexibility, defined here as a reduced ability to generate varied

[1] additional details on effect sizes are provided in Table 4.3

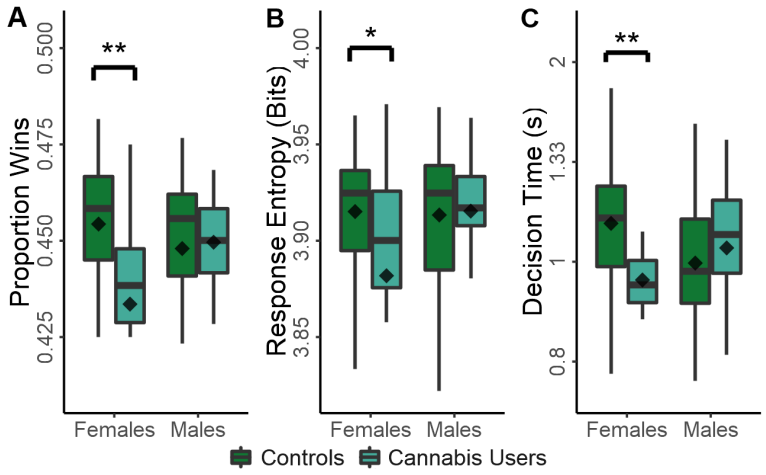


Figure 4.2: Effect of recreational drug use use on measures of task performance in males and females: proportion of wins (A); response entropy (B); and decision times (C). Plots show the conventional descriptive statistics: mean (diamond), median (horizontal line in the box), 25th/75th percentiles (box edges), and outliers (dots). Note: $*p < .05, **p < .01$,

responses following losses that are needed for optimal task performance.

### 4.3.2 Spatial cues drive lose-shift & win-stay responding

The optimal strategy in this task was to simply select targets at random on each trial. Deviation from this optimal strategy reveals the neural processes guiding choice behaviour. For example, lose-shift responding is maladaptive in this context but nonetheless is a prevalent strategy. We next investigated to what extent spatial and/or other visual features of targets affect the propensity for lose-shift and win-stay behaviours. The task design allows us to test if participants avoid the screen position of a target following a loss, as expected by the egocentric processing framework of LS, or if they avoid the target itself regardless of position. We compared the effects of local and global changes in target position via 2 (global change, no change) $\times$ 3

(no local change, displace, swap) mixed-effects models with repeated measures and a full random effect structure (see Methods). We then conducted planned comparisons of marginal means for significant effects revealed by ANOVA.

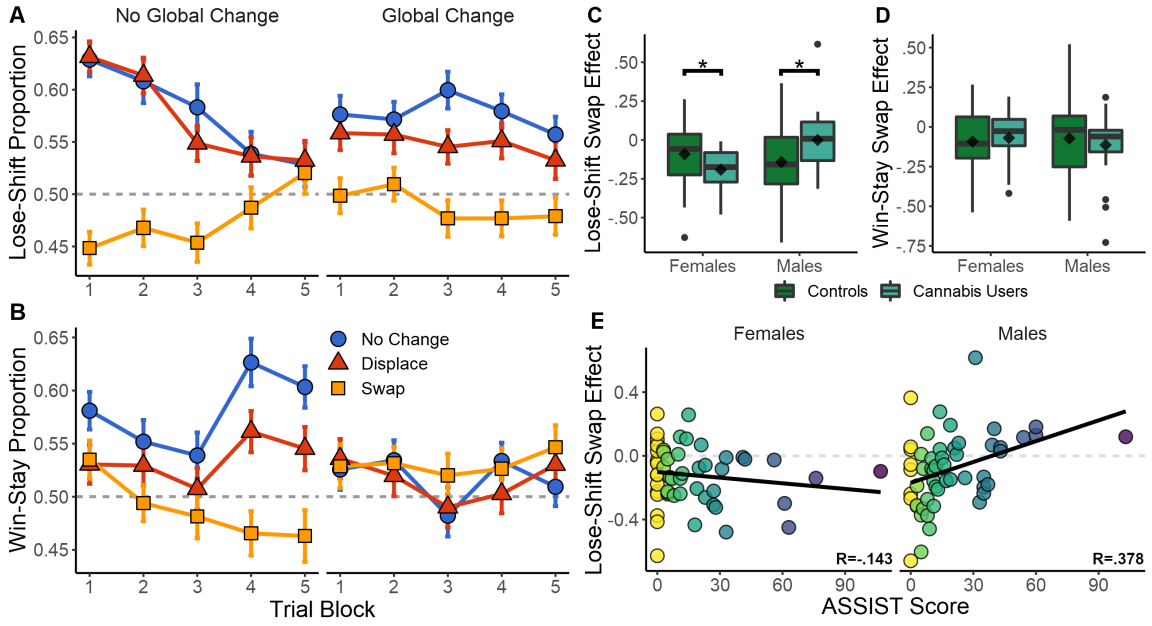


Figure 4.3: Effects of target reconfiguration and recreational drug use on reinforcement-driven behaviour. A-B: effect of local & global changes in choice position on lose-shift and win-stay tendencies for all participants. SEM in error bars. C-D: box plots of the difference in lose-shift and win stay when target positions are swapped as compared to no change. The effect of swapping targets on lose-shift is higher in women who use cannabis, but lower in men who use cannabis, than their sex-matched controls. E: correlation between total drug use (ASSIST) and swap effect on lose-shift.

Lose-shift behaviour was strongly affected by local changes in position across the 600 trials of the session parsed into 5 blocks [$F_{2,117.43} = 25.643$, $p < .001$]. The effects of global changes [$F_{1,320.75} = .368$, $p = .545$] and the local $\times$ global interaction [$F_{2,147.87} = 2.285$, $p = .105$] were not significant. As seen in the left panel of Figure 4.3.A, participants exhibited a high degree of lose-shift responding when choices did not move between trials, particularly in the first 3 blocks (360 trials). However,

swapping choice positions strongly reversed their associated lose-shift probabilities [$t(529)$ = -7.249, $p < .001$, $d$ = -.507]$^D$, particularly in the first 3 blocks. Because we are computing shifts with respect to each target (rather than position), a lose-shift probability less than 0.5 on swap trials indicates that participants are selecting the same target, but in a new location. This is a lose-stay response in terms of target identity, but a lose-shift in terms of spatial position. In other words, the blue and orange lines should overlap if lose-shift is computed with respect to target identity irrespective of location. Therefore, lose-shift is based on the previous position of an unrewarded target, rather than its identity as distinguished by other features (color and shape).

Although participants are able to eventually suppress lose-shift responding after hundreds of trials, this occurs only in the absence of global changes in target positions (cf. left and right panels of 4.3.A). The effects of local swaps persisted in the presence of global changes [$t(529)$ = -8.056, $p < .001$, $d$ = -.459]$^E$. Furthermore, the effect of local displacement was only significant in the presence of global changes [$t(529)$ = -2.827, $p$ = .005, $d$ = -.160]$^F$, and reduced lose-shift responses relative to global changes and no displacement. Global changes thus appear to immediately reduce the probability of lose-shifting early in the session, but also interfere learning to suppress this sub-optimal response near the session end.

The same analysis was repeated for win-stay responses (Fig. 4.3.B). A 2 × 3 ANOVA revealed a significant main effect of local changes [$F_{2,105.65} = 5.470$, $p = .005$] and a significant local × global interaction [$F_{2,159.33} = 11.070$, $p < .001$] . The main effect of global changes was not significant [$F_{1,1117.7} = 3.792$, $p = .052$]. Across the

entire session, both local swaps [$t(529)$ = -6.565, $p < .001$, $d$ = -.438]$^G$ and displacement [$t(529)$ = -4.388, $p < .001$, $d$ = -.221]$^H$ reduced win-stay responding when no global change was present. Unlike lose-shift responses, win-stay behaviour was initially unaffected by local changes. However, as trial blocks progressed, both win-stay behaviour and the effects of local changes increased. Global changes completely eliminated any effects of local changes throughout the session. The most parsimonious explanation is that subjects eventually learn to suppress the shift response, which reveals the stay response as the session progresses. To test if the lose-shift and win-stay are in competition, we computed their correlation using data from the no-change condition separated into 120-trial blocks. Initially they were uncorrelated [$r(104)$ = .008, $p$ = .935]; however, as trials progressed the win-stay and lose-shift exhibited an increasingly negative correlation [block 3: $r(104)$ = -.310, $p$ = .001], [block 5: $r(104)$ = -.406, $p < .001$]. Therefore, behaviour is initially biased toward shift responses but becomes increasingly biased toward stay responses as the session progresses.

Because the time between reinforcement and subsequent decisions affect lose-shifting [Gruber and Thapa, 2016, Ivan et al., 2018], we next analyzed whether changes in these response types could be explained by effects of target manipulation on decision times. Analysis of inverse-transformed decision times found significant effects of local [$F_{2,140.14}$ = 9.668, $p < .001$] and global [$F_{1,109.81}$ = 29.731, $p < .001$] changes in position, and a local × global interaction [$F_{2,140.15}$ = 9.672, $p < .001$]. Table 4.2 provides decision times for each response type following local and global changes. Regardless of response type, global changes significantly increased decision times. More importantly, local swaps increased the time of lose-shift, win-shift, and win-stay responses, particularly when no global changes were present. Although it

could be argued that this increase is due only to the extra time needed to move locations, the fact that these effects are not consistent between response types indicates otherwise. Instead, this finding suggests that actions are planned prior to target presentation, and must be updated when target positions change to unexpected locations. The difference in decision time when targets are moved is on the order of 0.1 s, which is too small to account for changes in lose-shift or win-stay responding as a decay in memory of the previous reinforcement [Ivan et al., 2018].

| | No Global Change | | Global Change | |
|---|---|---|---|---|
| Local | Mean±SE | $t(105)$ | Mean±SE | $t(105)$ |
| | | **Lose-Shift** | | |
| No Change | 0.955±.015 | - | 1.024±.017 | - |
| Displace | 0.975±.017 | 2.333* | 1.031±.018 | .936 |
| Swap | 1.025±.017 | 5.854‡ | 1.042±.019 | 1.897 |
| | | **Lose-Stay** | | |
| No Change | .980±.019 | - | 1.070±.020 | - |
| Displace | .995±.019 | 1.171 | 1.047±.018 | -1.691 |
| Swap | .987±.017 | .576 | 1.042±.019 | -2.372* |
| | | **Win-Shift** | | |
| No Change | 1.023±.020 | - | 1.084±.021 | - |
| Displace | 1.044±.021 | 1.907 | 1.081±.020 | -.287 |
| Swap | 1.065±.021 | 3.227† | 1.098±.021 | 1.447 |
| | | **Win-Stay** | | |
| No Change | 1.024±.018 | - | 1.056±.020 | - |
| Displace | 1.027±.020 | .329 | 1.064±.020 | .778 |
| Swap | 1.064±.023 | 3.644‡ | 1.058±.021 | .262 |

Table 4.2: Mean decision times for lose-shift, lose-stay, win-shift, & win-stay responses. The $t$-statistic value is reported for the paired comparison between the local position changes and no change condition. Note: $*p < .05, \dagger p < .01, \ddagger p < .001$ Standard errors ($\tilde{\sigma}X$) estimated from inverse-transformed RTs ($\tilde{Y}$) via $\tilde{\sigma}X = (1/\tilde{Y})^2 \times \tilde{\sigma}Y$.

### 4.3.3 Cannabis use modulates the lose-shift

We next analyzed the effects of cannabis use on lose-shift with a mixed-effects model testing the effects of sex (male, female), local changes (no change, displace, swap), and cannabis use (controls, habitual users). Models were fit separately to trials with and without global changes in position, in order to simplify model interpretation. A random-intercepts-only structure was used because the full random-effects structure resulted in an over-fit model.

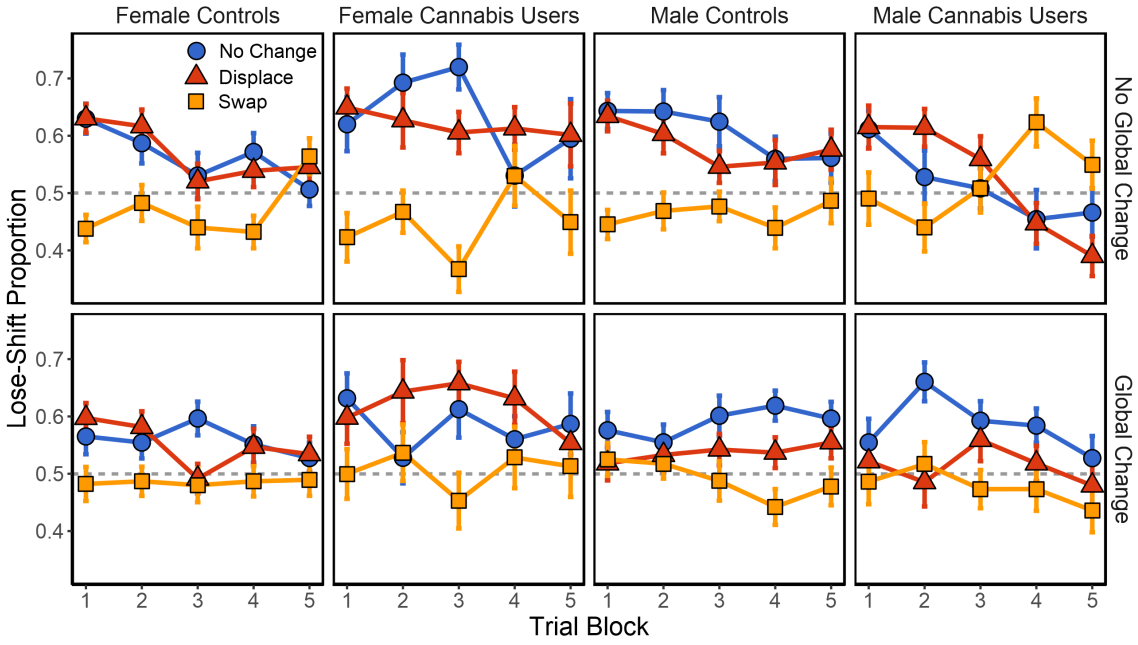


Figure 4.4: Effect of local & global changes in choice position on lose-shift tendencies in male & female cannabis users, relative to controls. SEM in error bars.

Following no global change, there was again a significant main effect of local changes [$F_{2,1476} = 43.347$, $p < .001$] on lose-shift behaviour. Significant local $\times$ sex [$F_{2,1476} = 5.711$, $p = .003$], sex $\times$ cannabis [$F_{1,102} = 8.342$, $p = .005$], and local $\times$ sex $\times$ cannabis [$F_{2,1476} = 13.008$, $p < .001$] interactions also were present. No other effects or interactions were significant ($p > .148$ in each case). As shown in Figure

4.4, male and female controls exhibit similar effects of positional changes on lose-shift behaviour. Specifically, both exhibited a strong-lose shift tendency that declined following local swaps and as trials progressed. In females, the difference between the no change and swap conditions increases with heavy cannabis use [$t(263) = 2.157$, $p = .032$, $d = .289$][I]. This effect on lose-shift behaviour may be due to an increased reliance on spatial choice cues or an increase in baseline lose-shift behaviour. We found that although female cannabis users lose-shift more in the no change condition [$t(263) = 2.402$, $p = .017$, $d = .321$], lose-shifting following local swaps did not differ from controls [$t(263) = -.951$, $p = .343$, $d = -.127$]. In other words, female cannabis users were no more reliant on spatial choice cues than female controls. Instead, female cannabis users exhibited elevated lose-shift responding at baseline (Fig. 4.3.C). Male cannabis users displayed the opposite behaviours. For example, although male controls show a large effect of spatial swaps, this behaviour was extinguished, and in some cases reversed with elevated drug use [$t(263) = 3.737$, $p < .001$, $d = .469$][J]. Importantly, relative to control subjects, male cannabis users lose-shifted less in the no-change condition [$t(263) = -3.489$, $p < .001$, $d = -.438$] and lose-shifted more following local swaps [$t(263) = 2.494$, $p = .013$, $d = .313$]. Therefore, male cannabis users were less reliant on spatial cues when responding after losses.

Modulation of lose-shift behaviour may not be specific to cannabis alone. As seen in Figure 4.3.E, Total drug use (as indexed by the ASSIST score) is positively correlated with the lose-shift swap effect in men [$r(51) = .378$, $p = .005$]. However, in women the two measures are not significantly correlated [$r(51) = -.143$, $p = .308$], suggesting that cannabis use provides a more informative metric. Furthermore, total drug use is more strongly correlated with cannabis use [$r(104) = .847$, $p < .001$]

than tobacco use [$r(104) = .762$, $p < .001$] or alcohol use [$r(104) = .651$, $p < .001$]. Therefore, in our population cannabis use is most strongly associated with total drug use, while also remaining a clinically relevant classification.

Following global changes in position, there were significant effects of local changes [$F_{2,1476} = 34.300$, $p < .001$] on lose-shift behaviour as well as significant local $\times$ sex [$F_{2,1476} = 5.633$, $p = .004$] and sex $\times$ cannabis [$F_{1,102} = 5.764$, $p = .018$], interactions. No other effects were significant ($p > .088$ in all cases).

Similar models were applied to win-stay behaviour. Although the effects of local changes remained significant [$F_{2,1578} = 22.755$, $p < .001$], the effect of sex was not significant [$F_{1,1578} = 3.477$, $p = .062$], nor were the effects of drug use, sex, or the interactions with local changes in position ($p > .219$ in all cases). As seen in Figure 4.3.D, processing of the win-stay did not differ with sex or drug use, and was not considered further. The same was true following global changes in position, where no effects were significant ($p > .135$ in all cases).

### 4.3.4 Computational Results

The results presented above demonstrate that target location is more important than target identity for choice adaptation based on reinforcement in the immediately previous trial. The importance of spatial configuration is evidenced by changes in win-stay and lose-shift probabilities following manipulations of cue position. We next sought to determine how choices, cue configurations, and reinforcement affected choice over multiple trials. We therefore used a biologically-relevant computational model to determine how learning rate, reward valuation, and loss aversion affected choice. Each participant's choice behaviour was analyzed using the *Q-learning with forgetting* (FQ)

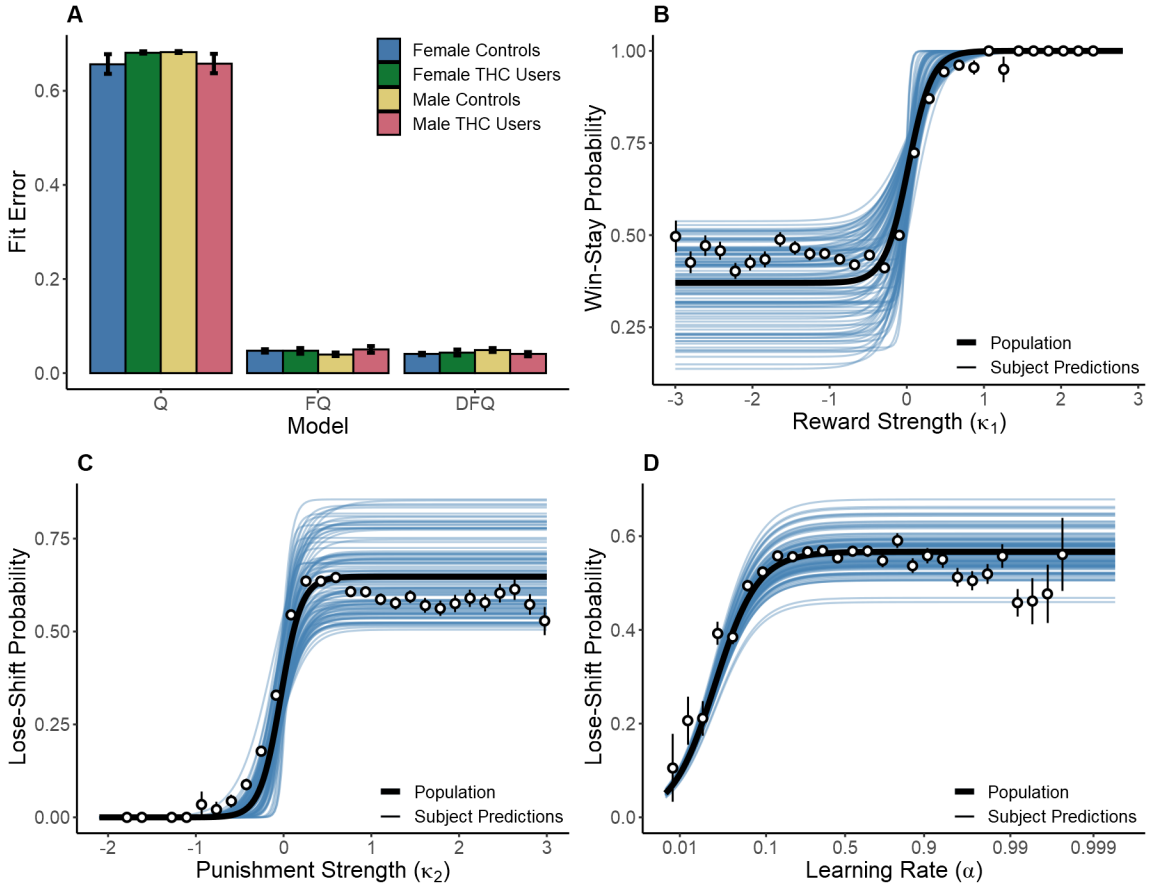


Figure 4.5: A: performance of the *Q-learning* (Q), *Q-learning with forgetting* (FQ), and *Q-learning with differential forgetting* (DFQ) models. B: relationship between $\kappa_1$ and win-stay behaviour with curves fit to individual subjects (blue lines), the population average (black line), and averages of binned raw data (points). C-D: relationship between $\kappa_2$ & $\alpha$ and lose-shift behaviour. Note: $\alpha$ was normalized via the logit transform prior to model fitting.

model [Barraclough et al., 2004], which uses learning rate ($\alpha$), inverse temperature ($\beta$), reward strength ($\kappa_1$), and punishment strength ($\kappa_2$) as parameters (hidden states) to estimate action values. We compared model performance against two other models: *Q-learning* [i.e., Q; Sutton and Barto, 2018] and *Q-learning with differential forgetting* [i.e., DFQ; Ito and Doya, 2009]. The Q-model only includes the $\alpha$ and $\beta$ parameters,

while DFQ includes a second $\alpha_2$ parameter to describe forgetting as a different process from learning. Hidden states were estimated for each subject, using the negative log-likelihood as a metric of model performance:

$$\text{Negative log-likelihood} = -\frac{1}{n} \times \sum_{i=1}^{n} log(P(i)) \tag{4.5}$$

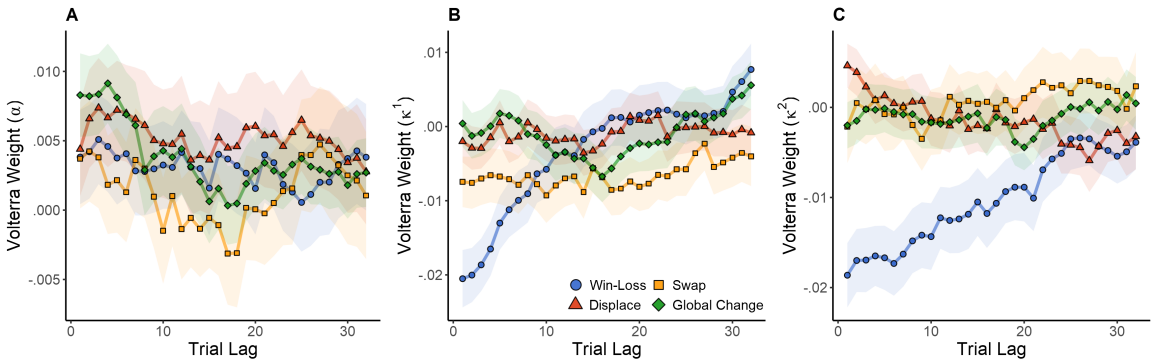


Figure 4.6: Effect of wins, local, & global changes in choice position on Q-learning parameters $\alpha$ (learning rate), $\kappa_1$ (reward strength), & $\kappa_2$ (punishment strength). The influence of wins and cue rearrangement during the previous 32 trials is estimated by Volterra decomposition, which provides a weight (loading) for each trial lag. SEM in shaded area.

where $n$ is the number of trials and $P(i)$ the probability that the model predicted each subject's choice made on trial $i$. As seen in Figure 4.5.A, both the FQ and DFQ models performed better than the Q-learning model. However, the FQ model fit no worse than DFQ ($p = .503$) while requiring one less parameter. Therefore, *Q-learning with forgetting* provided the best model of human choice in the present task. In no instance did model fit vary significantly as a result of sex or cannabis use ($p >$ .154 in all cases), indicating that comparison of the parameter values is well founded. Note that the parameters were free to vary within the session, and so take a range of values for each subject. Figures 4.5.B-D show a logistic-like relationship between

parameter values and win-stay/lose-shift response probability. We fit win-stay and lose-shift responses against $\kappa_1$, $\kappa_2$, and $\alpha$ using mixed-effects logistic regression with a logit-link function. The random effects of logistic asymptote, threshold, and slope were included to control for variability between participants.

We found a significant relationship between reward strength ($\kappa_1$) and win-stay behaviour. Consequently, the asymptote [$\beta$ = .370, $F_{1,30691}$ = 3069.663, $p < .001$], intercept [$\beta$ = .019, $F_{1,30691}$ = 10.986, $p$ = .001], and slope [$\beta$ = -.153, $F_{1,30691}$ = 237.979, $p < .001$] parameters were significant. As seen in Figure 4.5.B, when reward strength is low, subjects win-stay with a fixed baseline of 37.0% ($SD$ = 10.9%). However, when $\kappa_1$ is high ($> .019$), subjects almost exclusively win-stay.

Similar model results were obtained for lose-shift responses. At low values of $\kappa_2$ subjects exhibit a lose-stay policy. However, as $\kappa_2$ increases, they reach a stable lose-shift response rate of 64.8% (SD = 10.1%). Consequently, lose-shift asymptote [$\beta$ = .648, $F_{1,32587}$ = 3712.685, $p < .001$], intercept [$\beta$ = -.033, $F_{1,32587}$ = 20.921, $p <$ .001], and slope [$\beta$ = .118, $F_{1,32587}$ = 254.602, $p < .001$] were significant. Lose-shift behaviour was also associated with learning rates ($\alpha$, Fig. 4.5.D).

A mixed-effects model fitting lose-shift responses to learning rates ($\alpha$) indicated the asymptote in lose-shift responses [$\beta$ = .567, $F_{1,32587}$ = 14089.972, $p < .001$], intercept [$\beta$ = .026, $F_{1,32587}$ = 3641.362, $p < .001$], and slope [$\beta$ = .595, $F_{1,32587}$ = 122.090, $p < .001$] were significant. As subjects increase the rate at which new reinforcement updates past knowledge of choice-outcome associations, they lose-shift more before reaching an asymptote of 56.7% ($SD = 4.8\%$). However, a similar analysis of win-stay responding indicated it did not significantly vary with learning rates (p $> .128$ in all cases).

Given the relevance of the FQ model to human behaviour, we next sought to quantify how hidden states changed in response to reinforcement and cue positions using Volterra decomposition. The method accounted for the past effects of wins, local displacement, swaps, and global changes on changes in $\alpha$, $\kappa_1$, and $\kappa_2$ over the preceding $n \in (1, 32)$ trials. The effects of wins were calculated relative to those of losses, while local displacement, swaps, and global changes were calculated relative to the no change condition. Their impact on hidden states over time were tested with a mixed-effects model incorporating random intercepts and slopes for each subject. Each model was reparameterized to exclude a global intercept, but fit a separate intercept for each group (trial type). Therefore, for each model we tested whether each trial type differed from zero (null hypothesis of no effect) to determine if it had a significant impact on RL parameters.

Initially, we collapsed data over sex and cannabis use to determine what variance between subjects is explained by the model. For learning rate ($\alpha$), there was a significant effect of trial type (of past trials) on the learning rate of the present trial [$F_{4,404.68} = 7.828$, $p < .001$]. There was a significant change from baseline following local displacement [$t(121.5) = 2.916$, $p = .004$, $d = .529$]$^K$ or global changes [$t(121.5) = 2.108$, $p = .037$, $d = .383$]$^L$, but not local swaps or winning outcomes ($p > 0.091$ in both cases). As seen in Figure 4.6.A, all trial types resulted in a slight increase in learning rates relative to baseline (Volterra weight intercept $> 0$). The plot indicates that the effect persists for a maximum of about 10 previous trials.

There was also a significant effect of past trial type on reward strength [$F_{4,404.68} = 19.629$, $p < .001$]. Wins [$t(125.9) = -2.209$, $p = .029$, $d = -.394$] and local swaps [$t(125.9) = -4.266$, $p < .001$, $d = -.761$]$^M$ both caused significant decreases in reward

strength (Volterra weight intercept $< 0$). Therefore, as multiple wins (or swaps) are experienced, future rewards become progressively less impactful on choice. Local displacement and global changes had no effect on reward strength ($p > .461$ in both cases). Punishment strength ($\kappa_2$) was also affected by trial type [$F_{4,404.68} = 67.857$, $p < .001$]. As with $\kappa_1$, wins (relative to losses) decreased the strength of future punishments [$t(142.4) = -8.025$, $p = .029$, $d = -1.345$]. Consequently, losses increased the strength of future punishment, so that experiencing multiple losses would have a cumulative effect. As seen in Figure 4.6.B & C, reward strength quickly recovered in response to wins. However, $\kappa_2$ exhibited a much more prolonged change, suggesting that the effects of losses were more impactful over a longer time course. No other trial type had a significant effect on $\kappa_2$ ($p > .321$ in all cases).

In sum, these data indicate that recent rewards and manipulation of choice target locations increase the learning rate. Wins reduce the sensitivity of subjects to future reward ($\kappa_1$) and punishment ($\kappa_2$), whereas losses increase the sensitivity. We interpret this to indicate that subjects who have been winning on recent trials persist in their long-term strategy (e.g., executive control) rather than engaging in reflexive responding strongly influenced by the immediately previous reinforcement (e.g., sensorimotor control).

We next tested whether cannabis use and sex modulated the response of reinforcement learning parameters to wins, local displacements, swaps, and global changes. We used a mixed effects model with random slopes and intercepts for each subject. In this case, a global intercept was used because we were testing differences between conditions, rather than between each group relative to the null hypothesis of no change within each condition.

Cannabis use and sex had a significant effect on the change in learning rates ($\alpha$) following local displacement, as evidenced by a cannabis $\times$ sex interaction [$F_{1,102} = 4.748$, $p = .032$], while there were no main effects of sex or cannabis use ($p > .178$). The same sex $\times$ cannabis interaction was also present in the response to global changes [$F_{1,102} = 7.443$, $p = .007$]. However, there were no differences in the response to winning outcomes or local swaps ($p > .108$ in all cases). The immediate responses to each trial type (in the following trial, or at lag=1) are shown in Figure 4.7.A. Males exhibited a significant increase in learning rates immediately following local displacement [$t(51) = 2.325$, $p = .024$, $d = .653$]$^N$. Therefore, local displacement increased the rate at which new information updates choice value estimates. Male cannabis users exhibited a similar increase in response to local swaps, though the effect was not statistically significant [$t(51) = 1.801$, $p = .078$, $d = .506$]. Conversely, learning rates fell in response to global changes for male cannabis users, relative to male controls [$t(51) = -2.786$, $p = .007$, $d = -.782$]$^O$. For $\kappa_1$, there was a significant effect of sex on the response to displacement [$F_{1,102} = 4.517$, $p = .036$], as seen in Figure 4.7.B. In addition, there was a significant sex $\times$ cannabis interaction in the effect of global changes on $\kappa_1$ [$F_{1,102} = 6.242$, $p = .014$]. However, for $\kappa_2$, male and female cannabis users did not differ from controls in their response to wins, local displacement, swaps, and global changes ($p > .138$ in all cases).

In sum, male cannabis users tended to increase learning following local, but not global changes of target positions, which is different from all other groups. The parameter values for female cannabis users were not different from controls, which suggests that their reduced task performance is related mostly to processing of the previous reinforcement (e.g., lose-shift) rather than effects spanning multiple trials.

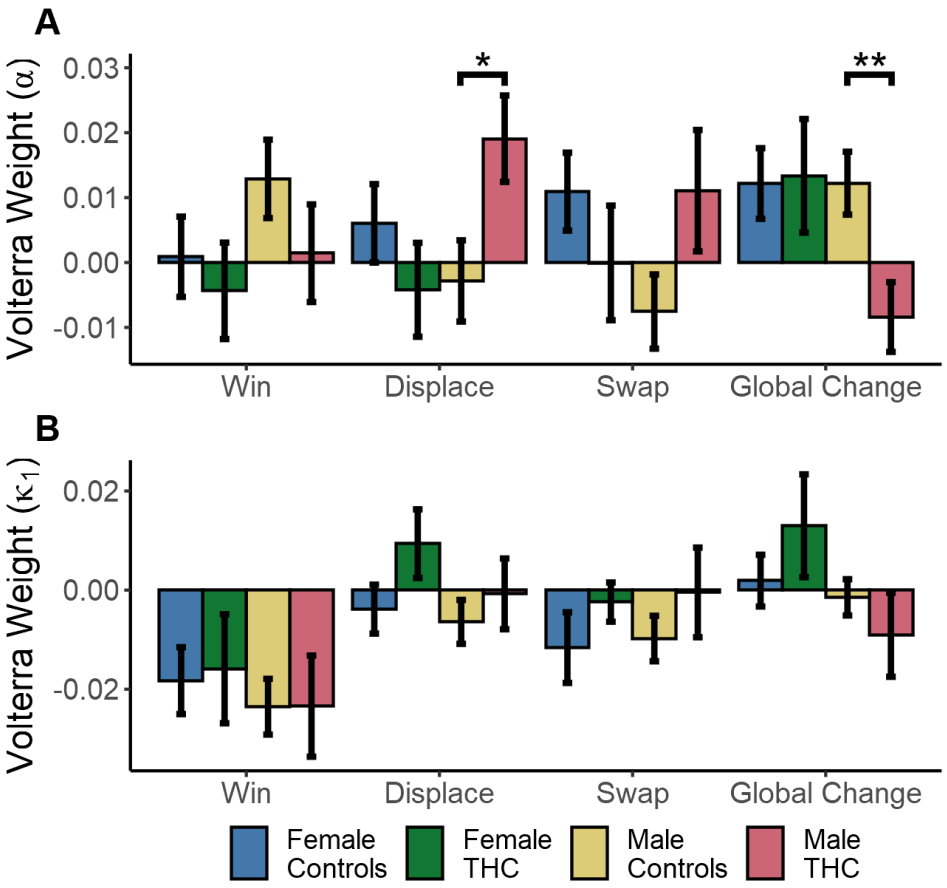


Figure 4.7: Cannabis × Sex interactions on mean reinforcement learning parameter values estimated by Volterra decomposition. Changes in learning rates ($\alpha$, A) & reward valuation ($\kappa_1$, B) immediately (trial lag 1) following wins, local choice displacement, swaps, and global changes in position. SEM in error bars.

## 4.4 Discussion & Conclusions

The current results provide novel behavioural evidence that the lose-shift response is strongly influenced by sensorimotor systems that encode the location of choice targets, and that the regulation of such responding is compromised differently in men and women who frequently use cannabis. A high proportion of lose-shift responding is sub-optimal in the present task because it is predictable, and can therefore be exploited by

the computer opponent. Indeed, the propensity for lose-shift responding is negatively correlated with task performance here. Nonetheless, subjects engage this response above chance levels for several hundred trials before learning to suppress it. This suggests that it is a default strategy, consistent with previous work in humans [Ivan et al., 2018], and analogous to what has been observed in animals performing a similar task [Gruber and Thapa, 2016]. As lose-shift responses eventually converge to chance levels in trials with no global changes, the probability of win-stay responses increase above chance levels. We found that this negative correlation between lose-shift and win-stay was significant and strikingly similar to correlations found with rodents [Gruber and Thapa, 2016], suggesting that lose-shift and win-stay are expressed by neural systems in competition with one another.

We show here that participants overwhelmingly perform the lose-shift according to target position, rather than target identity. In other words, participants avoided the prior position of the previously chosen target when it was unrewarded. This novel observation reveals a strong spatial component to the lose-shift. These data are consistent with the notion that lose-shift is a product of sensorimotor systems. Loss-driven response shifting is reduced following lesions to sensorimotor striatum in animals [Skelin et al., 2014, Gruber et al., 2017, Thapa and Gruber, 2018] or damage to putamen/insula in humans [Danckert et al., 2011], homologous structures that are strongly involved with sensorimotor control. Moreover, decision times are lower for lose-shifting than for lose-stay responses, even following global changes in position necessitate equally distance arm movements. There are multiple reasons this may occur. First, some cortical visual areas may process information about spatial position independently from other object characteristics [Mishkin et al., 1983, Haxby et al.,

1991] and this dorsal "where" pathway processes information more quickly than the ventral "what" pathway [Deubel et al., 1998, Goodale and Milner, 1992]. Secondly, the dorsal pathway may be used to compute actions prior to stimulus presentation. In the perceptual learning literature, activity in both the motor and visual cortices builds up prior to stimulus onset, and reflects stimulus expectation and the associated motor responses [de Lange et al., 2013]. Moreover, pre-response fluctuations in beta-power motor activity are also predictive of choice alternation (i.e., lose-shift), regardless of associated motor action [Pape and Siegel, 2016]. There is evidence that loops involving premotor cortex and the lateral striatum map vision and other sensory modalities into an egocentric space. The ventral premotor cortex contains neurons that both drive motor actions, but also encode locations of visual, tactile, and auditory stimuli [Fadiga et al., 2000]. Consequently, they form a motor vocabulary for mapping several modalities into an actions in a common egocentric space. Even when stimuli are removed, these neurons still respond to the position of remembered objects in relation to the body [Graziano and Gross, 1998]. The putamen (LS in rodents) also contains these bimodal visuomotor cells [Graziano and Gross, 1996], and therefore has the capacity to mediate lose-shift from a remembered location. In the context of our study, the spatial rearrangement of choice targets subverts this motor preparation, requiring choices to be recalculated following stimulus onset. This is evidenced by the increase in response times following local swaps and displacement. Interestingly, local swaps had a larger and more consistent effect on response times than did displacement, suggesting that a greater level of motor recalculation is required. Specifically, we speculate that it requires more time for the executive control to overcome the intrinsic inhibition of a previously unrewarded action (than a novel one) in the motor system

in order to intentionally select it.

The influence of local and global changes in choice position also highlights the importance of egocentric and allocentric processing of space. While local changes in choice orientation modulate the lose-shift, these effects persist even when all choices are moved to a new global position relative to the observer. Conversely, the win-stay is much less affected by spatial position. Local changes do have an effect on behaviour, but these are eliminated by concurrent global changes. These results highlight the importance of allocentric processing on the lose-shift. Choice positions are calculated relative to one another, allowing their associated values to be maintained across large global movements in choice position. Conversely, while processing of the win-stay is less reliant on spatial information, egocentric reference frames are more important than allocentric, where choice value is calculated relative to the subject. Consequently, local and global changes have a large effect on win-stay behaviour.

In addition to driving different decision strategies, the putamen/LS and ventral striatum (VS, including the nucleus accumbens) also respond differently to psychoactive drugs. Relative to the VS, the LS exhibits a much higher density of dopamine transporters [Coulter et al., 1997], endocannabinoid receptors [Herkenham et al., 1991], opioid receptors [Benfenati et al., 1991], and alcohol-sensitive NMDA receptors [Liste et al., 1995]. Consequently, THC administration temporarily increases dopamine release throughout the striatum, but particularly in the LS [Jentsch et al., 1998, Sakurai-Yamashita et al., 1989]. The effects of acute ethanol exposure are similar, though greater in the VS [Vena et al., 2016, Clarke et al., 2015]. Conversely, long-term sensitization to alcohol and cannabis reduces availability of striatal dopamine receptors [Budygin et al., 2007, Martinez et al., 2005, Albrecht et al., 2013] and

cannabinoid receptors [Villares, 2007], especially in the LS. Chronic exposure also inhibits the prefrontal cortex, shifting choice control to the LS [Lucantonio et al., 2014, Everitt and Robbins, 2005, 2013]. We expect this effect to impair the ability of participants to use executive control to suppress lose-shift responding by the sensorimotor systems. We do not have sufficient primary evidence to hypothesize how the change in receptor densities by repeated alcohol/THC exposure affects lose-shift processing within the LS and/or other components of the sensorimotor system.

In the present study, we find that self-reported level of recreational use of cannabis affects task performance, but that this effect differs on the basis of biological sex. Elevated cannabis use in men decreased spatial modulation of the lose-shift, possibly through dopaminergic desensitization of the LS. As seen in Figure 4.4, baseline lose-shifting is also reduced, falling below 50% in trial blocks 4 and 5. With this reduction, lose-shift responding after swaps increases to 63%. Therefore, either the calculation of the lose-shift is affected or its suppression by executive systems is enhanced in male cannabis users, while spatial processing remains unaffected. Conversely, female cannabis users exhibit decreased task performance and choice entropy - behaviours thought to rely on the suppression of sensorimotor responding by the prefrontal cortex. Furthermore, they show a moderate and significant increase in baseline lose-shift responding [$F_{1,51} = 4.109$, $p < .048$], revealed by a mixed effects model between female controls and cannabis users in the no change condition (Fig. 4.3.C). It thus appears that females with high cannabis use exert less executive control over sensorimotor systems in our task.

While it is tempting to describe this sex difference as a consequence of different drug effects on the brain in males and females, we cannot strongly support this

inference based on the present data. Several alternatives are possible. It is possible that the effect is due to a confounding factor that promotes high levels of recreational drug use and also impairs sensorimotor regulation. Unfortunately, the WHO ASSIST is not sufficient to infer whether these are the case in the present study. However, it is known that females are more susceptible to drug tolerance (including cannabis) and sensitization than are males [Wakley et al., 2014, Robinson, 1988]. Drug use is also comorbid with mood and anxiety disorders, particularly depression [Zilberman et al., 2003], which causes heightened loss aversion [Beevers et al., 2013]. These differences are possibly due to the effects of estrogen, which enhances striatal dopamine release in response to psychoactive drugs [Becker, 1999] and alters the effects of drugs on the prefrontal cortex. Females rats with high estrogen levels exhibit dysfunction of the prefrontal cortex relative to males and low-estrogen females when exposed to dopamine-enhancing drugs [Shansky et al., 2004]. Estrogen also heightens the effects of cocaine and amphetamine, causing an abnormal BOLD response in rats [Sárvári et al., 2014, Febo et al., 2005]. Alcohol and cannabis consumption also increase oestradiol levels, and can inhibit testosterone production in males [Yonker et al., 2005, Purohit, 2000, Maskarinec et al., 1978, Kolodny et al., 1974, Harclerode, 1984]. In males, increased estrogen and reduced testosterone levels cause declines in spatial cognitive ability Janowsky et al. [1994]. Therefore, the heightened susceptibility of the PFC to the combined effects of estrogen and drug abuse provides an explanation for why only women with high ASSIST scores show a dominance of sensorimotor control, without compromising the spatial dependence of lose-shift. Specifically, this population had accelerated decision speeds, lower proportion of wins, and a tendency for lower entropy of response sequences. On the other hand, the lose-shift remained

sensitive to swapping cue locations, which is similar to controls, but opposite of what is observed in males with high ASSIST scores. Our analysis of behaviour through a reinforcement learning framework also revealed a cannabis $\times$ sex interaction. Whereas the other analysis presented here focuses on the effects of the previous trial, the Q-learning model allowed us to examine effects that span many trials. It was the men who used cannabis who stood out in this analysis; they had increased learning when previous cues were displaced locally, and decreased learning when previous cues were switched globally. We expect such learning is part of a reinforcement learning scheme in 'goal-directed' brain circuits linked more closely to executive function than sensorimotor control [Balleine and O'Doherty, 2010, Gruber and McDonald, 2012], suggesting that not only is there an enhanced suppression of sensorimotor control by executive function in male cannabis users, but that adaptation by the executive system is also different than the other groups. It is worth noting, however, that our sample (as is common in the field) was predominantly young university students, who are presumably well educated and high functioning. We urge caution in extrapolating our findings to the general public.

The interpretation of data in this study faces several challenges besides the aforementioned limitations of the ASSIST. First, alcohol and cannabis use are highly concordant (Spearman's correlation of $\rho$=.533, $p < .001$ in our sample), and likely additive in their effects. Second, the sexually dimorphic effects observed here may be due to confounding interactions between drug use, IQ, and/or psychiatric disorders that have different prevalence among the sexes. However, the sexually dimorphic distribution of endocannabinoid receptors in the striatum and prefrontal cortex [De Fonseca et al., 1994] likely also play an important role. For instance, errors

when reconstructing spatio-temporal sequences were reduced in men and increased in women following THC treatment [Makela et al., 2006]. We previously reported that lose-shift is decreased by acute administration of THC in female rats [Wong et al., 2017a]. It is possible that the down regulation of receptors in have users may cause the inverse, which would be consistent with the data here.

In sum, the data presented here indicate that lose-shift responding is a useful gauge of the cognitive control over sensorimotor responding in humans, and that this is impacted differently in men and women that heavily use cannabis. These linkages are important factors to account for the impact of lose-shift responding in real-world economic decision making, such as gambling [Abouzari et al., 2015, Worthy et al., 2013], in addition to clinical/laboratory testing of cognitive flexibility with tasks such as the Wisconsin Card Sorting Task that involve loss-based shifting of response policies.

| Test | Dependent Var | Independent Var | *df* | $\Delta$ | *t* | *p* | *d* | *LCI* | *UCI* |
|---|---|---|---|---|---|---|---|---|---|
| A | Wins | Female Controls v THC | 51 | -.021 | -3.123 | .003 | -.934 | -1.562 | -.307 |
| B | Entropy | Female Controls v THC | 51 | -.033 | -2.201 | .032 | -.658 | -1.273 | -.044 |
| C | RT | Female Controls v THC | 51 | .142 | 3.024 | .004 | .904 | .279 | 1.531 |
| D | LSW | Swap v No Change | 529 | -.102 | -7.249 | <.001 | -.507 | -.652 | -.361 |
| E | LSW | Swap v No Change (Global) | 529 | -.089 | -8.056 | <.001 | -.459 | -.586 | -.333 |
| F | LSW | Displace v No Change (Global) | 529 | -.028 | -2.827 | .005 | -.160 | -.272 | -.048 |
| G | WST | Swap v No Change | 529 | -.093 | -6.565 | <.001 | -.438 | -.575 | -.301 |
| H | WST | Displace v No Change | 529 | -.046 | -4.388 | <.001 | -.221 | -.321 | -.121 |
| I | LSW Swap Effect | Female Controls v THC | 263 | .090 | 2.157 | .032 | .289 | .024 | .553 |
| J | LSW Swap Effect | Male Controls v THC | 263 | -.152 | -3.737 | <.001 | -.469 | -.720 | -.219 |
| K | $\alpha$ | Displace v Baseline | 121.5 | .005 | 2.916 | .004 | .529 | .126 | .884 |
| L | $\alpha$ | Global v Baseline | 121.5 | .004 | 2.108 | .037 | .383 | .025 | .715 |
| M | $\kappa_1$ | Swap v Baseline | 125.9 | -.007 | -4.266 | <.001 | -.761 | -1.113 | -.390 |
| N | $\alpha$ Displace Effect | Males Controls v THC | 51 | .022 | 2.325 | .024 | .653 | .075 | 1.231 |
| O | $\alpha$ Global Effect | Males Controls v THC | 51 | -.021 | -2.786 | .007 | -.782 | -1.366 | -.198 |

Table 4.3: Table summarizing key significant effects. Note: $\Delta$ is difference between means, $d$ is Cohen's $d$ along with 95% confidence intervals.

# Chapter 5

# Neural Correlates of the Win-Stay and Lose-Shift

## 5.1 Introduction

Chapters 2 - 4 provide behavioural evidence that choice value is calculated in spatial-motor coordinates: the location of choices, their positions relative to one another, and associated motor actions have a strong bearing on the decision to win-stay or lose-shift. These results are consistent with findings that the sensorimotor striatum drives win-stay and lose-shift responding in rats [Skelin et al., 2014]. However, human anatomical data would further prove that economic decision-making relies on spatial-motor processing. Therefore, the current experiment used high-density electroencephalography (EEG) to measure neural activity from sixty-seven subjects while they played the Matching Pennies game used in previous chapters. The eLORETA method of source localization was used to improve the otherwise poor spatial resolution of EEG, allowing identification of the cortical structures activated during

win-stay and lose-shift responding. While EEG-based source localization cannot directly assess striatal activity, it provides knowledge concerning the cortical networks which project to the striatum and are necessary decision-making.

Decision-making consists of several processing stages that influence win-stay and lose-shift behaviour. Following rewards and punishments several EEG components reflect changes in action-outcome associations after unexpected outcomes. Chief among these is the feedback-related negativity (FRN), a negative frono-central scalp potential elicited during punishment (relative to reward) [Nieuwenhuis et al., 2001]. The FRN is believed to reflect neural activity that re-weights action values when expectations of reward are violated [Holroyd and Coles, 2002] and is generated by a prediction error signal in the anterior cingulate [Holroyd et al., 2003]. As the FRN is related to the salience of rewards and punishments, the decision to win-stay or lose-shift is preceded by a change in FRN amplitudes [Cohen and Ranganath, 2007].

While limbic and associative systems adjust behaviour in response to feedback, motor circuits are equally important to decision-making. Motor activity is often treated as the final output of decision-making, rather than having an active role in determining choice value. However, in tasks where choices are represented by distinct motor actions, punishment results in greater negativity over the region of motor cortex corresponding to that action [Cohen and Ranganath, 2007]. Motor representations of choice value persist even during motor execution. Lateralization of motor beta oscillations differs between shift and stay behaviour during action execution [Pape and Siegel, 2016], even for identical motor responses. Therefore, motor activity is contextual and not a simple representation of fixed actions. For example, the expectation of future perceptual stimuli modulates the motor activity associated with responding

to that stimuli [de Lange et al., 2013]. The use of egocentric or allocentric spatial frames of reference influence activity in the premotor cortex, supplementary motor area, and thalamus [Jordan et al., 2004]. Therefore, motor circuits not only code for actions, but also their value and the context in which they are made.

In the present chapter we further explore the systems associated with win-stay and lose-shift behaviour in humans. Rewards leading to a win-stay response are preceded by greater activation of somatosensory, premotor, and visuospatial cortices. Punishments leading to a lose-shift response recruits similar signaling pathways, in addition to limbic and associative circuitry. Therefore, extensive cortical networks update the value associated with win-stay and lose-shift responding, of which sensorimotor and visuospatial systems are an important element.

Execution of motor actions also results in activity unique to sensorimotor response strategies. Win-stay responses were associated with greater activation of the primary motor and somatosensory cortices, particularly at beta frequencies known to be indicative of motor processing. This activity was restricted to the hemisphere contralateral to the hand used for motor action, suggesting win-stay responding is driven by increased motor signaling specific to that hand. Consequently, the win-stay may be processed in egocentric spatial coordinates, there being a 1:1 correspondence between choice location and sensorimotor activity.

Lose-shifting was also preceded by increased activity in limbic and associative circuits, relative to the lose-stay. In particular, delta activity was present in the anterior cingulate and frontal cortex ipsilateral to the hand being used. Delta activity in these higher-order circuits is associated with concentration and inhibition of motor responses Harmony [2013]. These regions are also associated with allocentric spatial

processing [Zaehle et al., 2007, Gramann et al., 2006]. Therefore, the lack of a direct correspondence between motor action and activity supports our previous findings (ch. 2 - 4) that the lose-shift is represented in allocentric spatial coordinates.

Evidence suggests complex networks are responsible for win-stay and lose-shift responding, rather than a single neural structure. Therefore, we investigated how rewards and punishments influenced connectivity across the entire brain and whether it could predict subject behaviour. We found that reduced connectivity across the cortex preceded sensorimotor responding. Moreover, subjects who had a greater tendency to win-stay or lose-shift exhibited the greatest decrease in connectivity between a variety of motor, visual, frontal, and limbic systems. Together these data (28400 trials from 71 subjects) demonstrate extensive networks drive the win-stay and lose-shift of which motor and visuospatial systems are key elements.

## 5.2 Methods

### 5.2.1 Behavioural Task

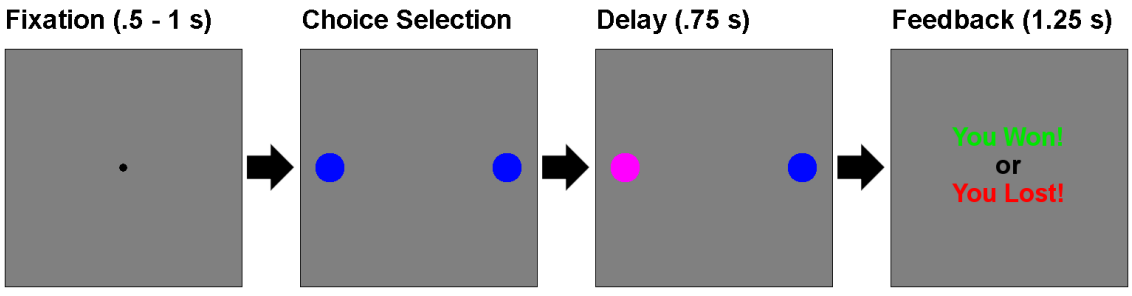


Figure 5.1: Time course the matching pennies task.

During the experiment, subjects played the "Matching Pennies" game against a

computer opponent. The task display consisted of two blue buttons presented on a 15.4" CRT monitor at a refresh rate of 85 Hz (Figure 5.1). Each trial began with the presentation of a fixation point for a random interval between 0.5 - 1.0 seconds. Afterwards, two choices would be presented and participants would select either the left or right target using a keyboard. After a delay of 0.75 s, they would receive visual feedback indicating "You Win" or "You Lose" for 1.25 s paired with an auditory tone. Prior to choice selection, the computer attempted to predict which target would be chosen using the algorithms described in Chapter 2. If the participant selected this target, the trial was a loss. Otherwise it was a win. The algorithm attempted to minimize the number wins for participants. The optimal strategy is to be unpredictable in choice, which leads to wins on 50% of trials. Because win-stay and lose-shift responses are predictable, subjects should learn to suppress these behaviours as the session progresses. When coupled with simultaneous EEG recordings, this task provides measures of the neural processes associated with the lose-shift, win-stay, and suppression of these sensorimotor responses.

Participants also completed the Wechsler adult intelligence scale (WAIS) IV. The WAIS-IV consists of ten tests, grouped into four sub-domains that assess verbal comprehension, perceptual reasoning (PRI), working memory (WMI), and processing speed (PSI). The verbal comprehension index (VCI) describes the ability to think verbally and the extent of one's acquired knowledge, including history, geography, and word definitions. Ability in this domain largely relies on function of the left inferior frontal cortex [i.e., Broca's Area Gläscher et al., 2009]. The perceptual reasoning index (PRI) corresponds to integration of visuospatial information with motor action, and relies on processing in the right parietal, occipito-parietal and superior

temporal cortices. The working memory index (WMI) encapsulates the manipulation of auditory information in short term memory, and relies on the superior parietal cortex. The prefrontal cortex, dorsal cingulate, premotor cortex, and posterior parietal cortex are also vital to working memory across a variety of task domains [Owen et al., 2005]. Finally, the processing speed index (PSI) indicates the speed at which visual information can be interpreted and met with a motor response. It relies on a number of regions distributed throughout the brain including the left precentral, inferior parietal, and lingual gyri, the left postcentral sulcus, the right middle frontal gyrus, and right inferior frontal gyrus [Gläscher et al., 2009].

### 5.2.2 Procedure

All procedures and experimental tasks were approved by the McMaster University Research Ethics Board. Seventy-one undergraduates (32 males, mean age = 19.54, $SD$ = 3.48) from McMaster University participated in the study over two 2-hour sessions in exchange for course credit and/or payment. All were right handed, and had normal or corrected-to-normal vision and hearing. On day 1, after providing informed consent and being fitted with an EEG system, participants played 400 trials of matching pennies. They were informed that they would win nothing each time the computer predicted their choice and 3 ¢ each time it could not, rounded up to the nearest $5 upon completion of the experiment. Participants were given no guidance as to optimal decision-making strategies. Afterwards they participated in 396 trials of another task, for publication in a separate paper.

The second session occurred at the same time of day within 7 days of the first. Subjects first completed the Wechsler adult intelligence scale (WAIS) IV. Performance on

the 10 primary WAIS-IV tests were converted to VCI, PRI, WMI, and PSI scores and scaled against the American reference group. After task completion, participants were screened with the South Oaks Gambling Screen, the alcohol, smoking and substance involvement screening test (ASSIST) v3.0, Adult ADHD Self-Report Scale v1.1, depression, anxiety, and stress scales (DASS), the 42-item revised obsessive-compulsive inventory (OCI-R), and an additional demographic questionnaire. Habitual cannabis users were defined as those meeting the criteria for brief or intensive treatment (score $> 3$) on the ASSIST cannabis subtest. Males had a mean cannabis-use score of 5.34, with 14 of 32 meeting the criteria for cannabis use requiring intervention. Females had a mean score of 5.64, with 14 of 39 meeting cannabis use criteria. While behavioural data was collected from all subjects, 4 lacked EEG scans due to hairstyle, 2 lacked IQ tests due to cancellations, and 11 lacked DASS and OCI-R questionnaires.

### 5.2.3 EEG Recording and Analysis

**EEG preprocessing** EEG data were recorded from 67 subjects using a BioSemi ActiveTwo system with 128 scalp, two mastoid, and two horizontal ocular electrodes. All data pre-processing was conducted in Matlab, using the EEGLAB toolbox [Delorme and Makeig, 2004]. Data were band-pass filtered from 1 to 100 Hz using zero-phase, non-causal filters (0.5/112.5 Hz cutoff, 1/25 Hz transition bandwidth). Sinusoidal 60 Hz line noise and 85 Hz CRT monitor noise was removed using the Cleanline function [Mullen, 2012] with a 2 Hz bandwidth at each frequency. Transient, high-amplitude artifacts were removed using artifact subspace reconstruction [Mullen et al., 2013] with a $\sigma$ of 25. Remaining noisy channels ($SD \gtrapprox 30\ \mu V$) were replaced using spherical interpolation. Data were then re-referenced to a theoretical

zero reference (i.e., neutral potential) calculated at an infinite distance from all electrodes using the reference electrode standardization technique [REST; Yao, 2001, Dong et al., 2017]. While REST is a relatively new reference technique, it has been shown to result in less distorted estimates of connectivity than the average reference [Chella et al., 2016], lower error in estimations of source activity [Marzetti et al., 2007], greater estimates of brain lateralization [Zheng et al., 2018], and less spatial smearing of scalp activity, due to the average reference acting as a spatial low-pass filter [Kayser and Tenke, 2015, Zheng et al., 2018].

Additional non-neural sources of noise were isolated via adaptive mixture independent component analysis [AMICA; Palmer et al., 2012]. Components corresponding to muscle activity, eye movements, electrode noise, line noise, and other artifacts were identified and removed manually with aid of the ICLabel package Pion-Tonachini et al. [2019]. On average, cleaned data retained 61 independent components for each subject, with 70 being removed during cleaning. Epochs were extracted for averaging and analysis, time-locked to choice execution and feedback presentation, and referenced to a -300 to -100 ms, pre-stimulus baseline.

**Scalp voltage potentials** To investigate feedback-related processing, scalp voltage potentials were measured at channel Cz following feedback presentation, where the greatest feedback-related negativity amplitudes occur [Scheffers and Coles, 2000]. The difference between loss and win-associated activity was compared at each time point in the 1000 ms following feedback presentation using paired t-tests. In particular, voltage potentials related to reward processing were analyzed including the early feedback-related negativity (eFRN), feedback-related negativity (FRN), and error related positivities (P3a and P3b). Whole-head maps of scalp voltage potentials were

also measured throughout the course of reward processing.

The decision to engage in shift or stay responding is largely determined by the impact of rewards and punishments on choice value. For example, unexpected, highly salient losses should cause a large neural response that leads to lose-shift behaviour. Therefore, differences in voltage potentials leading to shift vs stay responses were assessed using paired t-tests (i.e., LSW - LST and WST - WSW differences in activity). These responses were termed the loss-predictive FRN (LpFRN) and win-predictive FRN (WpFRN) respectively. The LpFRN and WpFRN were assessed at channels Cz and Oz respectively.

**eLORETA source localization** Source localization was used to reconstruct three-dimensional maps of brain activity from scalp voltage potentials. During the experiment electrode positions and fiducials (i.e., Nasion, LPA, and RPA) were digitized using a Polhemus Patriot. Electrode positions were co-registered to the MNI head template and projected onto the template surface. Locations were then transformed into Talairach coordinates for export to the LORETA-Key software Pascual-Marqui et al. [1994]. Exact LORETA [eLORETA; Pascual-Marqui, 2007] was used to compute the three-dimensional distribution of current source densities (CSD) across the entire brain from scalp-recorded electric potentials. During this analysis the brain is partitioned in 6239 voxels at a $5\,\mathrm{mm}^3$ spatial resolution. The use of eLORETA for source localization is beneficial in that it allows identification of cortical generators that drive behaviour at a higher spatial resolution than ordinary EEG without sacrificing temporal precision. In addition, eLORETA has advantages over other inverse solutions such as sLORETA [Pascual-Marqui et al., 2002], exhibiting zero localization error and fewer false-positives than sLORETA, beamforming, minimum norm

estimates, or dynamic statistical parametric mapping Pascual-Marqui et al. [2018]. EEG-based source localization does not include sub-cortical circuits such as the basal ganglia. However, given the distinct cortical inputs to the striatum, eLORETA should provide insights into the sub-cortical circuits necessary to the win-stay and lose-shift.

eLORETA solutions were calculated during feedback processing and motor response selection using a signal-noise ratio of 100. Whole brain maps of the LpFRN and WpFRN compared differences in shift and stay related current source densities at each time point in the 1000 ms following feedback processing. These event-related changes in CSDs result from oscillations produced by interacting neural circuits. Therefore, frequency response maps were calculated in the $\delta$ (1-4 Hz), $\theta$ (4-8 Hz), $\alpha_1$ (8-10.5 Hz), $\alpha_2$ (10.5-13 Hz), $\beta_1$ (13-20 Hz), $\beta_2$ (20-30 Hz), $\gamma_1$ (30-50 Hz), and $\gamma_2$ (50-80 Hz) bands over the second following feedback. Differences in oscillatory dynamics during the execution of shift and stay responses were also calculated in the -100 to 500 ms interval time-locked to choice selection. This interval was chosen so as not to overlap with feedback-related processing. Separate maps were generated for responses made with the left and right hands.

**Statistical Analyses** Paired $t$ tests assessed LSW - LST and WST - WSW differences in scalp voltage potentials, current source densities, and non-linear connectivities. These tests were applied independently to each time point, voxel, EEG channel, or frequency band in question. Current source densities were log-transformed prior to testing as they follow a non-normal distribution, being equivalent to amplitude squared. No additional normalization, spatial, or temporal smoothing of brain activity was used. The relationship between brain connectivity and rates of win-stay and lose-shift responding was assessed via correlation.

The exploratory nature of this chapter, and the large number of statistical tests conducted (i.e., 6239 voxels and 8 frequency bands) necessitated a relatively low threshold for statistical significance. However, as brain activity in adjacent voxels or time points are statistically dependent, the Bonferroni correction was deemed unnecessarily conservative. Instead a significance threshold of $p < .0005$ was used for each test. This level of $\alpha$ is more conservative than the $p < .001$ threshold used in most fMRI studies. Given N = 67 and $\alpha = .0005$, Type II error rates ($\beta$) of .4, .2, & .1 would result in effect sizes ($d_z$) of .48, .55, and .61 respectively. Therefore, our study is sensitive enough to detect moderate effect sizes.

For all eLORETA results, tables summarizing the top four local maxima (distinct brain structures) are provided. For each region, the number of above-threshold voxels ($p < .0005$), max activity $z$-score, and the MNI-305 coordinates of the local maxima are given [Evans et al., 1992, Collins et al., 1994]. The laterality index ($LI$) for each region is also given, describing the relative activation of the left and right hemispheres. It was calculated as:

$$LI = \frac{Q_{LH} - Q_{RH}}{Q_{LH} + Q_{RH}} \tag{5.1}$$

where $Q_{LH}$ and $Q_{RH}$ are the quantities of above threshold voxels ($p < .0005$) in the left and right hemispheres, not counting those present on the midline. A $LI$ of $\pm 1$ indicates that only the left/right hemispheres are active, while a $LI$ of zero indicates they are equally active.

We also assessed the effects of sex and cannabis use on subject behaviour, cognitive ability, mental-health outcomes, and task-related neural activity. Wins experienced, lose-shift responses, win-stay responses, response entropy, and decision times were

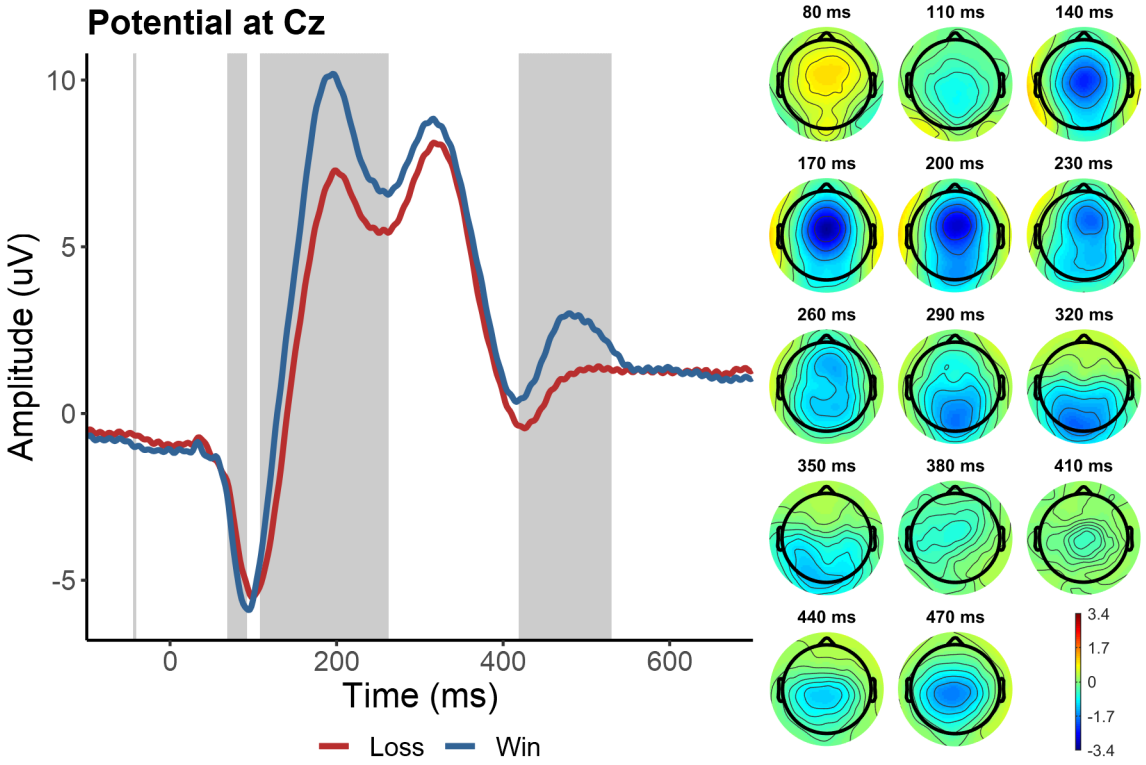


Figure 5.2: Left: neural response to losses and wins, time locked to feedback presentation, measured at channel Cz. Shaded regions indicate time-points in which loss-win differences were significantly different at $p < .0005$. Right: loss - win difference in voltage potentials across the entire scalp.

averaged over the entire experimental session for each participant. The effects of cannabis use and biological sex on behaviour were tested using $2 \times 2$ ANVOVAs that employed Type-III sums-of-squares and the sum-to-zero constraint. Similar ANOVAs were applied to the WAIS-IV verbal comprehension, perceptual reasoning, working memory, and processing speed scores, and self-reported depression, anxiety, stress, and ADHD symptomatology scores. In addition, we tested whether sex and cannabis use were associated with differences in scalp volatage potentials and source activity during reward processing, win-stay responding, and lose-shift responding. Differences between cannabis users and controls among males and females were also assessed using

separate, planned $t$ tests.

## 5.3 Results

### 5.3.1 Time course of reward processing

Rewards and punishments have a large effect on our future decisions and are accompanied by several neural markers including the early feedback-related negativity (eFRN), FRN, P3a, and P3b. The FRN is thought to represent violations of reward expectancy [Holroyd and Coles, 2002] generated by a prediction error signal propagating from the midbrain to anterior cingulate [Holroyd et al., 2003]. Consequently, it is correlated with the valence of outcomes [Yeung and Sanfey, 2004]. However, the P3a relates to attention and the context in which actions and outcomes are viewed [Donchin and Coles, 1988, Linden, 2005]. In the context of economic decision-making the P3a reflects comparison of current outcomes against the choice-reward model being used. It is sensitive to outcome magnitude [Yeung and Sanfey, 2004] and is produced by cortical generators in the cingulate, insula, inferior temporal, and superior parietal cortices [Linden, 2005]. In addition, the P3a is correlated with activity in the ventral striatum [Pfabigan et al., 2014]. Given the ventro-medial striatum also codes action-outcome associations needed for model updating Skelin et al. [2014], Burton et al. [2015], the P3a likely reflects processes related to context evaluation in the ventral striatum. The early FRN is thought to be an early stage of the P3a, on which the FRN is superimposed [San Martín, 2012]. Consequently, it also is sensitive to outcome magnitude [Goyer et al., 2008]. Finally, the P3b is thought to reflect memory updating processes [Donchin, 1981] modulated by attention and outcome

probability [Duncan-Johnson and Donchin, 1977]. Generators of the P3b include the supramarginal gyrus and the caudal superior temporal gyrus [San Martín, 2012].

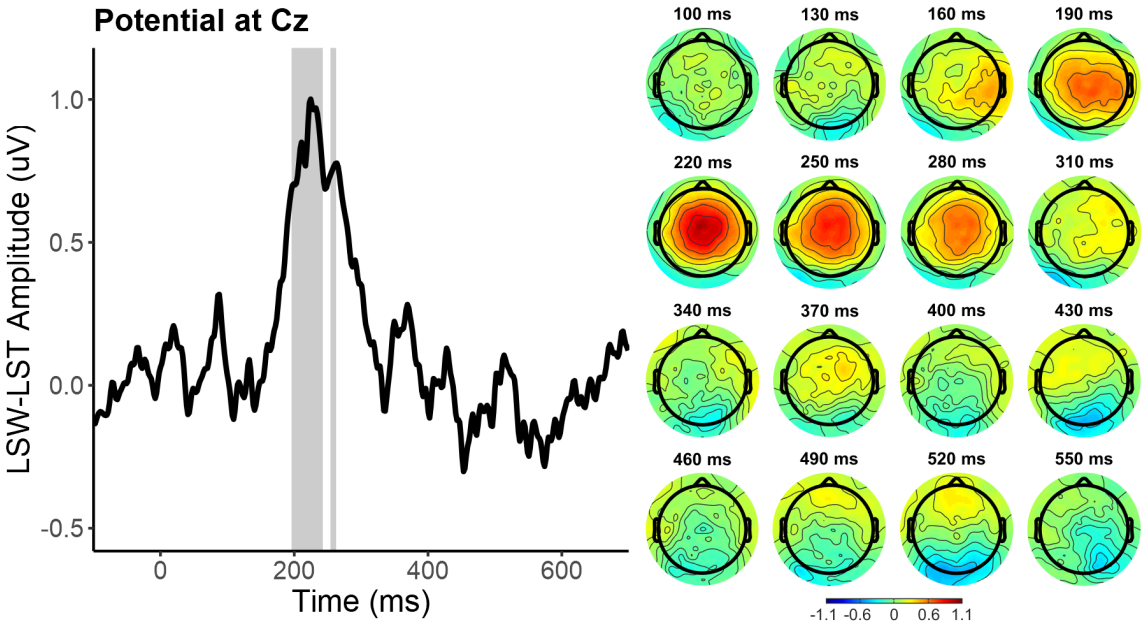


Figure 5.3: Difference in neural response to punishment that led to lose-shifting vs lose-stay behaviour at channel Cz (left) and across the entire scalp (right).

Scalp potentials at channel Cz were analyzed for the eFRN, FRN, P3a, and P3b using paired $t$ tests to assess differences in win and loss-related activity. As seen in Figure 5.2, an evoked potential (N100) peaking at 98 ms captured the early response to the audiovisual reward feedback. Although the N100 significantly differed between wins and losses ($p < .0005$), it occurred too early to reflect reward processing [Johannes et al., 1995]. Instead voltage differences were likely driven by the greater luminance of the "You Win" text (115.2 cd/$m^2$) relative to "You Lose" (36.6 cd/$m^2$) [Series, 2011].

The early FRN (P200), FRN (N250), P3a, and P3b each peaked at 197, 263, 316, and 478 ms post-feedback. Scalp potentials following losses significantly differed from

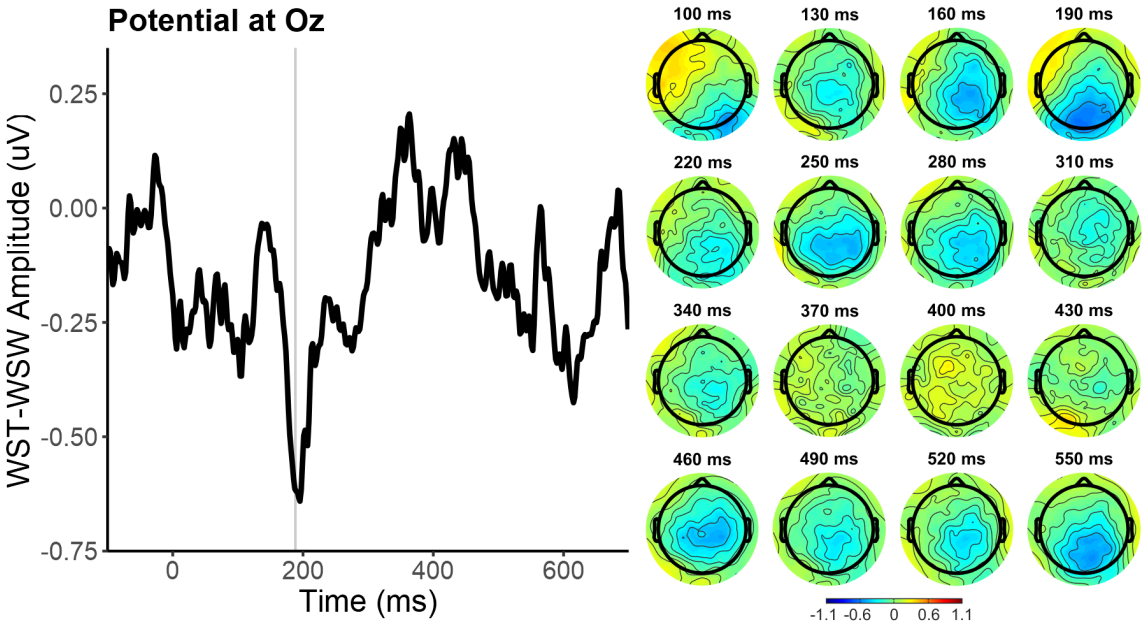


Figure 5.4: Difference in neural response to reward that led to win-stay vs win-shift behaviour at channel Oz (left) and across the entire scalp (right).

wins during the eFRN, FRN, and P3b, but not during the P3a. In probabilistic tasks P3a amplitudes are affected by outcome type and novelty San Martín et al. [2010]. However, in our task the likelihoods of reward and punishment were approximately equal ($M = 47.21\%$ win rate), mitigating the effects of outcome type on P3a amplitudes. The significant differences in eFRN, FRN, and P3b amplitudes associated with wins and losses was typical of previous studies on reward processing San Martín [2012].

The decision to engage in habitual lose-shift or win-stay behaviour was also preceded by altered reward processing. Cohen and Ranganath [2007] demonstrated that the decision to lose-shift was preceded by a modulation of the feedback-related negativity. Therefore, we compared punishment-related activity that preceded a lose-shift against that leading to a lose-stay response. As reported by Cohen and Ranganath

we found that lose-shifting was preceded by significantly greater Cz amplitudes 195-265 ms following feedback on the previous trial (Fig. 5.3). This loss-predictive FRN (LpFRN) peaked at 226 ms post-feedback and occurred between the eFRN and FRN. As with the FRN, the LpFRN was centered over the vertex of the head. This result is consistent with the hypothesis that the tendency to shift choices after a loss is due to a modulation of the same error signal in the ACC that generates the FRN.

Although the lose-shift was preceded by a modulation of the FRN, win-stay behaviour was not. As seen in Figure 5.4 the time course of win-stay predictive activity fluctuated slightly over the occipital lobe, suggesting visuospatial processing may be important to the win-stay. However, activity only differed significantly from the win-shift at 188 ms post-feedback. Cortical circuits may not be responsible for win-stay related reward signaling; however, a more likely explanation is that the WpFRN is not strongly phase-locked to feedback presentation. If so, the WpFRN would manifest as changes in oscillatory activity rather than as an event-related potential. Regardless, our results are consistent with the idea that separate cortical circuits are responsible for win-stay and lose-shift behaviour.

### 5.3.2 Source Localization

Scalp ERPs helped identify the time course of win-stay and lose-shift processing, but lack the spatial precision needed to locate the cortical generators of these potentials. Therefore, eLORETA was used to localize cortical generators of sensorimotor responding. Lose-shift vs lose-stay and win-stay vs win-shift differences in trial-averaged current source densities were assessed using paired $t$ tests.

As seen in Figure 5.5 and Table 5.1 cortical activities derived using eLORETA were

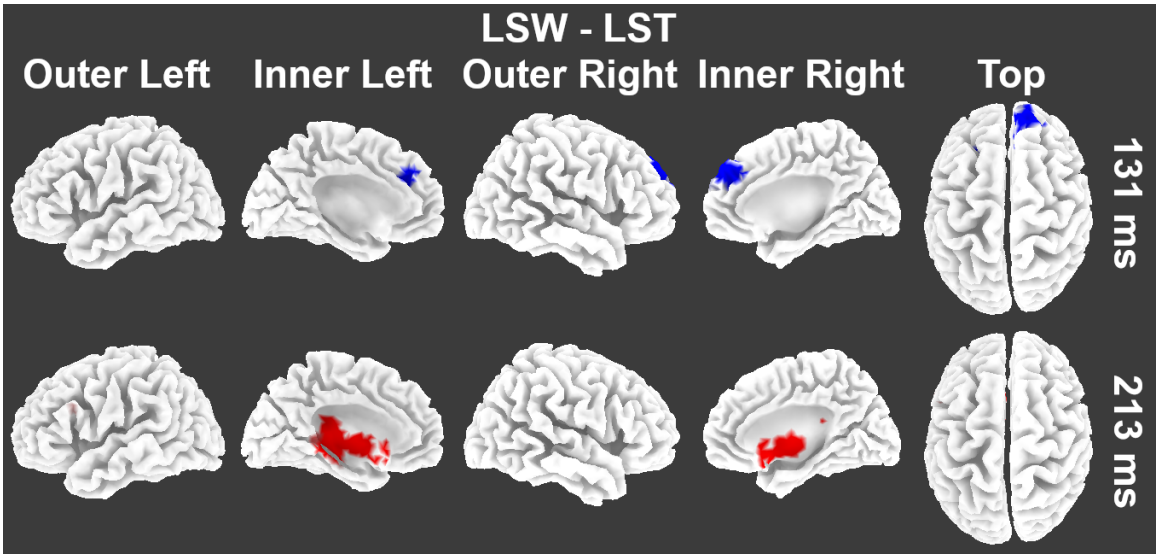


Figure 5.5: Cortical surface maps of voxels exhibiting significant differences ($p < .0005$) in feedback-evoked current source densities leading to a lose-shift vs lose-stay. Regions in red indicate significantly greater lose-shift associated activity (relative to the lose-stay) while those highlighted in blue indicate significantly lesser current source densities. Cortical activities are time-locked to feedback presentation.

consistent with scalp voltage potentials. Greater activation of the anterior cingulate 213 ms after feedback presentation led to lose-shifting. Therefore, the loss-predictive FRN (LpFRN) is a modulation of the same error signal responsible for the FRN. Limbic structures including the insula and subcallosal gyrus also were more active preceding the lose-shift, as was the parahippocampal gyrus. The primary motor cortex (precentral gyrus) also was significantly active 227 ms after punishment (Table 5.1), suggesting the motor cortex actively codes for loss aversion.

As with scalp ERPs, win-stay responding was not strongly associated with potentials phase-locked to reward presentation. However, the right supramarginal gyrus located in the inferior parietal lobe was significantly less active prior to win-stay

| Time (ms) | Region | Voxels | LI | Max $z$ | X | Y | Z |
|---|---|---|---|---|---|---|---|
| | Lose-Shift > Lose-Stay | | | | | | |
| 213 | Anterior Cingulate | 11 | -.429 | 4.284 | 0 | 0 | -5 |
| | Parahippocampal Gyrus | 7 | 1 | 3.987 | -10 | -35 | 0 |
| | Subcallosal Gyrus | 5 | -1 | 3.767 | 5 | 5 | -15 |
| | Insula | 4 | 1 | 3.737 | -35 | 5 | 20 |
| 227 | Precentral Gyrus | 6 | 1 | 4.494 | -40 | 0 | 30 |
| | Insula | 17 | 1 | 4.402 | -35 | 5 | 20 |
| | Inferior Frontal Gyrus | 7 | 1 | 4.181 | -45 | 0 | 25 |
| | Posterior Cingulate | 3 | 1 | 3.643 | -5 | -30 | 25 |
| | Lose-Shift < Lose-Stay | | | | | | |
| 131 | Superior Frontal Gyrus | 21 | -1 | -4.277 | 15 | 55 | 40 |
| | Medial Frontal Gyrus | 28 | 0 | -4.195 | 10 | 50 | 40 |
| | Cingulate Gyrus | 2 | 1 | -3.522 | -10 | 30 | 30 |
| | Win-Stay < Win-Shift | | | | | | |
| 184 | Inferior Parietal Lobule | 5 | -1 | -3.762 | 50 | -60 | 40 |
| | Supramarginal Gyrus | 2 | -1 | -3.703 | 65 | -45 | 35 |

Table 5.1: Structures exhibiting the greatest differences in lose-shift vs lose-stay and win-stay vs win-shift associated current source densities over the time course of feedback processing. For each time point the top four structures exhibiting the greatest differences in activation following are given. Voxels: # of voxels in region activated above threshold ($p < .0005$). LI: left/right (+1/-1) laterality of activation. X,Y,Z: MNI coordinates of maximum activation (max $z$) within structure.

behaviour (Table 5.1). These regions are responsible for spatial processing and localization of the limbs in space, further validating the importance of spatial processing to the win-stay and lose-shift.

### 5.3.3 Source Frequency Analysis

Having validated eLORETA against scalp voltage potentials, we next investigated frequency generators relating to the win-stay and lose-shift. Unlike event-related potentials, oscillatory dynamics can capture unsynchronized activity and consequently

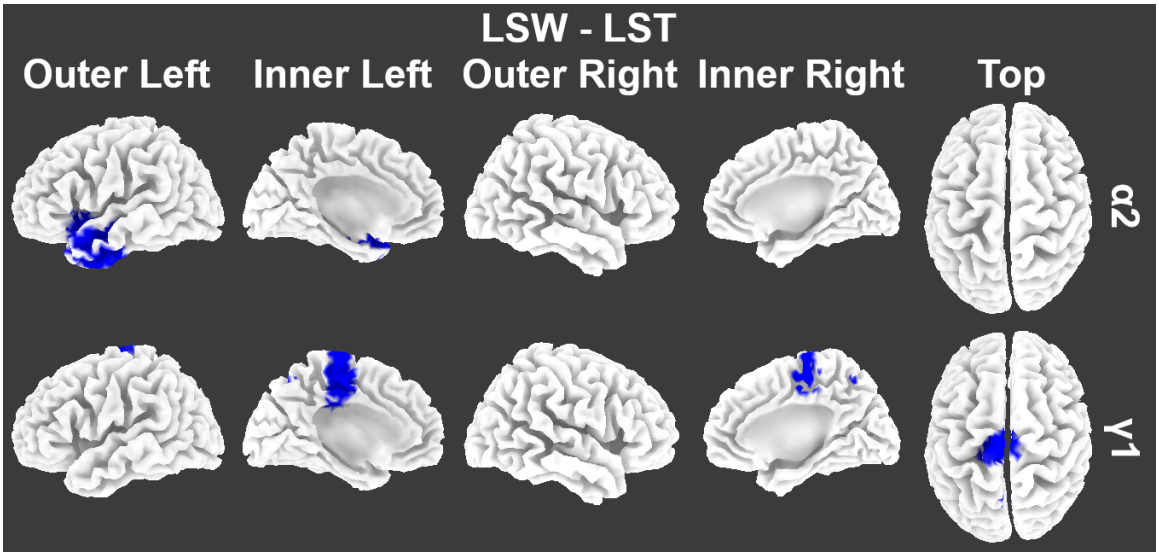


Figure 5.6: Difference in frequency response between LSW-LST outcomes following feedback presentation. Regions highlighted in blue indicate significantly less oscillatory activity elicited preceding a lose-shift response (relative to a lose-stay).

are not restricted to event-related phase resetting of cortical activity [Makeig et al., 2004]. Therefore, localizing the generators of these oscillatory dynamics provides a more informative and biologically relevant description of decision and reward-related activity [Makeig et al., 2002].

Previous studies of reward processing have focused on theta ($\theta$, 4-8 Hz) oscillations generated in the anterior cingulate and prefrontal cortex [Oberg et al., 2011, Cohen et al., 2007]. These oscillations may be linked the integration of place and reward information via hippocampal and striatal theta [van der Meer and Redish, 2011, Cohen, 2011, Herweg et al., 2016]. However, delta ($\delta$, <4 Hz) and gamma ($\gamma$, >30 Hz) oscillations are also worth consideration, being related to reward processing, response selection, and motor inhibition [Başar-Eroglu et al., 1992, Aoki et al., 1999, Gaetz et al., 2013]. Therefore, we calculated differences in the shift vs stay frequency responses over the second following feedback presentation. Lose-shift and

| Time (ms) | Region | Voxels | LI | Max $z$ | X | Y | Z |
|---|---|---|---|---|---|---|---|
| | Lose-Shift < Lose-Stay | | | | | | |
| $\alpha_2$ | Insula | 15 | 1 | -4.386 | -40 | 10 | -10 |
| | Superior Temporal Gyrus | 50 | 1 | -4.329 | -40 | 5 | -15 |
| | Inferior Frontal Gyrus | 18 | 1 | -4.204 | -35 | 10 | -15 |
| | Middle Temporal Gyrus | 27 | 1 | -3.924 | -50 | 0 | 15 |
| $\gamma_1$ | Medial Frontal Gyrus | 47 | .250 | -3.962 | -5 | -20 | 60 |
| | Precentral Gyrus | 5 | 1 | -3.875 | -10 | -20 | 70 |
| | Paracentral Lobule | 16 | .556 | -3.843 | -5 | -20 | 50 |
| | Cingulate Gyrus | 22 | 1 | -3.782 | -10 | -20 | 45 |
| | Win-Stay > Win-Shift | | | | | | |
| $\theta$ | Medial Frontal Gyrus | 48 | .219 | 4.124 | -10 | -10 | 65 |
| | Superior Frontal Gyrus | 7 | .714 | 4.119 | -10 | -10 | 70 |
| | Precuneus | 38 | 1 | 3.954 | -10 | -50 | 60 |
| | Postcentral Gyrus | 17 | .647 | 3.920 | -10 | -55 | 65 |
| $\alpha_1$ | Precuneus | 159 | .119 | 4.720 | 0 | -55 | 65 |
| | Postcentral Gyrus | 32 | -.125 | 4.659 | 10 | -55 | 65 |
| | Paracentral Lobule | 52 | -.304 | 4.551 | 10 | -50 | 65 |
| | Superior Parietal Lobule | 61 | .410 | 4.435 | 15 | -55 | 60 |
| $\alpha_2$ | Superior Parietal Lobule | 49 | .796 | 4.610 | -30 | -60 | 65 |
| | Postcentral Gyrus | 75 | .627 | 4.559 | -25 | -50 | 70 |
| | Precuneus | 38 | .444 | 4.258 | 5 | -55 | 65 |
| | Paracentral Lobule | 50 | .143 | 4.251 | 10 | -50 | 65 |
| $\beta_1$ | Superior Parietal Lobe | 69 | .797 | 5.766 | -25 | -50 | 60 |
| | Postcentral Gyrus | 81 | .802 | 5.737 | -25 | -50 | 65 |
| | Paracentral Lobule | 44 | .588 | 5.700 | -20 | -45 | 55 |
| | Precuneus | 153 | .600 | 5.679 | -25 | -50 | 50 |
| $\beta_2$ | Precuneus | 41 | .444 | 4.078 | -20 | -55 | 45 |
| | Superior Parietal Lobule | 2 | 1 | 3.743 | -25 | -55 | 45 |

Table 5.2: Structures exhibiting the greatest differences in lose-shift vs lose-stay and win-stay vs win-shift associated oscillatory activity measured over the 1000 ms following feedback presentation. For each frequency band, the top four regions exhibiting activity at a significance threshold of $p < .0005$ are detailed.

win-stay associated activity was assessed in the $\delta$ (1-4 Hz), $\theta$ (4-8 Hz), $\alpha_1$ (8-10.5 Hz), $\alpha_2$ (10.5-13 Hz), $\beta_1$ (13-20 Hz), $\beta_2$ (20-30 Hz), $\gamma_1$ (30-50 Hz), and $\gamma_2$ (50-80 Hz) bands.

As seen in Table 5.2 and Figure 5.6 lose-shifting was preceded by reductions in $\alpha_2$

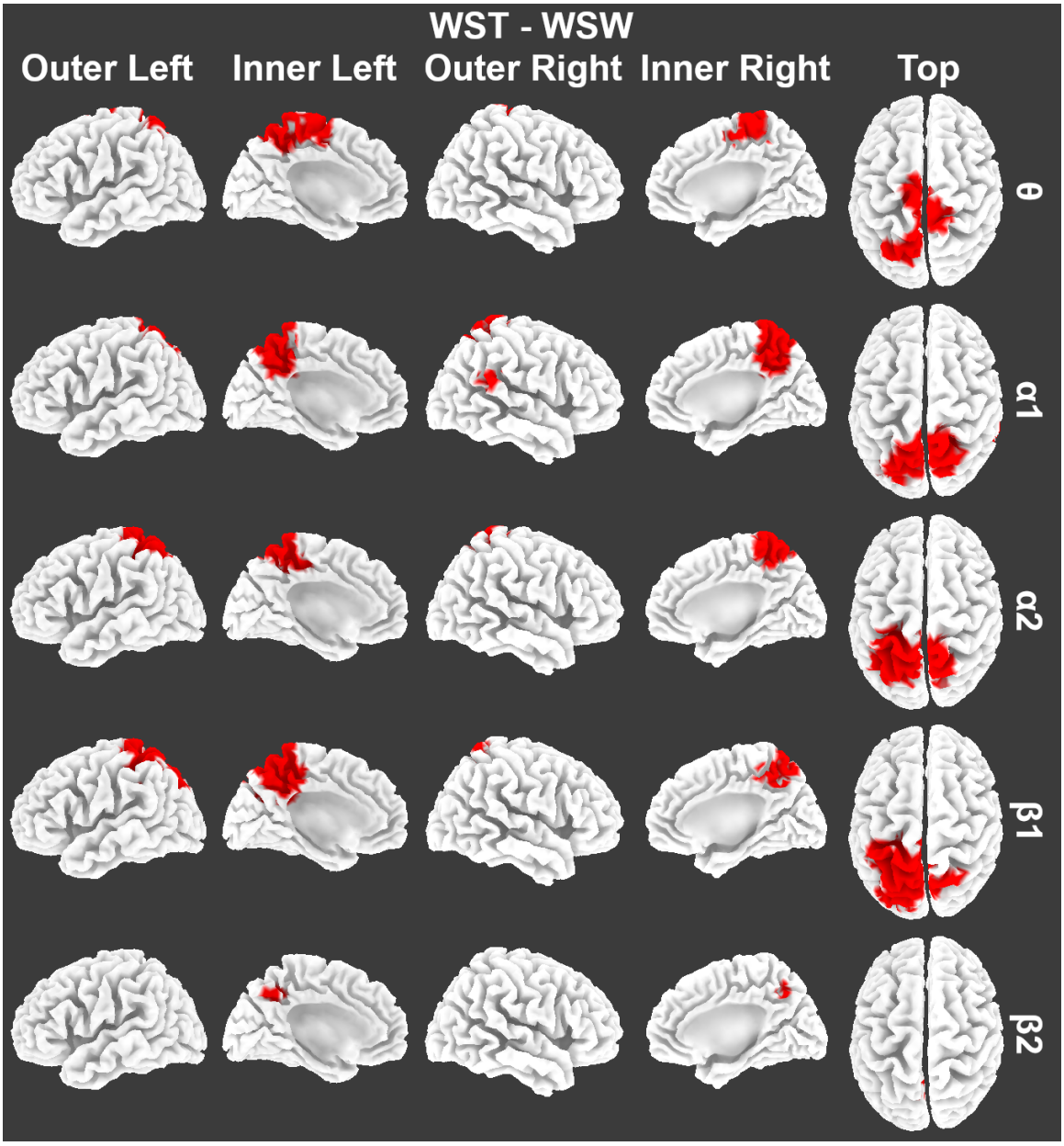


Figure 5.7: Difference in oscillatory activity following outcomes that lead to a win-stay vs a win-shift response measured over the 1000 ms following feedback presentation. In all cases win-stay responses were preceded by greater oscillatory activity relative to win-shifting, as depicted in red.

and $\gamma_1$ oscillations. Alpha activity was localized within the left insula, inferior frontal gyrus, superior and middle temporal gyri, regions active during decision-making tasks. The insula is associated with loss aversion and risk-taking [Markett et al., 2016, Weller et al., 2009], the left IFG with choice biases [Reckless et al., 2014], the left STG with decision effort and lose-shift responding [Yang et al., 2016, Paulus et al., 2002a], and both the insula and superior temporal gyrus with action selection and shift responding [Paulus et al., 2005]. Reduced alpha in the left insula, superior temporal gyrus, and middle frontal gyrus is also associated with greater attention [Clemente et al., 2014] and increased emotional valence in the left hemisphere [Reuderink et al., 2013].

Lose-shifting was also associated with reduced $\gamma_1$ in the bilateral medial frontal gyrus, paracentral lobule, middle cingulate gyrus, and left primary motor cortex (precentral gyrus). Paulus et al. [2002a] found that activation of these regions, particularly the cingulate, medial frontal gyrus, and precentral gyrus is associated with lose-shifting. The paracentral lobule and precentral gyrus also contain motor and sensory innervations, suggesting value of the lose-shift is partially represented by sensorimotor systems. In addition, reduced low gamma is associated with processing of negative outcomes [HajiHosseini et al., 2012], visuospatial attention [Gruber et al., 1999], arousal [Lakatos et al., 2004], and top-down modulation of decision processes [Castelhano et al., 2014]. Finally, gamma oscillations in the medial frontal gyrus differ as a function of D1 and D4 dopamine receptor genotype [Demiralp et al., 2007], which in turn drive lose-shift responding [Floresco, 2013, Onge et al., 2011].

Unlike what we found concerning event-related potentials, win-stay responding was associated with large increases in oscillatory activity during the 1000 ms following reward presentation, particularly in the $\beta_1$ band. This result implies that

activation of cortical circuits responsible for the win-stay was not phase-locked to rewards. As seen in Figure 5.7 this activity was centered over the somatosensory and visuospatial cortices. The postcentral gyrus, superior parietal, and paracentral lobes were consistently active in the alpha and beta range (Table 5.2). While the functional roles of the precuneus in decision making are poorly understood, it is known to be connected with the putamen, caudate nucleus, thalamus, and associative cortices [Cavanna and Trimble, 2006]. Sensory inputs from the hands project to both the putamen and the superior parietal lobe (SPL) as a whole [Zaehle et al., 2007, Cunningham et al., 2017]. Consequently, these regions are associated with spatial processing of visual targets and motor coordination. Activity in the putamen and SPL is also tied to egocentric [Gramann et al., 2006] and allocentric Frings et al. [2006] coding of space. The significant activation of the postcentral gyrus and paracentral lobule, which are components of the somatosensory and premotor cortices, further suggests that spatial-motor processing is important to the win-stay response.

Overall, these results show that oscillatory activity associated with win-stay responses are found in frequency bands and anatomical structures distinct from those activated prior to lose-shift responding. However, both behaviours were associated with activity in regions involved with spatial-motor processing, highlighting the connection between abstract representations of choice and their associated motor action.

### 5.3.4 Choice-Related Frequency Response

The previous results indicate sensorimotor response strategies rely on value updating by spatial-motor circuitry. Motor systems may also actively influence the value of win-stay or lose-shifting during response execution, rather than being a passive output for decision circuits. We hypothesize that the context in which motor responses are performed influence their associated brain activity. In particular, for the same response (left/right hand) following the same outcome (win/loss), decision-related activity will differ based on whether that choice is a shift or stay. Also, if context-specific activity directly corresponds to the hand being used (i.e., left hand & right motor cortex) then sensorimotor responses are coded in egocentric spatial coordinates. Oscillatory activity during the -100 to 500 ms interval surrounding execution of win-stay and lose-shift responses were compared against win-shift and lose-stay activity. Separate $t$ tests were applied to responses made with the left and right hands.

Execution of lose-shift and win-stay responses were both accompanied by increased oscillatory activity relative to lose-stay or win-shift responding. As seen in Figure 5.8 and Table 5.3, lose-shifting was not associated with changes in motor activity. Instead, delta oscillations were present in the frontal gyrus and anterior cingulate ipsilateral to the hand being used. These regions are associated with modulation of motor activity [Fonken et al., 2016] and inhibition of ongoing actions [Braver et al., 2001, Swick et al., 2008, Sharp et al., 2010]. For example, unilateral stimulation and lesions of the anterior cingulate induces respective movement and neglect of contralateral limbs Watson et al. [1973], Luppino et al. [1991]. Delta activity in these regions also is associated with attentional shifts and sensorimotor inhibition during concentration [Harmony, 2013, de Vries et al., 2018]. Therefore, lose-shift responding may require

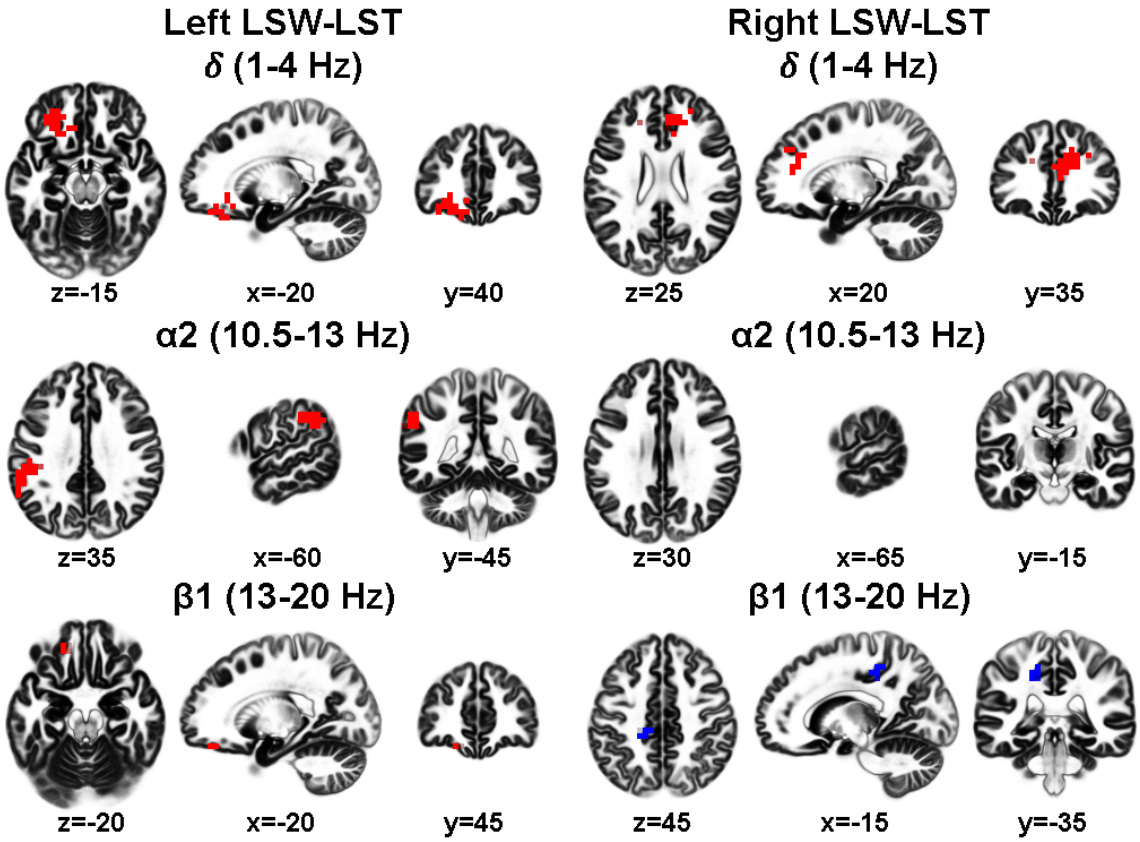


Figure 5.8: Difference in frequency response between LSW-LST activity for choices made with the left and right hands during execution of motor actions. Regions highlighted in red indicate significantly greater oscillatory activity elicited during lose-shift responses.

active inhibition of the previous motor response.

Win-stay responding was strongly associated with activation of the motor (precentral) and somatosensory (postcentral) cortices, particularly in the beta band. As seen in Figure 5.9 and Table 5.4, this activity was restricted to the cortex contralateral to the hand used. Somatomotor beta is associated with planning and execution of motor actions [Baker, 2007, Zaepffel et al., 2013], particularly in the contralateral hemisphere. As with reward processing, greater activation was also present in the

| Frequency | Region | Voxels | LI | Max $z$ | X | Y | Z |
|---|---|---|---|---|---|---|---|
| | Left Lose-Shift > Lose-Stay | | | | | | |
| $\delta$ | Middle Frontal Gyrus | 17 | 1 | 4.193 | -20 | 40 | -15 |
| | Anterior Cingulate | 2 | 1 | 4.092 | -10 | 35 | -10 |
| | Inferior Frontal Gyrus | 22 | 1 | 4.073 | -25 | 35 | -10 |
| | Superior Frontal Gyrus | 4 | 1 | 4.044 | -20 | 40 | -20 |
| $\alpha_2$ | Supramarginal Gyrus | 8 | 1 | 4.213 | -60 | -45 | 35 |
| | Inferior Parietal Lobule | 15 | 1 | 4.024 | -60 | -45 | 40 |
| | Postcentral Gyrus | 3 | 1 | 3.712 | -50 | -30 | 35 |
| | Middle Frontal Gyrus | 2 | 1 | 3.522 | -20 | 40 | -15 |
| $\beta_1$ | Superior Frontal Gyrus | 2 | 1 | 3.559 | -20 | 45 | -20 |
| | Right Lose-Shift > Lose-Stay | | | | | | |
| $\delta$ | Medial Frontal Gyrus | 16 | -1 | 4.740 | 20 | 35 | 25 |
| | Anterior Cingulate | 20 | -1 | 4.437 | 15 | 35 | 20 |
| | Insula | 2 | -1 | 4.138 | 30 | 20 | 15 |
| | Superior Frontal Gyrus | 6 | -1 | 3.871 | 20 | 40 | 35 |
| | Right Lose-Shift < Lose-Stay | | | | | | |
| $\beta_1$ | Cingulate Gyrus | 5 | 1 | -3.665 | -15 | -35 | 45 |
| | Paracentral Lobule | 2 | 1 | -3.606 | -15 | -40 | 50 |

Table 5.3: Difference in lose-shift vs lose-stay oscillatory power during motor responses executed with the left or right hands.

superior and inferior parietal lobes, two regions that are necessary for motor coordination and spatial processing of visual targets [Zaehle et al., 2007]. Therefore, motor and visuospatial systems directly encode the greater value associated with win-stay responses.

During both lose-shift and win-stay responding, use of the non-dominant hand elicited greater and more widespread cortical activity than did use of the dominant hand. Lose-shifting with the left hand was uniquely associated with $\delta$ activity in the left inferior frontal gyrus and $\alpha_2$ oscillations in the left supramarginal gyrus and inferior parietal lobule. The inferior frontal gyrus is associated with habitual response inhibition and attentional control [Swick et al., 2008, Hampshire et al., 2010]. The

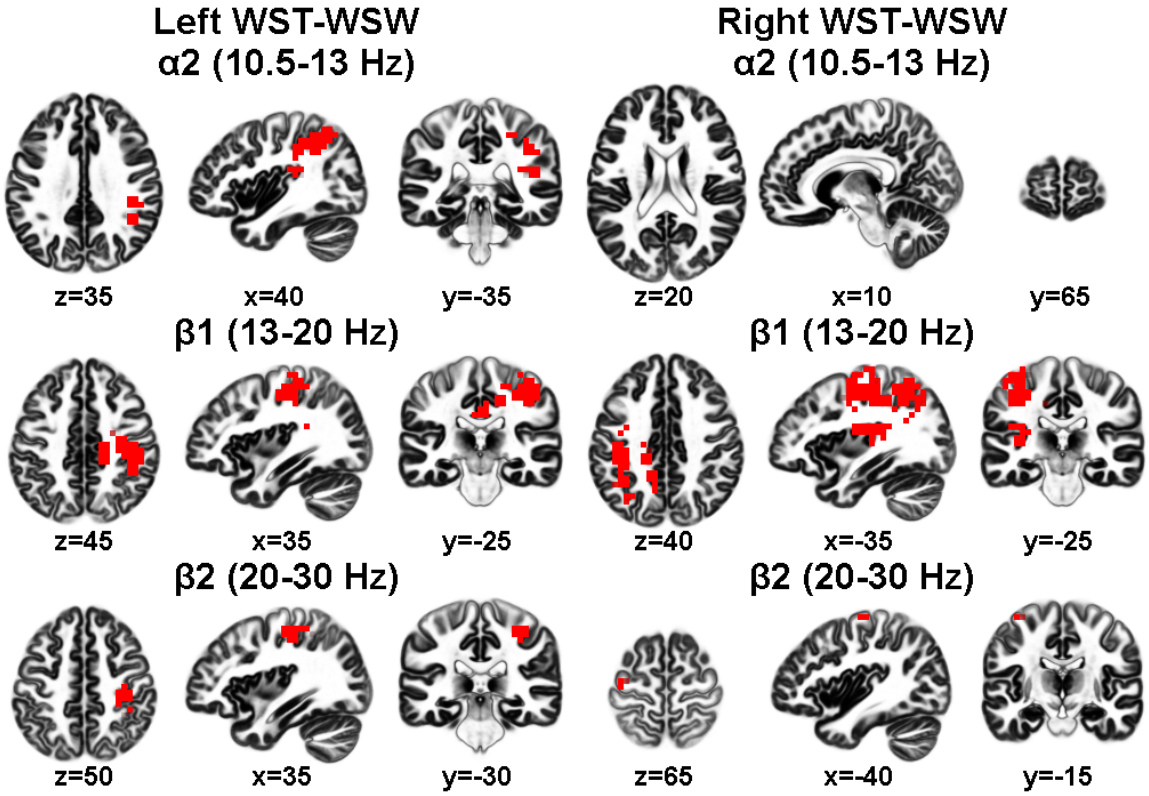


Figure 5.9: Difference in frequency response between win-stay vs win-shift activity for choices made with the left and right hands during execution of motor actions. Regions highlighted in red indicate significantly ($p < .0005$) greater activity elicited during win-stay responding.

supramarginal gyrus, a part of the somatosensory association cortex, is necessary for localization of the hands in space [Ben-Shabat et al., 2015]. Finally, the left and right inferior parietal lobes are associated with motor imagery, attention to the contralateral visual fields, and initiating motor actions towards the right or left visual fields [Mattingley et al., 1998, Buxbaum et al., 2007]. Win-stay responses made with the left hand elicited $\alpha_1$ - $\beta_2$ activity in the right inferior parietal lobule and cingulate gyrus. However, right-handed win-stay responses uniquely elicited greater activity in the left insula, which is known to mediate the default bias, is necessary

| Frequency | Region | Voxels | LI | Max $z$ | X | Y | Z |
|---|---|---|---|---|---|---|---|
| | Left Win-Stay > Win-Shift | | | | | | |
| $\alpha_1$ | Inferior Parietal Lobule | 3 | -1 | 3.673 | 50 | -35 | 35 |
| | Superior Parietal Lobule | 5 | -1 | 3.588 | 35 | -60 | 55 |
| | Precuneus | 1 | -1 | 3.511 | 30 | -55 | 50 |
| $\alpha_2$ | Inferior Parietal Lobule | 38 | -1 | 4.497 | 40 | -35 | 35 |
| | Supramarginal Gyrus | 4 | -1 | 4.334 | 40 | -45 | 35 |
| | Precuneus | 3 | -1 | 4.285 | 30 | -50 | 50 |
| | Superior Parietal Lobule | 13 | -1 | 4.145 | 35 | -55 | 50 |
| $\beta_1$ | Postcentral Gyrus | 46 | -1 | 5.211 | 35 | -25 | 45 |
| | Precentral Gyrus | 39 | -1 | 5.049 | 30 | -20 | 45 |
| | Inferior Parietal Lobule | 22 | -1 | 4.610 | 40 | -35 | 40 |
| | Cingulate Gyrus | 46 | -.737 | 4.570 | 20 | -25 | 40 |
| $\beta_2$ | Postcentral Gyrus141 | -1 | 3.927 | 35 | -30 | 50 | |
| | Precentral Gyrus | 6 | -1 | 3.806 | 30 | -30 | 50 |
| | Inferior Parietal Lobule | 1 | -1 | 3.551 | 40 | -40 | 50 |
| | Right Win-Stay > Win-Shift | | | | | | |
| $\beta_1$ | Postcentral Gyrus | 45 | 1 | 5.051 | -35 | -25 | 40 |
| | Insula | 31 | 1 | 4.998 | -35 | -25 | 20 |
| | Precentral Gyrus | 49 | 1 | 4.765 | -35 | -20 | 45 |
| | Superior Parietal Lobule | 58 | 1 | 4.635 | -25 | -70 | 55 |
| $\beta_2$ | Precentral Gyrus | 32 | 1 | 3.694 | -40 | -15 | 65 |

Table 5.4: Difference in win-stay vs win-shift oscillatory power during motor responses executed with the left or right hands.

for switching between tasks, and is associated with rates of win-stay and lose-shift responding [Paulus et al., 2002b, Yu et al., 2010, Varjačić et al., 2018]. Given that right-handed subjects demonstrate a strong preference to make decisions with their dominant hand (Chapters 2 - 4), switching from this habitual response and using the left hand is accompanied by more widespread activity.

Overall, the cortical activity elicited during execution of motor responses provides

further evidence that the win-stay and lose-shift behaviours are calculated in egocentric and allocentric spatial coordinates respectively. Win-stay responses were associated with activity in the motor and somatosensory cortices that in turn project to the putamen, circuits necessary for egocentric spatial processing [Kesner and DiMattia, 1987, Palencia and Ragozzino, 2005]. Conversely, lose-shift behaviour elicited greater activity in associative and limbic circuits that project to the nucleus accumbens and caudate nucleus, which are involved in allocentric spatial processing [De Leonibus et al., 2005, Ragozzino et al., 2002, Postle and D'Esposito, 2003, Possin et al., 2017].

## 5.4 Cannabis use and decision-making

### 5.4.1 Greater lose-shift inhibition in male cannabis users

In Chapters 2 - 4 we demonstrated that lose-shift responding is elevated in female cannabis users while male cannabis users are better able to inhibit sensorimotor response strategies. These opposing behavioural trajectories may be due to different usage patterns among males and females. For example, Wong et al. [2017b] demonstrated that acute amphetamine exposure results in elevated striatal dopamine, attenuating loss aversion and lose-shift behaviour. However, chronic use inhibits prefrontal control over habitual responding, resulting in increased lose-shift behaviour. To test this hypothesis and further replicate our findings in Chapters 2 - 4, we tested whether decision-making was influenced by cannabis use and sex.

For each subject, the proportion of winning trials, lose-shift responses, win-stay responses, log-transformed decision times, and response entropy were averaged over the 400-trial session. The effects of cannabis use and biological sex on behaviour was

then tested using 2 × 2 ANOVAs. Type-III sums-of-squares and zero-sum contrasts were used to control for unequal subject numbers in each group. Differences between cannabis users and controls among the male and female sub-populations were assessed Welch's unequal variance $t$ tests.

As seen in Figure 5.10.A, cannabis users exhibited significantly better task performance than did controls [$F_{1,67} = 6.707$, p = .012]. Moreover, the effect of sex and the sex × cannabis use interaction were not significant ($p > .403$ in both cases), indicating the effects of cannabis use were consistent between males and females. $T$ tests also indicated that female cannabis users significantly differed from controls [$t(31.862) = 2.197$, p = .035, $d = .692$] while male users and controls did not significantly differ [$t(27.439) = 1.602$, p = .121, $d = .574$]. Therefore, cannabis use was associated with improved performance in the present task.

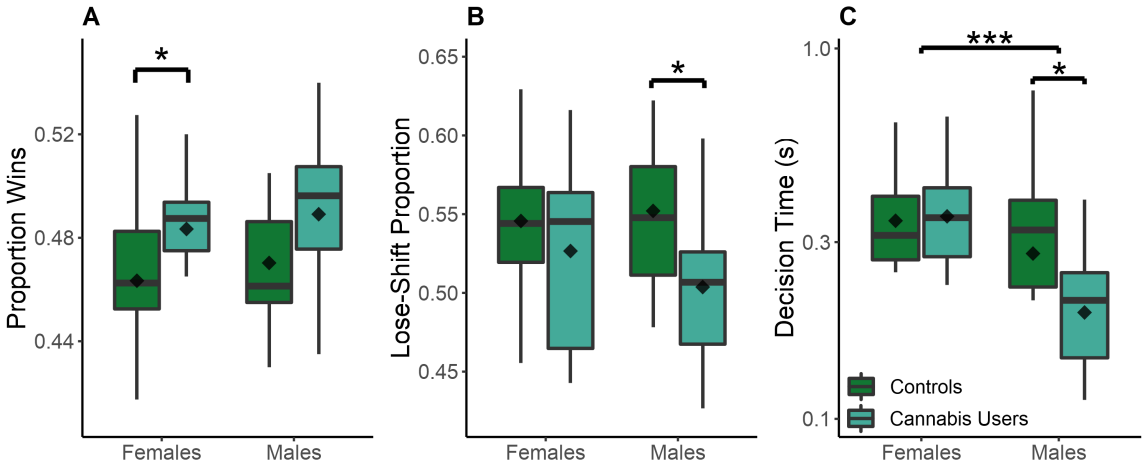


Figure 5.10: Effects of sex and cannabis use on proportion of wins experienced (A), lose-shift behaviour (B), and log decision times (C).

Cannabis users also exhibited improved suppression of habitual lose-shift behaviour [$F_{1,67} = 6.729$, p = .012]. As with wins, lose-shift responding did not significantly vary with sex or the sex × cannabis use interaction ($p > .264$). As seen in Figure 5.10.B, male cannabis users lose-shifted more than male controls [$t(28.339) =$ -2.586, p = .015, $d =$ -.919] while the difference between female cannabis users and controls was not significant [$t(22.145) =$ -.993, p = .331, $d =$ -.355]. Rates of win-stay behaviour ($p > .355$) and response entropy ($p > .129$) did not significantly vary with cannabis use or sex. However, men exhibited much faster decision times than did women [$F_{1,67} = 14.135$, p < .001], primarily because male cannabis users were much faster than controls [$t(27.05) = 2.341$, p = .027, $d = .776$]. However, the main effects of cannabis use [$F_{1,67} = 2.529$, p = .116], and the sex × cannabis interaction [$F_{1,67} =$ 3.412, p = .069] fell short of significance (Fig. 5.10.C).

The fact that cannabis users out-performed controls at the Matching Pennies task and were better able to suppress lose-shift behaviour bears in direct contrast with Chapters 2 - 4. Previously we found that only male cannabis users exhibited improved task performance, while females lose-shifted more and were less random in their choices (which resulted in poorer performance). This discrepancy may be due to differences in task presentation. In the present experiment participants made quick decisions using a keyboard. The experiments in Chapters 2 - 4 used a touch-screen, resulting in much slower decisions. However, faster decision times are known to result in a greater tendency towards lose-shift responding [Ivan et al., 2018] rather than a reduction, indicating task presentation cannot account for differences in behaviour.

A more likely explanation for this discrepancy is the dates during which the experiments were conducted. In Canada, recreational cannabis use was legalized on

October 17, 2018. The experiments in Chapters 2 - 4 were conducted prior to this date. Conversely, the present experiment was conducted during the year following legalization. As such, the present study had a much greater likelihood of including new cannabis users who had not yet developed a pattern of chronic cannabis use. Wong et al. [2017b] demonstrated that chronic and acute amphetamine use have opposite effects on lose-shift behaviour, similar to our findings. Moreover, while females are much less likely to use cannabis than males, females who do use it are much more likely to develop cannabis use disorder [Calakos et al., 2017] and suffer cognitive deficits and mood disorders as a result [Crane et al., 2013]. Therefore, we hypothesized the present study consisted of a greater proportion of new cannabis users that had not yet developed a pattern of chronic use, relative to chapters 2 - 4.

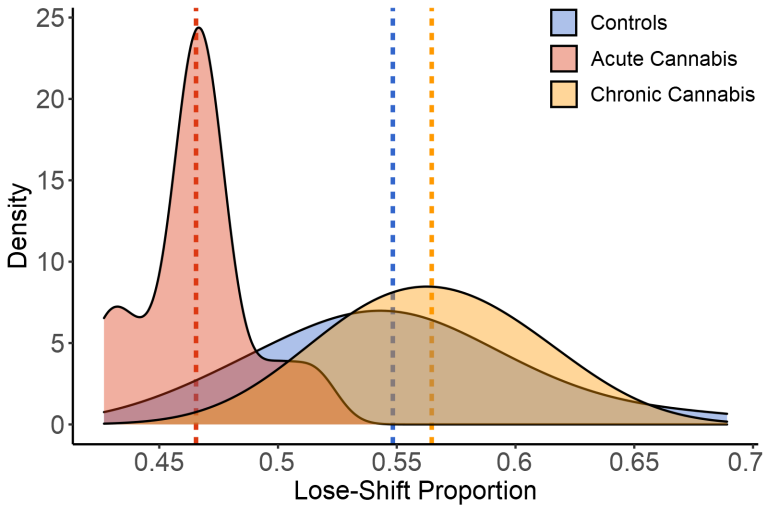


Figure 5.11: Density plots of lose-shift rates among controls, acute, and chronic cannabis users. Dashed vertical lines indicate the mean of each group.

Unfortunately we did not ask participants when they started using cannabis, but only their frequency of usage within the last three months. However, if our population consisted of acute and chronic users a corresponding pattern should exist in their

behaviour. In particular, they should consist of sub-populations exhibiting low and high rates of lose-shift responding. Therefore, we tested whether lose-shifting among cannabis users followed a bimodal distribution using the R MCLUST package [Fraley and Raftery, 1999]. Likelihood ratio tests indicated that a bimodal distribution fit lose-shifting rates better than a unimodal distribution [$\chi^2(1, N = 28) = 6.828$, $p = .046$]. Consequently, we partitioned cannabis users into subgroups exhibiting lose-shift behaviour consistent with chronic and acute usage. As seen in Figure 5.11, control subjects lose-shifted on about 55% of trials. Cannabis users consisted of acute and chronic use sub-populations that lose-shifted less and more often than controls respectively. Moreover, those we labelled as acute users had a mean ASSIST cannabis score of 10.71 (SD = 8.08) while chronic users had a greater mean score of 16.36 (SD = 11.86), indicating lose-shifting behaviour was an effective indicator of cannabis use severity.

To ensure these two sub-populations were unique to the present study, we re-analyzed lose-shift behaviour from Chapters 2 and 4 using the MCLUST package. Analysis of nineteen cannabis users from Chapter 2 indicated unimodal and bimodal distributions fit the lose-shift behaviour equally well on both days one [$\chi^2(1, N = 19) = .020$, $p = .823$] and two [$\chi^2(1, N = 19) = .827$, $p = .702$] of the experiment. Similar analyses of lose-shifting responses in the Chapter 4 experiment, using only trials when no local or global changes in choice position were present, also indicated a unimodal distribution adequately fit the data [$\chi^2(1, N = 37) = -.008$, $p = .958$]. Therefore, subject behaviour and cannabis use severity in the present study supports our hypothesis that the majority of subjects were relatively new cannabis users, due to legalization in Canada. Acute cannabis use is associated with increased striatal

dopamine, lessening the adverse effects of losses [Bossong et al., 2009].

**Co-morbidities between cannabis use, cognitive ability, & mental health** While cannabis use is consistently associated with altered decision-making, it may not necessarily be the cause of this behaviour. For example, it may exhibit co-morbidities that could be the true cause for behavioural differences between cannabis users and controls. Cannabis use is known to coincide with higher rates of depression and anxiety in young adults [Patton et al., 2002] that are also associated with altered striatal function [Malone Jr et al., 2009, Marchand and Yurgelun-Todd, 2010]. Other confounding factors such as tobacco use and socio-economic status are also associated with cannabis use [Rogeberg, 2013, Mokrysz et al., 2016]. For example, lower socio-economic status in childhood and adolescence coincides with a increased rate of cannabis use and dependence Daniel et al. [2009], Legleye et al. [2012].

However, there is evidence that cannabis use can cause many of these outcomes. For example, while cannabis use in young adulthood predicts anxiety and depression later in life, the reverse is not true of early onset anxiety or depression [Hayatbakhsh et al., 2007]. The association between early cannabis use and mental health later in life is also found in twins where social, genetic, and economic factors are controlled for [Lynskey et al., 2004, Smolkina et al., 2017]. Cannabis use is also associated with cognitive impairment, even after controlling for educational background [Meier et al., 2012]. This cannabis-induced impairment results in reduced hippocampal volumes and abnormalities in the temporal, prefrontal, and cerebellar regions [Lubman et al., 2015]. Therefore, it is worth investigating cannabis use as a causal factor responsible for altered decision-making, mental health outcomes, and changes in cognitive ability. However, the same is not true of ADHD, as childhood ADHD is known to predict

future cannabis use [Bidwell et al., 2014], while cannabis is a popular form of self medication for ADHD [Mitchell et al., 2016].

Therefore, we tested whether cannabis use was associated with variation in mental-health outcomes and cognitive ability among the present population. Subjects completed the depression, anxiety, and stress scales (DASS) [Lovibond and Lovibond, 1996], the revised obsessive-compulsive inventory (OCI-R) [Hajcak et al., 2004], and the adult ADHD self-report scale (ASRS) v1.1 [Kessler et al., 2005]. Cognitive ability was assessed using the Wechsler Adult Intelligence Scale [WAIS-IV Drozdick et al., 2012]. In particular, we measured verbal ability, memory capacity, processing speed, visuospatial reasoning, and inductive logic using the VCI, WMI, PSI, and PRI indices of intelligence. The effects of sex and cannabis use on cognitive and mental-health outcomes were tested using 2 × 2 ANOVAs.

As seen in Figure 5.12.A-B, verbal comprehension ($p > .380$ in all cases) and perceptual reasoning ($p > .057$) scores did not significantly vary with sex, cannabis-use, or the sex × cannabis interaction. However, when females were assessed separately cannabis-users exhibited lower perceptual reasoning scores than controls [$t(17.599) =$ -2.697, p = .015, $d =$ -1.051]. Cannabis use and the sex × cannabis interaction also did not influence working memory scores (Fig. 5.12.C; $p > .151$). However, males exhibited higher scores than females [$F_{1,65} = 11.093$, p = .001]. Conversely, females exhibited faster processing speeds than males [$F_{1,65} = 7.128$, p = .010], as seen in figure 5.12.D. Cannabis users also exhibited slower processing speeds than controls [$F_{1,65} = 4.315$, $p = .042$], particularly among female cannabis users relative to controls [$t(30.993) =$ -2.518, p = .017, $d =$ -.801]. However no sex × cannabis interaction was present ($p = .630$).

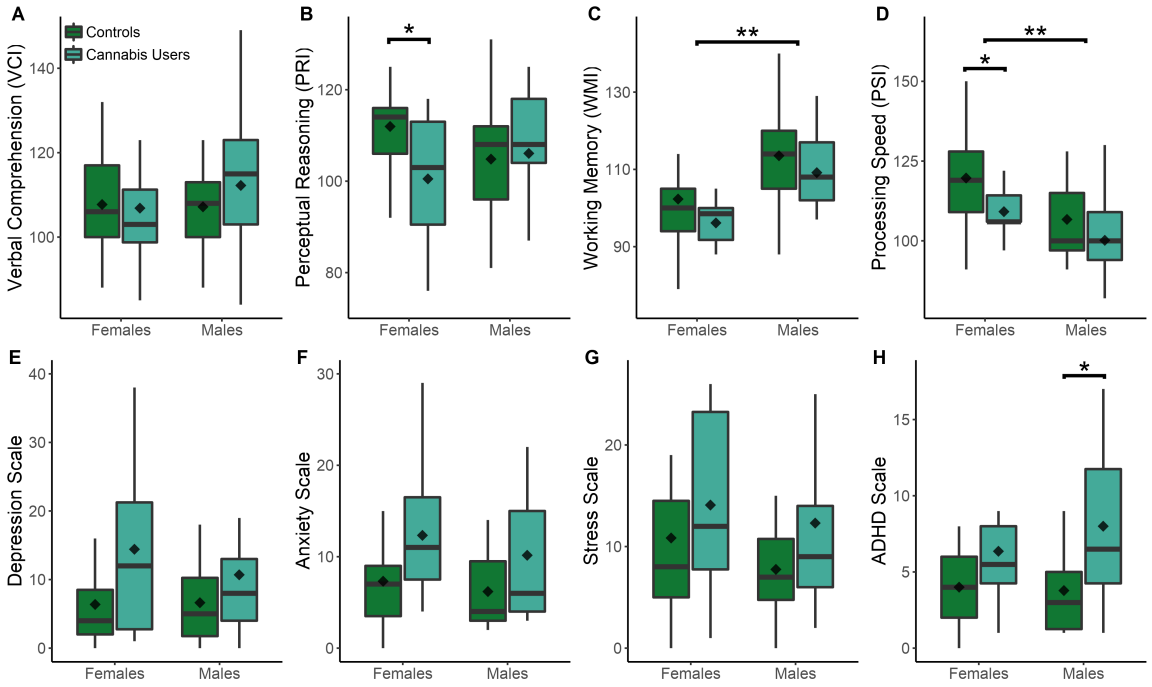


Figure 5.12: Effects of sex and cannabis use on IQ scores (A-D) and scales of depression, anxiety, stress, and ADHD (E-H).

Cannabis use was also associated with negative mental health outcomes (Figs. 5.12.E-H). Cannabis users reported a greater severity of depressive symptomatology [$F_{1,56} = 7.145$, $p = .010$, $d = .699$], anxiety [$F_{1,56} = 6.552$, p = .013, $d = .664$], and ADHD [$F_{1,67} = 14.374$, p < .001, $d = .924$] relative to controls. In each case, the effects of sex and the sex × cannabis interaction were not significant ($p > .094$). Moreover, symptoms of stress disorders ($p > .057$) and obsessive-compulsive disorder ($p > .094$) were not significantly associated with cannabis-use, sex, or the interaction.

Overall, these results indicate cannabis use corresponds to negative lower perceptual reasoning abilities and processing speed, particularly in females. This sexual dimorphism is also found in non-human animals; for example, females rats sensitized to THC exhibit impaired spatial learning compared to males [Cha et al., 2007].

Sex hormones are known to modulate the structure and sensitivity of the female endocannabinoid system, rendering females more susceptible to the adverse effects of cannabis Struik et al. [2018]. However, male and female cannabis users were equally susceptible to the negative effects of cannabis use on mental-health. Despite cannabis users exhibiting increased lose-shift responding, decreased cognitive ability, and negative mental-health outcomes, these three outcomes are not related. Symptoms of depression, anxiety, stress, ADHD, and OCD were uncorrelated with all intelligence scores ($p > .200$ in all cases), rates of lose-shifting ($p > .115$), win-stay responding ($p > .255$), and response entropy ($p > .562$). Intelligence scores were also not significantly correlated with lose-shift behaviour ($p > .058$), win-stay behaviour ($p > .396$), or response entropy ($p > .072$). However, OCD scores were negatively correlated with the verbal abilities (VCI; $r(58) = -.292$, $p = .024$). Win-stay responding was also negatively correlated with ADHD symptom severity ($r(69) = -.243$, $p = .041$), a finding supported by Abouzari et al. [2015]. Overall though, we find the effects of cannabis-use on lose-shift behaviour are not due to differences in mental health.

**Cannabis use affects feedback-related processing** As demonstrated in Figure 5.2, losses are followed by reduced scalp voltage potentials over Cz relative to wins. Consequently, highly salient losses result in more negative FRN (loss-win) amplitudes [Pfabigan et al., 2011]. We have demonstrated that cannabis use in the present task coincides with reduced lose-shifting, a sign of blunted loss aversion in the midbrain and downstream targets including the striatum and anterior cingulate [Tanda et al., 1997, Wong et al., 2017a]. Reduced error signaling in the midbrain should result in a blunted (i.e., less negative) feedback-related negativity [Holroyd et al., 2003]. Therefore, we tested the effects of sex and cannabis use on eFRN, FRN, P3a, and

P3b amplitudes measured at channel Cz (loss-win difference) 2 × 2 ANOVAs.

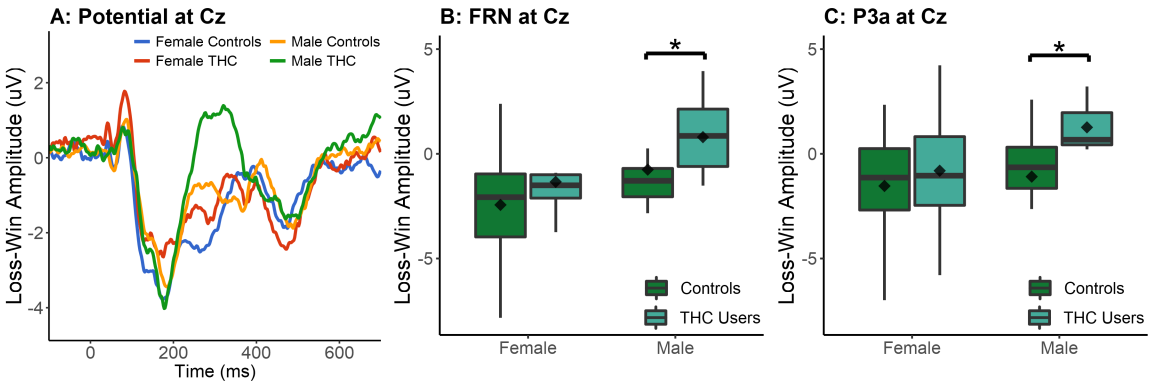


Figure 5.13: A: effects of sex and cannabis use on loss-win scalp amplitudes at channel Cz during the 700 ms following reward presentation. B-C: boxplots of loss-win amplitudes in male and female cannabis users and controls, measured during the FRN and P3a neural responses.

As seen in Figure 5.13.A, cannabis use was associated with less negative FRN amplitudes in men [$t(28.571)$=2.075, $p$=.047, $d$=.703], while female cannabis users did not significantly differ from controls ($p$ = .146). Consequently, the effect of cannabis use on the FRN [$F_{1,63}$ = 5.641, $p$ = .021] was significant. FRN amplitudes were also greater in men relative to women [$F_{1,63}$ = 11.703, $p$ = .001], though the sex × cannabis interaction was not significant ($p$ = .687). The loss-win difference in P3a amplitudes was also greater in cannabis users [$F_{1,63}$ = 5.774, $p$ = .019], particularly in males [$t(24.005)$=2.504, $p$=.019, $d$=.930], as seen in Figure 5.13.B. However, the P3a did not significantly vary with sex [$F_{1,63}$ = 3.854, $p$ = .054] or the sex × cannabis use interaction ($p$ = .210). The early-FRN and P3b also did not significantly vary with sex or cannabis use ($p$ > .285 in all cases).

Overall, cannabis users (particularly males) are less punishment-adverse as indicated by their reduced lose-shifting and elevated FRN amplitudes. This change in

neural response is consistent with recent cannabis use, rather than a long-term pattern of chronic use. Moreover, males are less adverse to punishment than females .The effect of sex could be due to differences in hair, skull density, and skin conductance, rather than reward processing. However, given that the mean loss-win differences in FRN and P3a amplitudes were positive for male cannabis users and negative for female users, the observed effects must be due to differences in reward processing.

### 5.4.2 Behavioural Correlates of Reward Processing

We have demonstrated that cannabis users exhibit both reduced lose-shift behaviour and decreased FRN amplitudes. The FRN, as a measure of action value [Holroyd and Coles, 2002] may be directly correlated with rates of lose-shift responding across subjects. Moreover, the eFRN, FRN, P3a, and P3b represent different stages of reward processing that may each be related to different aspects of decision-making. Therefore, we tested the correlations between scalp voltage potentials and behaviour (e.g., lose-shifting, entropy, win-stay responding). Reinforcement learning parameters were also derived for each subject using the *Q-learning with forgetting model* (i.e., $\alpha$, $\beta$, $\kappa_1$, and $\kappa_2$) and compared against ERP amplitudes. For each subject the eFRN, FRN, P3a, P3b were calculated as the average loss-win difference from 187-207, 253-273, 306-326, and 468-488 ms at channel Cz. The LpFRN was averaged over the LSW-LST difference from 216-236 ms. The WpFRN was averaged over the WST-WSW difference from 184-191 ms post-feedback.

As seen in Figure 5.14.A, the FRN was significantly correlated with task performance [$r(65) = .300$, $p = .014$]. Participants who did not respond strongly to losses won more against the computer, due to being more random in their decisions (Fig.

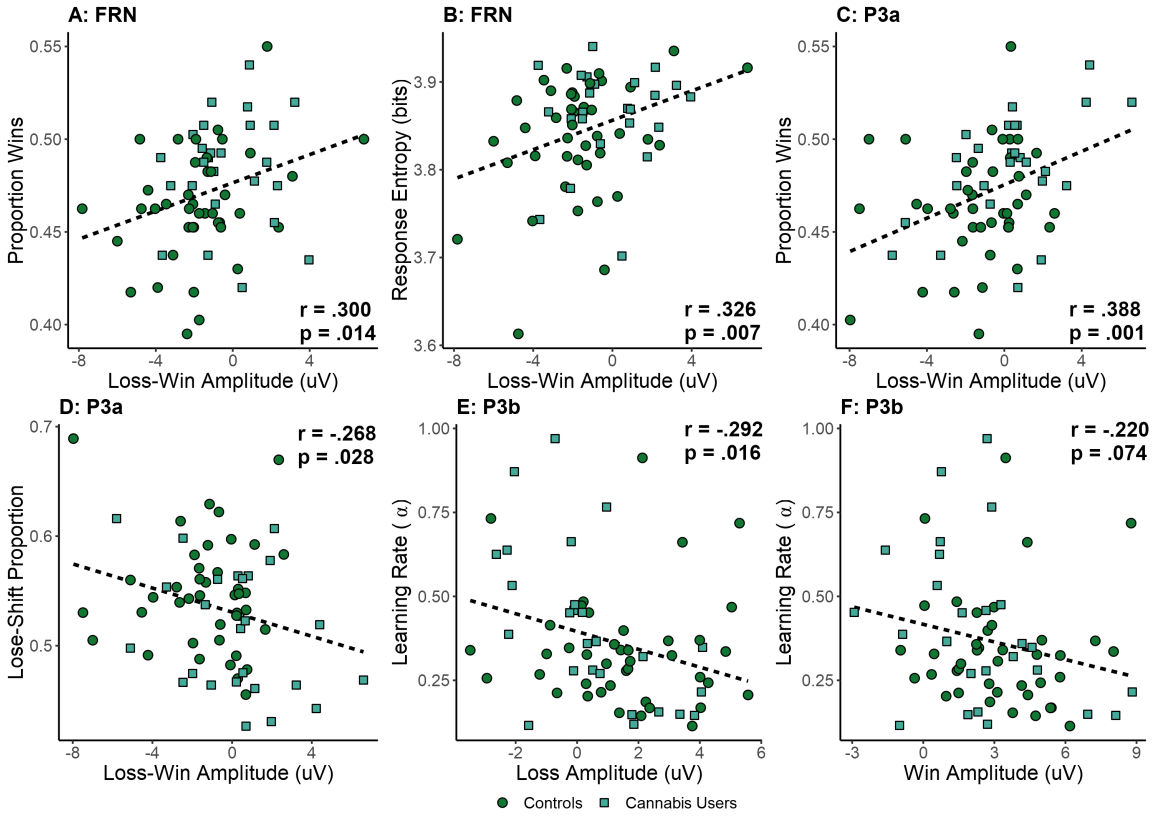


Figure 5.14: Correlations between FRN (N250) and task performance (A), response entropy (B), and decision times (C), and between the P3a and wins (D), lose-shift responding (E), and decision times (F).

5.14.B). Consequently, there was a significant correlation between FRN amplitudes and response entropy [$r(65) = .326$, $p = .007$]. Task performance was also negatively correlated with log decision times [$r(65) = -.291$, $p = .017$]. FRN amplitudes were not related to any other measures of learning or behaviour ($p > .139$ in all cases). Neither was response entropy significantly correlated with the eFRN, P3a, P3b, or LpFRN ($p > .102$ in each case). Consequently, the ability to maintain a mixed-response strategy is uniquely associated with the prediction error signal generated anterior cingulate.

Surprisingly, lose-shift tendencies were not related to LpFRN amplitudes [$r(65)$ = -.030, $p$ = .808]. Instead, the loss-win difference during the P3a was significantly correlated with lose-shifting [$r(65)$ = -.268, $p$ = .028], as seen in Figure 5.14.D. Task performance [$r(65)$ = .388, $p$ = .001] and decision times [$r(65)$ = -.329, $p$ = .007] were also correlated with P3a amplitudes (Fig. 5.14.C). Consequently, the lose-shift behaviour and maintenance of mixed-response strategies are two distinct processes, captured at different times during feedback processing. In particular, the lose-shift is uniquely associated with processes related to model updating, reward valuation, and surprise.

Finally, learning rates were significantly correlated with P3b amplitudes following losses [$r(65)$ = -.292, $p$ = .016] and weakly correlated with those following wins [$r(65)$ = -.220, $p$ = .074], but not related to the loss-win amplitude difference [$r(65)$ = -.060, $p$ = .631]. The P3b reflects memory updating processes while $\alpha$ describes the rate at which new information overrides previously learned action-outcome associations. As seen in Figures 5.3.E-F, subjects exhibiting more positive P3b amplitudes are also less susceptible to overwrite learned action-outcome associations in response to recent rewards or punishments.

Win-stay responding was also significantly associated with P3b amplitudes following losses [$r(65)$ = -.308, $p$ = .011] and weakly correlated with those following wins [$r(65)$ = -.235, $p$ = .055]. While the association between win-stay responding and P3b amplitudes following losses seems odd, it may indicate a more general association with memory-updating processes during the P3b. No other correlations between event-related potentials and task performance, behaviour, and reinforcement learning parameters were present ($p > .139$ in all cases).

**Cannabis use associated with greater activity in the anterior cingulate** Cannabis use is known to increase activity in the dopaminergic midbrain, affecting downstream targets including the striatum and anterior cingulate [Tanda et al., 1997, Wong et al., 2017a]. Consequently, the cannabis-associated change in FRN and P3a amplitudes may be due to oscillatory activity in the anterior cingulate. Therefore, loss-win current source densities were calculated during the second following feedback presentation at 6239 voxels using eLORETA. The frequency response in the $\delta$ (1-4 Hz), $\theta$ (4-8 Hz), $\alpha_1$ (8-10.5 Hz), $\alpha_2$ (10.5-13 Hz), $\beta_1$ (13-20 Hz), $\beta_2$ (20-30 Hz), $\gamma_1$ (30-50 Hz), and $\gamma_2$ (50-80 Hz) bands was calculated for each subject and voxel. The effects of cannabis use and sex on FRN-associated oscillatory activity in each frequency band and voxel were calculated using $2 \times 2$ ANOVAs with a significance threshold of $p < .0005$.

As hypothesized, cannabis use was associated with oscillatory activity in the cingulate gyrus. As seen in Figure 5.15 and Table 5.5, delta activity in the cingulate gyrus, medial frontal gyrus, and superior frontal gyrus was significantly greater in cannabis users. Feedback-related oscillations also differed between the sexes. Males exhibited greater gamma-band activity across the parahippocampal gyrus, uncus, temporal gyrus, and fusiform gyrus (Table 5.5). In every voxel and frequency band the sex $\times$ cannabis use interaction was not significant at the $p < .0005$ level. Moreover, analysis of the LpFRN and WpFRN indicated oscillatory activity did not significantly vary with sex or cannabis use. Finally, analysis of activity during motor execution of win-stay or lose-shift indicated few significant effects. Sex, cannabis-use, and the sex $\times$ cannabis interactions had no significant effects on motor execution of lose-shift

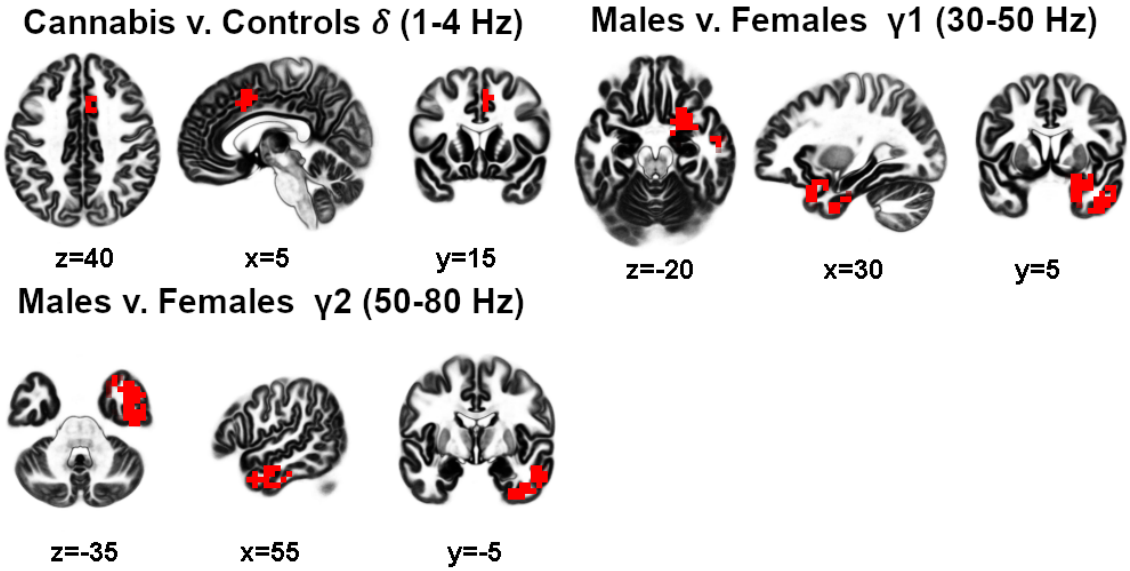


Figure 5.15: Results of ANOVA indicating voxels exhibiting significantly different FRN-associated oscillatory activity between cannabis users and controls in the $\delta$ (1-4 Hz) band and between males and females in the $\gamma_1$ (30-50 Hz) and $\gamma_2$ (50-80 Hz) bands. Red voxels indicate cannabis users or males exhibited significantly greater activity than controls or females at a $p < 0.0005$ threshold.

responses with the right hand or win-stay responses with either hand. However, lose-shifting with the left hand resulted in marginally lower $\theta$ and $\alpha_1$ activity in cannabis users. In particular, theta in two voxels within the left IPL, one voxel in the post-central gyrus, and alpha in two voxels within the right parietal lobe was significantly greater in controls relative to cannabis users. However, these differences are minimal. Therefore, cannabis use is associated with altered feedback processing as a whole, and not differences in motor activity, win-stay, or lose-shift related feedback processing.

In sum, cannabis use is associated with reduced lose-shift responding, negative mental health outcomes, a decline in cognitive ability, attenuated scalp voltage potentials, and increased oscillatory activity in the cingulate and frontal gyri. Men and women also exhibit differences in gamma activity localized within the uncus and

| Frequency | Region | Voxels | LI | Max $z$ | X | Y | Z |
|---|---|---|---|---|---|---|---|
| | THC > Controls | | | | | | |
| $\delta$ | Cingulate Gyrus9 | -1 | 3.663 | 5 | 15 | 40 | |
| | Medial Frontal Gyrus | 3 | -1 | 3.608 | 10 | 10 | 50 |
| | Superior Frontal Gyrus | 1 | -1 | 3.482 | 10 | 10 | 55 |
| | Males > Females | | | | | | |
| $\gamma_1$ | Parahippocampal Gyrus | 6 | -1 | 3.741 | 30 | 5 | -20 |
| | Uncus | 27 | -1 | 3.730 | 30 | 5 | -25 |
| | Inferior Temporal Gyrus | 28 | -1 | 3.717 | 55 | -5 | -35 |
| | Superior Temporal Gyrus | 25 | -1 | 3.713 | 35 | 5 | -25 |
| $\gamma_2$ | Inferior Temporal Gyrus | 31 | -1 | 3.832 | 55 | -5 | -35 |
| | Middle Temporal Gyrus | 38 | -1 | 3.795 | 55 | 5 | -35 |
| | Fusiform Gyrus | 13 | -1 | 3.783 | 55 | -5 | -30 |
| | Uncus | 9 | -1 | 3.605 | 35 | -5 | -40 |

Table 5.5: Results of 2 × 2 ANOVA comparing the effects of cannabis use and sex on eLORETA-derived oscillatory activity associated with the FRN. Top four distinct regions listed. LI = (L-R)/(L+R) number of above threshold ($p < .0005$) voxels. The interaction was not significant for any voxels.

parahippocampal, temporal, and fusiform gyri. This sex-related difference may be due to emotional arousal, which is related to gamma activity [Yang et al., 2018]. Both the temporal lobe and parahippocampal gyrus are known to produce gamma oscillations [Hirai et al., 1999] and activity in the parahippocampal gyrus and uncus vary with the emotional valence of a stimuli. Finally, men and women are known to differ in their emotional response to positive and negative stimuli, which in turn increases gamma activity Yang et al. [2018].

## 5.5 Discussion & Conclusions

In the present study we used electrophysiology to investigate the anatomical correlates of the win-stay and lose-shift behaviours. The results demonstrate that the

win-stay and lose-shift are uniquely associated with activity in associative, sensorimotor and visuospatial systems (relative to the win-shift and lose-stay). Lose-shift responding was preceded by a strong modulation of the feedback-related negativity, a marker of violated reward expectation generated in the anterior cingulate cortex [Holroyd and Coles, 2002, Holroyd et al., 2003]. Source localization using eLORETA confirmed that this loss-predictive FRN (LpFRN) was generated in the anterior cingulate, as well as the parahippocampal gyrus and insula, replicating previous results [Cohen and Ranganath, 2007]. Each of these regions are strongly connected with the nucleus accumbens [Danckert et al., 2011, Chen et al., 2012, Paulus et al., 2002a] and are necessary for representing choices in allocentric spatial coordinates De Leonibus et al. [2005], Bohbot et al. [2015], Parslow et al. [2004]. Gramann et al. [2006] also demonstrated that the anterior cingulate is active during allocentric spatial navigation. Therefore, our results support the hypothesis that the lose-shift is computed in an allocentric frame of reference.

We also demonstrated that lose-shifting is preceded by decreased oscillatory activity following punishment. In particular, $\gamma_1$ oscillations (30-50 Hz) were reduced in the cingulate gyrus, medial frontal gyrus, and motor cortex. Gamma activity is found throughout the cortex, hippocampus, and striatum [Headley and Paré, 2017] and historically has been difficult to study due to overlapping with EMG muscle activity. However, in the motor cortex $\gamma$ is present during movement and learning of motor skills [Nowak et al., 2018] and is associated with improved task performance [Aoki et al., 1999]. Motor $\gamma$ is also highly coherent with activity in the basal ganglia, particularly the ventral striatum [Kalenscher et al., 2010], suggesting the basal ganglia generate motor $\gamma$ through the thalamus [Lee and Jones, 2013]. In the cingulate

gyrus $\gamma$ is associated with conscious awareness of visual targets [Luo et al., 2008] and visual search [Leung and Borst, 1987], particularly during movement Cheron et al. [2016], Uemoto et al. [2017], Calabrò et al. [2017].

Unlike the lose-shift, win-stay behaviour was not preceded by different scalp amplitudes relative to the win-shift. Therefore, the win-stay related processing is supported by different circuits than the lose-shift that produce activity that is not time-locked to reward presentation. However, win-stay responses were preceded by increased oscillatory activity in the $\theta$, $\alpha$, and $\beta$ frequency bands (4-30 Hz). This activity was primarily localized within the somatosensory cortex, paracentral lobule, medial frontal gyrus, and superior parietal lobule (including the precuneus). Gramann et al. [2006] found that the precuneus, paracentral lobule, and medial frontal gyrus are activated during egocentric spatial navigation [also see Vavrečka and Lhotská, 2009, Ruotolo et al., 2019]. The precuneus and SPL are also connected with the putamen, caudate nucleus, and thalamus [Cavanna and Trimble, 2006], structures necessary for egocentric navigation and modulation of motor action. Moreover, sensory inputs from the hands project to both the putamen and precuneus/SPL [Zaehle et al., 2007, Cunningham et al., 2017].

Given that spatial-motor action determines choice value, we hypothesized unique neural activity accompanies execution of win-stay and lose-shift responses. We found that both behaviours resulted in greater oscillatory activity that was specific to the hand being used. For example, win-stay responses executed with the right hand elicited greater activity in the corresponding (i.e., left) motor (precentral) and somatosensory (postcentral) cortices, relative to win-shift responses executed with the

same hand. This activity was dominant at $\beta$ frequencies associated with motor action [Baker, 2007, Zaepffel et al., 2013]. Therefore, choice value (as relates to the win-stay) is directly represented as the motor action needed to make that choice. This direct correspondence between value and motor action supports our hypothesis that the win-stay is coded for in egocentric spatial coordinates. Moreover, the motor and somatosensory cortices provide direct inputs to the putamen [Brasted et al., 1999, Pan et al., 2010, Malach and Graybiel, 1986], which supports egocentric frames of reference [Kesner and DiMattia, 1987, Palencia and Ragozzino, 2005].

Conversely, motor execution of lose-shift (i.e., loss adverse) responses were not coded directly in the motor cortex. Instead, the anterior cingulate and frontal gyrus ipsilateral (i.e., not corresponding) to the hand being used was more active within the $\delta$ band. These regions connect to the nuclues accumbens [Gerfen, 1984, Voorn et al., 2004, Kelley et al., 1982] and caudate nucleus [Khibnik et al., 2014, Fuccillo, 2016], and are associated with spatial processing in an allocentric reference frame [Gramann et al., 2006, Vavrečka and Lhotská, 2009]. Moreover, delta activity in these structures is associated with modulation of motor activity [Fonken et al., 2016] and inhibition of ongoing actions [Braver et al., 2001, Swick et al., 2008, Sharp et al., 2010] in the contralateral limbs Watson et al. [1973], Luppino et al. [1991]. Therefore, adoption of a lose-shift response primarily requires active inhibition of the previous motor response. That the lose-shift was not supported by greater levels of motor activity, but by higher frontal circuits known to projects to the NAc and DMS, supports our hypothesis that it is represented in allocentric spatial coordinates.

Unlike what was found in Chapters 2 - 4, the current experiment found that

cannabis use was associated with improved task performance due to better suppression of lose-shift behaviour in both men and women. The different results may have been due to the fact that the current study, unlike the earlier ones, was conducted immediately following the legalization of THC in Canada. Wong et al. [2017a] demonstrated that acute THC use reduces rates of lose-shift responding through dopaminergic desensitization of the striatum to losses. However, long-term drug sensitization increases lose-shift responding by weakening inhibitory control of the prefrontal cortex over the sensorimotor striatum [Wong et al., 2017b]. Therefore, the sexually dimorphic effects of cannabis use on lose-shift behaviour are not due to functional differences in how cannabis affects the brain. Instead, females are known to be more susceptible to THC sensitization [Wakley et al., 2014, Robinson, 1988] due to the sexually dimorphic distribution of endocannabinoid receptors in the striatum and prefrontal cortex [De Fonseca et al., 1994]. Estrogen also increases striatal dopamine release in response to psychoactive drugs [Becker, 1999]. Consequently, high estrogen levels in rats are associated with greater prefrontal dysfunction in response to THC and other recreational drugs [Shansky et al., 2004, Sárvári et al., 2014, Febo et al., 2005] and a decline in spatial ability [Janowsky et al., 1994, Makela et al., 2006]. Conversely, males exhibit greater dopamine release in the putamen, nucleus accumbens, and caudate nucleus during first-time drug use, relative to females [Munro et al., 2006].

Alternatively, male and female cannabis users may exhibit different patterns of use. However, among the 130 participants who met the WHO-ASSIST criteria for cannabis use in chapters 2 - 4 and the present study, males ($M = 10.73$, $SD = 6.79$) and females ($M = 10.77$, $SD = 8.17$) did not differ in scores of cannabis use

severity [$t(128) = .030$, $p = .976$]. Therefore, intermittent cannabis use is associated with decreased lose-shift responding while long-term sensitization causes increased behaviour.

Cannabis users also exhibited attenuated loss aversion, as evidenced by attenuated voltage potentials over the central scalp following feedback. In particular, males exhibited the greatest increase in loss-win ERPs during the FRN (253-273 ms post-feedback). The FRN is known to represent violation of reward expectation and subsequent re-weighting of action-outcome associations [Holroyd and Coles, 2002] by a signal generated in the anterior cingulate [Holroyd et al., 2003]. Attenuation of this signal was uniquely associated with increases in task performance due to improved response entropy. Consequently, better suppression of reward expectation was associated with the ability to generate random, mixed-strategy responses.

Loss-win differences in P3a amplitudes (306-326 ms post-feedback) were also significantly different in cannabis users relative to controls. The P3a is thought to reflect changes in the choice-reward model used to generate responses, rather than re-weighting of action-outcome associations within an existing model [Donchin and Coles, 1988, Linden, 2005]. Attenuated P3a amplitudes were uniquely correlated with a decrease in lose-shift responding, which is also known to be associated with alternation between response strategies [Ragozzino et al., 2002, Ragozzino, 2007, McDonald et al., 2008]. Moreover, P3a amplitudes are correlated with activity in the ventral striatum [Pfabigan et al., 2014]. Consequently, we hypothesize that the lose-shift response is supported by circuits in the ventral striatum that are also necessary for alternating between response strategies.

These cannabis-induced changes in reward processing were associated with greater

oscillatory activity in the anterior cingulate, medial, and superior frontal gyri during feedback processing, particularly in the $\delta$ (1-4 Hz) frequency band. Delta activity in these regions originates from a signal in the putamen [Foti et al., 2015] and is associated with reward processing [Wacker et al., 2009, Lucchiari and Pravettoni, 2010], impulsivity of decisions [Bernat et al., 2011], and inhibition of ongoing actions in order to shift to a new choice strategy [Braver et al., 2001, Swick et al., 2008, Sharp et al., 2010]. Moreover $\delta$ oscillations support the P3a potential [Başar-Eroglu et al., 1992] that is correlated with lose-shift responding. Therefore, given that cannabis users exhibit greater $\delta$ activity and attenuated P3a amplitudes, $\delta$ reflects processing relevant to the lose-shift response.

Males and females also exhibited activity differences during feedback processing. In particular, males exhibited greater gamma-band activity (30-80 Hz) in the right parahippocampal gyrus, uncus, temporal gyrus, and fusiform gyrus. Gamma activity in these regions are associated with emotional arousal [Hirai et al., 1999, Yang et al., 2018], suggesting the sexes may differ in their emotional response to rewards relative to punishment. However, $\gamma$ oscillations occur throughout the cortex, hippocampus, and ventral striatum [Kalenscher et al., 2010, Jenkinson et al., 2013, Headley and Paré, 2017]. In different contexts $\gamma$ is elicited during visual search [Tallon-Baudry et al., 1997, Leung and Borst, 1987, Bragin et al., 1995], visual awareness of a stimuli [Luo et al., 2008] particularly during movement [Cheron et al., 2016, Uemoto et al., 2017, Calabrò et al., 2017, Nowak et al., 2018], arousal [Litvak et al., 2012], top-down attentional processes [Kahana, 2006], and reward processing [Van Der Meer and Redish, 2009, Masimore et al., 2005]. Therefore, a targeted investigation is required to determine how sex-differences in gamma-activity relate to decision-making processes.

While not associated with sex or cannabis use, P3b scalp amplitudes (468-488 ms) following both wins and losses were correlated with learning rates ($\alpha$) and rates of win-stay responding. P3b amplitudes as associated with a number of factors including outcome probability, value, task complexity, and the amount of information obtained from an outcome [Duncan-Johnson and Donchin, 1977, Donchin, 1981, Johnson, 1986, Kok, 2001]. The correlation between win-stay behaviour and P3b amplitudes may simply be that greater valuation of, or surprise elicited by, wins is associated with greater P3b amplitudes. Individuals exhibiting greater P3b amplitudes also exhibited less change in learned action-outcome associations in response feedback. Therefore, the correlation between learning rates and P3b amplitudes likely reflects the rate at which learned action-outcome associations are updated.

In sum, the research presented here provide an in-depth analysis of the anatomical correlates of sensorimotor responding in humans. Win-stay responding is supported by the precuneus, motor cortex, and somatosensory circuits, both during reward processing and execution of motor actions. Moreover, motor and precuneal activity elicited during win-stay responding directly corresponds to the hand being used, even when compared against a win-shift response made with the identical hand, indicating it is represented in egocentric spatial coordinates. Conversely, lose-shift responses are supported by circuits in the cingulate and frontal gyri known to support allocentric processing of space. We also find that while lose-shift responding may be impacted differently in male and female cannabis users, this is primarily because females more easily sensitize to THC. In new cannabis users, males and females exhibit reduced loss aversion and attenuated neural signatures of feedback processing. Given the recent legalization of recreational cannabis in Canada, these factors are important to

consider in developing policies and regulations regarding cannabis use.

# Chapter 6

# General Discussion

## 6.1 Introduction

When making decisions, most animals tend to win-stay and lose-shift [Thorndike, 1911]. Together, these responses constitute a simple, but effective strategy for adapting our choices in a continuously changing environment [Nowak and Sigmund, 1993]. Since they were first codified by Thorndike [1898], these behaviours have been found to rely on the striatum, including the putamen, nucleus accumbens, and caudate nucleus [Packard et al., 1989, Packard and White, 1991, McDonald and White, 2013, Skelin et al., 2014, Gruber et al., 2017, Thapa and Gruber, 2018]. The striatum is primarily associated with integrating information from the limbic, associative, and sensorimotor systems in order to guide motor action [Murray et al., 2011, Northcutt, 2008]. Moreover, these striatal circuits represent actions in a number of spatial coordinate systems. The putamen and parts of the caudate nucleus are implicated in egocentric, self-oriented representations of space [Palencia and Ragozzino, 2005] while the nucleus accumbens and sections of the caudate are necessary for world-centered,

allocentric spatial processing [Zaehle et al., 2007, Gramann et al., 2006]. The dual role of the striatum in maintaining the win-stay/lose-shift strategy and spatial-motor processing suggests that these strategies are calculated according to spatial-motor actions, rather than as visually identifiable choices. In addition, evidence that the win-stay and lose-shift rely on different striatal sub-regions [McDonald and White, 2013, Skelin et al., 2014, Gruber et al., 2017, Thapa and Gruber, 2018] suggests these strategies are represented in different spatial reference frames.

In this dissertation I have provided evidence for three hypotheses. First, that the value of decisions are determined by their associated spatial-motor actions, not with specific choices or visual stimuli. Second, the win-stay and lose-shift are processed in different spatial reference frames. Reward-seeking (i.e., win-stay) behaviour is calculated in egocentric coordinates, relative to the participant. However, loss-aversion (i.e., lose-shift) is represented in an allocentric reference frame, calculated according to choice positions relative to one another. Third, I have demonstrated that habitual cannabis use alters our ability to inhibit lose-shift responses. In particular females are the most susceptible to cannabis sensitization.

## 6.2 Summary of results

In Chapter 2 I demonstrated that different locations are associated with unique win-stay and lose-shift tendencies and that the relationship between choice location and value depends on the motor action required for a choice. Win-stay responses were much more likely for locations near the hand used to make a response, while lose-shifting was prevalent for locations contralateral to the hand. In Chapter 3 we also found that win-stay and lose-shift tendencies are not tied to specific choices. Instead

wins and losses influence the value of many choices distributed throughout space. Moreover, humans exhibit two distinct lose-shift strategies oriented towards different regions in their environment. Following a loss participants often explore the region directly adjacent to their previous choice, similar to foraging. Alternatively, they may also completely avoid a previous loss, moving to a new region of their environment. In Chapter 4 I demonstrate that even when choices are visually distinct, the position of a previous win or loss determines the value of future choices in that location. When available choices switch positions, wins and losses remain associated with the original locations of each choice, rather than their associated symbols. Chapter 5 provides anatomical data supporting the association between spatial-motor action and choice value. We demonstrated that reward processing necessary for the win-stay and lose-shift, and execution of these responses, are associated with unique activation of motor, somatosensory, and visuospatial systems.

While both the win-stay and lose-shift are processed in spatial coordinates, the reference frame used differs for each action. In particular, we found that the win-stay is calculated in egocentric coordinates, supported by the putamen and associated sensorimotor circuitry. The lose-shift is calculated in allocentric coordinates supported by the nucleus accumbens, caudate nucleus, associative, and limbic circuits. For example, while both the win-stay and lose-shift were influenced by choice location in Chapter 2, the spatial bias in win-stay responding was only present during use of the right hand. Consequently, an individual's self-oriented frame of reference has considerable influence on win-stay behaviour. In Chapter 4 win-stay tendencies associated with a choice were disrupted by any change in its location. However, loss-aversion

corresponded to choice locations relative to one another, and not relative to the participant. When all choices moved to a new location, their associated lose-shift tendencies were retained so long as their relative positions remained unchanged. Therefore, the win-stay and lose-shift are calculated in egocentric and allocentric reference frames respectively. The results in Chapter 5 provide anatomical evidence for this dissociation. The execution of win-stay responses (relative to win-shift) was accompanied by greater activity in the motor cortex, sensorimotor cortex, and precuneus corresponding (contralateral) to the hand being used. These structures project to the putamen [Brasted et al., 1999, Pan et al., 2010, Malach and Graybiel, 1986], which is associated with egocentric processing of space [Kesner and DiMattia, 1987, Palencia and Ragozzino, 2005] and the win-stay response [Packard and White, 1991, McDonald and White, 2013]. Loss-aversion was not coded in the motor cortex. Instead, lose-shifting was associated with inhibition of the unused hand through greater activation of prefrontal structures [Harmony, 2013] that project to the nucleus accumbens [Gerfen, 1984, Voorn et al., 2004, Kelley et al., 1982], and are associated with allocentric processing of space [Zaehle et al., 2007, Gramann et al., 2006].

The different reference frames used to represent the win-stay and lose-shift conflicts with previous anatomical research on these behaviours. For example, in rats lose-shift behaviour is associated with the putamen, nucleus accumbens, caudate nucleus, hippocampus, and anterior cingulate in different contexts [Packard et al., 1989, Paulus et al., 2002a, Chen et al., 2012, Grospe et al., 2018, Thapa and Gruber, 2018]. However, these structures are also necessary for win-stay behaviour in a variety of contexts [Packard and White, 1991, Paulus et al., 2002a, McDonald and White, 2013,

Gruber et al., 2017]. This conflict suggests that multiple overlapping networks compete to drive the win-stay and lose-shift behaviours [Packard and White, 1991]. The extent to which each network drives behaviour depends on the context in which decisions are made [McDonald et al., 2008]. For example, in rats the putamen supports lose-shift and win-stay responses [Skelin et al., 2014]. These behaviours are strongly correlated, suggesting they are driven by a habitual, shifting-based strategy [Gruber et al., 2017]. The putamen is known to support these habitual, associative responses [Burton et al., 2014]. However, in the present thesis, lose-shifting was not correlated with win-stay behaviour and may be part of a goal-directed strategy to beat the computer opponent. Given that the nucleus accumbens drives goal-directed control of behaviour [Burton et al., 2015], the context in which decisions are made determines their associated neural circuitry.

In all four chapters I explored the relationship between recreational cannabis use and altered decision-making. The experiments in Chapters 2 - 4 found that recreational cannabis use had a sexually dimorphic effect on lose-shift behaviour. Female cannabis-users exhibited reduced task performance, an inability to suppress habitual lose-shift responding, and less behavioural flexibility in adapting to the decisions of the computer opponent. Conversely, male cannabis-users were better able to suppress habitual lose-shift responses. However, in Chapter 5 both male and female cannabis users exhibited better task performance and reduced lose-shift behaviour. As the experiment in Chapter 5 was conducted in the year following legalization of recreational cannabis in Canada, these results highlight the opposite effects short-term and habitual cannabis use may have on behaviour. Immediately following legalization, the proportion of cannabis users in Ontario rose from 13.5% to 20% of the population

Rotermann [2019]. Acute cannabis use is known to result in desensitization to losses resulting from increased dopamine release in the striatum [Wong et al., 2017a], reducing lose-shift behaviour. However, chronic drug abuse increases lose-shift responses due to inhibited dopamine production and weakened prefrontal control over sensorimotor response systems [Volkow et al., 2010]. Consequently, elevated lose-shifting among females cannabis users in Chapters 2 - 4 is indicative of cannabis sensitization while participants in Chapter 5 exhibited signs of acute use. Therefore, the sexually dimorphic effects of cannabis use on lose-shift behaviour are not due to functional differences in how cannabis affects the brain. Instead, females are known to be more susceptible to THC sensitization [Wakley et al., 2014, Robinson, 1988] due to the sexually dimorphic distribution of endocannabinoid receptors in the striatum and prefrontal cortex [De Fonseca et al., 1994]. Estrogen also increases striatal dopamine release in response to psychoactive drugs [Becker, 1999]. Consequently, high estrogen levels in rats are associated with greater prefrontal dysfunction in response to THC and other recreational drugs [Shansky et al., 2004, Sárvári et al., 2014, Febo et al., 2005] and a decline in spatial ability [Janowsky et al., 1994, Makela et al., 2006]. Conversely, males exhibit greater dopamine release in the putamen, nucleus accumbens, and caudate nucleus during first-time drug use, relative to females [Munro et al., 2006].

## 6.3 Future research

While this thesis provides considerable insight into the spatial-motor basis of decision-making, it also highlights a number of questions that have yet to be answered. All the experiments in Chapters 2 - 5 demonstrate that when motor actions are required to

make decisions, the location of a choice determines its value. Future research should answer how the win-stay and lose-shift are processed when choices contain no spatial component, for example, when made verbally.

The results in Chapters 2, 4, and 5 also indicate that the win-stay is supported by egocentric systems including the putamen while lose-shift responding is supported by allocentric circuits, such as the nucleus accumbens and caudate nucleus. These results contrast with the opposite findings in rats [Gruber et al., 2017]. However, in humans lose-shifting may be part of a goal-directed strategy to beat the computer opponent, while in rats they are driven by a habitual, sensorimotor response. Therefore, future research should whether manipulations to experimental context influence goal-directed and habitual control of the win-stay and lose-shift. If so, the lose-shift may be represented in an egocentric or allocentric reference frame, depending on the context in which decisions are made.

Finally, it is worth investigating how decision-making changes throughout the lifespan. Previously, Ivan et al. [2018] demonstrated that childhood development is marked by improved suppression of lose-shift processing, due to increased connectivity between the striatum and prefrontal executive systems. Conversely, normal aging is accompanied by degraded white matter integrity in the putamen, caudate nucleus, prefrontal cortex, and in connections from the PFC to the thalamus [Wang et al., 2010, Samanez-Larkin and Knutson, 2014] found that white matter integrity from the thalamus to PFC declines. Aging is also accompanied by depletion of striatal and prefrontal dopamine receptors [Floresco, 2013]. Each of these changes may influence win-stay and lose-shift behaviour in elderly adults. For example, older adults exhibit a reduced neural response to rewards in the ventral striatum [Marschner et al.,

2005, Schott et al., 2007], attenuated loss aversion [Beste et al., 2009], a greater reliance on rewards during learning [Eppinger et al., 2008], and impaired behavioural flexibility, marked by reduced lose-shift and win-stay behaviour [Means and Holsten, 1992]. Therefore, investigating win-stay and lose-shift responding in elderly adults will indicate how decision-making changes throughout the lifespan.

## 6.4 Conclusions

This dissertation demonstrates that choice location determines choice value. Both the win-stay and lose-shift responses are strongly influenced by the spatial-motor action needed to make that response. Consequently, wins and losses experienced at one spatial location influence future actions there, regardless of the visual cues or choices present. However, the win-stay and lose-shift are not calculated in the same manner. The win-stay is represented in self-referential (egocentric) spatial coordinates, supported by the putamen, motor, and somatosensory systems. Lose-shift behaviour is processed in world-centred (allocentric) coordinates, supported by associative systems and the nucleus accumbens. Lose-shift tendencies associated with a choice persist in different locations, so long as the spatial relationships between choices are maintained. Lose-shift responding also varies as a function of cannabis use, particularly in female users who often exhibit elevated lose-shift behaviour. Males also exhibit changes in lose-shift responding concordant with cannabis use. However, they are much less susceptible to the effects of cannabis sensitization following long term use.

# Bibliography

M. Abouzari, S. Oberg, A. Gruber, and M. Tata. Interactions among attention-deficit hyperactivity disorder (adhd) and problem gambling in a probabilistic reward-learning task. *Behavioural Brain Research*, 291:237–243, 2015.

H. Akhlaghpour, J. Wiskerke, J. Y. Choi, J. P. Taliaferro, J. Au, and I. B. Witten. Dissociated sequential activity and stimulus encoding in the dorsomedial striatum during spatial working memory. *Elife*, 5:e19507, 2016.

G. Albouy, V. Sterpenich, E. Balteau, G. Vandewalle, M. Desseilles, T. Dang-Vu, A. Darsaud, P. Ruby, P.-H. Luppi, C. Degueldre, et al. Both the hippocampus and striatum are involved in consolidation of motor sequence memory. *Neuron*, 58(2): 261–272, 2008.

D. S. Albrecht, P. D. Skosnik, J. M. Vollmer, M. S. Brumbaugh, K. M. Perry, B. H. Mock, Q.-H. Zheng, L. A. Federici, E. A. Patton, C. M. Herring, et al. Striatal d2/d3 receptor availability is inversely correlated with cannabis consumption in chronic marijuana users. *Drug and Alcohol Dependence*, 128(1-2):52–57, 2013.

F. Aoki, E. E. Fetz, L. Shupe, E. Lettich, and G. A. Ojemann. Increased gamma-range activity in human sensorimotor cortex during performance of visuomotor

tasks. *Clinical Neurophysiology*, 110(3):524–537, 1999.

P. Apicella, E. Scarnati, T. Ljungberg, and W. Schultz. Neuronal activity in monkey striatum related to the expectation of predictable environmental events. *Journal of Neurophysiology*, 68(3):945–960, 1992.

S. N. Baker. Oscillatory interactions between sensorimotor cortex and the periphery. *Current Opinion in Neurobiology*, 17(6):649–655, 2007.

B. W. Balleine and J. P. O'Doherty. Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology*, 35(1):48–69, 2010.

P. J. Banks, M. S. Tata, P. J. Bennett, A. B. Sekuler, and A. J. Gruber. Implicit valuation of the near-miss is dependent on outcome context. *Journal of Gambling Studies*, 34(1):181–197, 2018.

D. J. Barr, R. Levy, C. Scheepers, and H. J. Tily. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3):255–278, 2013.

D. J. Barraclough, M. L. Conroy, and D. Lee. Prefrontal cortex and decision making in a mixed-strategy game. *Nature Neuroscience*, 7(4):404, 2004.

C. Başar-Eroglu, E. Başar, T. Demiralp, and M. Schürmann. P300-response: possible psychophysiological correlates in delta and theta frequency channels. a review. *International Journal of Psychophysiology*, 13(2):161–179, 1992.

D. Bates, M. Mächler, B. Bolker, and S. Walker. Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*, 2014.

J. B. Becker. Gender differences in dopaminergic function in striatum and nucleus accumbens. *Pharmacology Biochemistry and Behavior*, 64(4):803–812, 1999.

C. G. Beevers, D. A. Worthy, M. A. Gorlick, B. Nix, T. Chotibut, and W. T. Maddox. Influence of depression symptoms on history-independent reward and punishment processing. *Psychiatry Research*, 207(1-2):53–60, 2013.

E. Ben-Shabat, T. A. Matyas, G. S. Pell, A. Brodtmann, and L. M. Carey. The right supramarginal gyrus is important for proprioception in healthy and stroke-affected participants: a functional mri study. *Frontiers in Neurology*, 6:248, 2015.

F. Benfenati, E. M. Pich, M. Zoli, R. Grimaldi, K. Fuxe, and L. F. Agnati. Changes in striatal $\mu$ and $\delta$ opioid receptors after transient forebrain ischemia: a quantitative autoradiographic study. *Brain Research*, 546(1):171–175, 1991.

E. M. Bernat, L. D. Nelson, V. R. Steele, W. J. Gehring, and C. J. Patrick. Externalizing psychopathology and gain–loss feedback in a simulated gambling task: Dissociable components of brain response revealed by time-frequency analysis. *Journal of Abnormal Psychology*, 120(2):352, 2011.

C. Beste, R. Willemssen, C. Saft, and M. Falkenstein. Error processing in normal aging and in basal ganglia disorders. *Neuroscience*, 159(1):143–149, 2009.

D. Bett, E. Allison, L. H. Murdoch, K. Kaefer, E. R. Wood, and P. A. Dudchenko. The neural substrates of deliberative decision making: contrasting effects of hippocampus lesions on performance and vicarious trial-and-error behavior in a spatial memory task and a visual discrimination task. *Frontiers in Behavioral Neuroscience*, 6: 70, 2012.

L. Bidwell, E. Henry, E. Willcutt, M. Kinnear, and T. Ito. Childhood and current adhd symptom dimensions are associated with more severe cannabis outcomes in college students. *Drug and Alcohol Dependence*, 135:88–94, 2014.

L. Blanco-Hinojo, J. Pujol, B. J. Harrison, D. Macià, A. Batalla, S. Nogué, M. Torrens, M. Farré, J. Deus, and R. Martín-Santos. Attenuated frontal and sensory inputs to the basal ganglia in cannabis users. *Addiction Biology*, 22(4):1036–1047, 2017.

V. D. Bohbot, J. J. Allen, A. Dagher, S. O. Dumoulin, A. C. Evans, M. Petrides, M. Kalina, K. Stepankova, and L. Nadel. Role of the parahippocampal cortex in memory for the configuration but not the identity of objects: converging evidence from patients with selective thermal lesions and fmri. *Frontiers in Human Neuroscience*, 9:431, 2015.

A. B. Bond, R. G. Cook, and M. R. Lamb. Spatial memory and the performance of rats and pigeons in the radial-arm maze. *Animal Learning & Behavior*, 9(4): 575–580, 1981.

M. G. Bossong, B. N. Van Berckel, R. Boellaard, L. Zuurman, R. C. Schuit, A. D. Windhorst, J. M. Van Gerven, N. F. Ramsey, A. A. Lammertsma, and R. S. Kahn. $\delta$9-tetrahydrocannabinol induces dopamine release in the human striatum. *Neuropsychopharmacology*, 34(3):759–766, 2009.

S. Boyd, L. O. Chua, and C. A. Desoer. Analytical foundations of volterra series. *IMA Journal of Mathematical Control and Information*, 1(3):243–282, 1984.

A. Bragin, G. Jandó, Z. Nádasdy, J. Hetke, K. Wise, and G. Buzsáki. Gamma (40-100

hz) oscillation in the hippocampus of the behaving rat. *Journal of Neuroscience*, 15(1):47–60, 1995.

P. J. Brasted, T. W. Robbins, and S. B. Dunnett. Distinct roles for striatal subregions in mediating response processing revealed by focal excitotoxic lesions. *Behavioral Neuroscience*, 113(2):253, 1999.

T. S. Braver, D. M. Barch, J. R. Gray, D. L. Molfese, and A. Snyder. Anterior cingulate cortex and response conflict: effects of frequency, inhibition and errors. *Cerebral Cortex*, 11(9):825–836, 2001.

E. S. Brush, M. Mishkin, and H. Rosvold. Effects of object preferences and aversions on discrimination learning in monkeys with frontal lesions. *Journal of Comparative and Physiological Psychology*, 54(3):319, 1961.

E. A. Budygin, E. B. Oleson, T. A. Mathews, A. K. Läck, M. R. Diaz, B. A. McCool, and S. R. Jones. Effects of chronic alcohol exposure on dopamine uptake in rat nucleus accumbens and caudate putamen. *Psychopharmacology*, 193(4):495–501, 2007.

A. C. Burton, G. B. Bissonette, N. T. Lichtenberg, V. Kashtelyan, and M. R. Roesch. Ventral striatum lesions enhance stimulus and response encoding in dorsal striatum. *Biological Psychiatry*, 75(2):132–139, 2014.

A. C. Burton, K. Nakamura, and M. R. Roesch. From ventral-medial to dorsal-lateral striatum: neural correlates of reward-guided decision-making. *Neurobiology of Learning and Memory*, 117:51–59, 2015.

L. J. Buxbaum, K. Kyle, M. Grossman, and B. Coslett. Left inferior parietal representations for skilled hand-object interactions: evidence from stroke and corticobasal degeneration. *Cortex*, 43(3):411–423, 2007.

R. S. Calabrò, A. Naro, M. Russo, A. Leo, R. De Luca, T. Balletta, A. Buda, G. La Rosa, A. Bramanti, and P. Bramanti. The role of virtual reality in improving motor performance as revealed by eeg: a randomized clinical trial. *Journal of neuroengineering and rehabilitation*, 14(1):53, 2017.

K. C. Calakos, S. Bhatt, D. W. Foster, and K. P. Cosgrove. Mechanisms underlying sex differences in cannabis use. *Current addiction reports*, 4(4):439–453, 2017.

R. N. Cardinal, J. A. Parkinson, J. Hall, and B. J. Everitt. Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. *Neuroscience & Biobehavioral Reviews*, 26(3):321–352, 2002.

J. Castelhano, I. C. Duarte, M. Wibral, E. Rodriguez, and M. Castelo-Branco. The dual facet of gamma oscillations: separate visual and decision making circuits as revealed by simultaneous eeg/fmri. *Human Brain Mapping*, 35(10):5219–5235, 2014.

A. E. Cavanna and M. R. Trimble. The precuneus: a review of its functional anatomy and behavioural correlates. *Brain*, 129(3):564–583, 2006.

Y. M. Cha, K. H. Jones, C. M. Kuhn, W. A. Wilson, and H. S. Swartzwelder. Sex differences in the effects of $\delta$9-tetrahydrocannabinol on spatial learning in adolescent and adult rats. *Behavioural pharmacology*, 18(5-6):563–569, 2007.

S. R. Chamberlain and B. J. Sahakian. The neuropsychiatry of impulsivity. *Current opinion in psychiatry*, 20(3):255–261, 2007.

F. Chella, V. Pizzella, F. Zappasodi, and L. Marzetti. Impact of the reference choice on scalp eeg connectivity estimation. *Journal of neural engineering*, 13(3):036016, 2016.

X.-P. Chen, W.-Z. Chen, F.-S. Wang, and J.-X. Liu. Selective cognitive impairments are related to selective hippocampus and prefrontal cortex deficits after prenatal chlorpyrifos exposure. *Brain Research*, 1474:19–28, 2012.

G. Cheron, G. Petit, J. Cheron, A. Leroy, A. Cebolla, C. Cevallos, M. Petieau, T. Hoellinger, D. Zarka, A.-M. Clarinval, et al. Brain oscillations in sport: toward eeg biomarkers of performance. *Frontiers in psychology*, 7:246, 2016.

H. F. Clarke, T. W. Robbins, and A. C. Roberts. Lesions of the medial striatum in monkeys produce perseverative impairments during reversal learning similar to those produced by lesions of the orbitofrontal cortex. *Journal of Neuroscience*, 28 (43):10972–10982, 2008.

R. B. Clarke, B. Söderpalm, A. Lotfi, M. Ericson, and L. Adermark. Involvement of inhibitory receptors in modulating dopamine signaling and synaptic activity following acute ethanol exposure in striatal subregions. *Alcoholism: Clinical and Experimental Research*, 39(12):2364–2374, 2015.

M. Clemente, A. Rodríguez, B. Rey, and M. Alcañiz. Assessment of the influence of navigation control and screen size on the sense of presence in virtual reality using eeg. *Expert Systems with Applications*, 41(4):1584–1592, 2014.

M. X. Cohen. Error-related medial frontal theta activity predicts cingulate-related structural connectivity. *Neuroimage*, 55(3):1373–1383, 2011.

M. X. Cohen and C. Ranganath. Reinforcement learning signals predict future decisions. *The Journal of Neuroscience*, 27(2):371–378, 2007.

M. X. Cohen, C. E. Elger, and C. Ranganath. Reward expectation modulates feedback-related negativity and eeg spectra. *Neuroimage*, 35(2):968–978, 2007.

D. L. Collins, P. Neelin, T. M. Peters, and A. C. Evans. Automatic 3d intersubject registration of mr volumetric data in standardized talairach space. *Journal of computer assisted tomography*, 18(2):192–205, 1994.

R. Cools, S. E. Gibbs, A. Miyakawa, W. Jagust, and M. D'Esposito. Working memory capacity predicts dopamine synthesis capacity in the human striatum. *Journal of Neuroscience*, 28(5):1208–1212, 2008.

C. L. Coulter, H. K. Happe, and L. C. Murrin. Dopamine transporter development in postnatal rat striatum: an autoradiographic study with [3h] win 35,428. *Developmental brain research*, 104(1-2):55–62, 1997.

N. A. Crane, R. M. Schuster, and R. Gonzalez. Preliminary evidence for a sex-specific relationship between amount of cannabis use and neurocognitive performance in young adult cannabis users. *Journal of the International Neuropsychological Society: JINS*, 19(9):1009, 2013.

S. I. Cunningham, D. Tomasi, and N. D. Volkow. Structural and functional connectivity of the precuneus and thalamus to the default mode network. *Human Brain Mapping*, 38(2):938–956, 2017.

J. Danckert, E. Stöttinger, N. Quehl, and B. Anderson. Right hemisphere brain damage impairs strategy updating. *Cerebral Cortex*, 22(12):2745–2760, 2011.

J. Z. Daniel, M. Hickman, J. Macleod, N. Wiles, A. LINGFORD-HUGHES, M. Farrell, R. Araya, P. Skapinakis, J. Haynes, and G. Lewis. Is socioeconomic status in early life associated with drug use? a systematic review of the evidence. *Drug and alcohol review*, 28(2):142–153, 2009.

M. Darvas and R. D. Palmiter. Restriction of dopamine signaling to the dorsolateral striatum is sufficient for many cognitive behaviors. *Proceedings of the National Academy of Sciences*, 106(34):14664–14669, 2009.

M. Darvas and R. D. Palmiter. Restricting dopaminergic signaling to either dorsolateral or medial striatum facilitates cognition. *Journal of Neuroscience*, 30(3): 1158–1165, 2010.

J. Daunizeau, V. Adam, and L. Rigoux. Vba: a probabilistic treatment of nonlinear models for neurobiological and behavioural data. *PLoS Comput Biol*, 10(1): e1003441, 2014.

F. R. De Fonseca, M. Cebeira, J. Ramos, M. Martin, and J. Fernandez-Ruiz. Cannabinoid receptors in rat brain areas: sexual differences, fluctuations during estrous cycle and changes after gonadectomy and sex steroid replacement. *Life sciences*, 54(3):159–170, 1994.

R. de la Fuente-Fernández, A. Kishore, D. B. Calne, T. J. Ruth, and A. J. Stoessl. Nigrostriatal dopamine system and motor lateralization. *Behavioural Brain Research*, 112(1-2):63–68, 2000.

F. P. de Lange, D. A. Rahnev, T. H. Donner, and H. Lau. Prestimulus oscillatory activity over motor cortex reflects perceptual expectations. *Journal of Neuroscience*, 33(4):1400–1410, 2013.

E. De Leonibus, A. Oliverio, and A. Mele. A study on the role of the dorsal striatum and the nucleus accumbens in allocentric and egocentric spatial memory consolidation. *Learning & Memory*, 12(5):491–503, 2005.

I. E. de Vries, J. van Driel, M. Karacaoglu, and C. N. Olivers. Priority switches in visual working memory are supported by frontal delta and posterior alpha interactions. *Cerebral Cortex*, 28(11):4090–4104, 2018.

A. Delorme and S. Makeig. Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1):9–21, 2004.

T. Demiralp, C. S. Herrmann, M. E. Erdal, T. Ergenoglu, Y. H. Keskin, M. Ergen, and H. Beydagi. Drd4 and dat1 polymorphisms modulate human gamma band responses. *Cerebral Cortex*, 17(5):1007–1019, 2007.

H. Deubel, W. X. Schneider, and I. Paprotta. Selective dorsal and ventral processing: Evidence for a common attentional mechanism in reaching and perception. *Visual Cognition*, 5(1-2):81–107, 1998.

E. Donchin. Surprise!... surprise? *Psychophysiology*, 18(5):493–513, 1981.

E. Donchin and M. G. Coles. Is the p300 component a manifestation of context updating? *Behavioral and Brain Sciences*, 11(3):357–374, 1988.

L. Dong, F. Li, Q. Liu, X. Wen, Y. Lai, P. Xu, and D. Yao. Matlab toolboxes for reference electrode standardization technique (rest) of scalp eeg. *Frontiers in Neuroscience*, 11:601, 2017.

L. W. Drozdick, D. Wahlstrom, J. Zhu, and L. G. Weiss. The wechsler adult intelligence scale—fourth edition and the wechsler memory scale—fourth edition. 2012.

C. C. Duncan-Johnson and E. Donchin. On quantifying surprise: The variation of event-related potentials with subjective probability. *Psychophysiology*, 14(5):456–467, 1977.

B. Eppinger, J. Kray, B. Mock, and A. Mecklinger. Better or worse than expected? aging, learning, and the ern. *Neuropsychologia*, 46(2):521–539, 2008.

A. C. Evans, D. L. Collins, and B. Milner. An mri-based stereotactic atlas from 250 young normal subjects. *Society for Neuroscience Abstracts*, 18(408), 1992.

J. Evenden and T. W. Robbins. Effects of unilateral 6-hydroxydopamine lesions of the caudate-putamen on skilled forepaw use in the rat. *Behavioural Brain Research*, 14(1):61–68, 1984.

B. J. Everitt and T. W. Robbins. Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nature Neuroscience*, 8(11):1481, 2005.

B. J. Everitt and T. W. Robbins. From the ventral to the dorsal striatum: devolving views of their roles in drug addiction. *Neuroscience & Biobehavioral Reviews*, 37 (9):1946–1954, 2013.

L. Fadiga, L. Fogassi, V. Gallese, and G. Rizzolatti. Visuomotor neurons: Ambiguity

of the discharge or 'motor'perception? *International Journal of Psychophysiology*, 35(2-3):165–177, 2000.

M. Febo, C. F. Ferris, and A. C. Segarra. Estrogen influences cocaine-induced blood oxygen level-dependent signal changes in female rats. *Journal of Neuroscience*, 25 (5):1132–1136, 2005.

S. B. Floresco. Prefrontal dopamine and behavioral flexibility: shifting from an "inverted-u" toward a family of functions. *Frontiers in Neuroscience*, 7:62, 2013.

S. B. Floresco, J. R. S. Onge, S. Ghods-Sharifi, and C. A. Winstanley. Cortico-limbic-striatal circuits subserving different forms of cost-benefit decision making. *Cognitive, Affective, & Behavioral Neuroscience*, 8(4):375–389, 2008.

Y. M. Fonken, J. W. Rieger, E. Tzvi, N. E. Crone, E. Chang, J. Parvizi, R. T. Knight, and U. M. Krämer. Frontal and motor cortex contributions to response inhibition: evidence from electrocorticography. *Journal of Neurophysiology*, 115(4):2224–2236, 2016.

D. Foti, A. Weinberg, E. M. Bernat, and G. H. Proudfit. Anterior cingulate activity to monetary loss and basal ganglia activity to monetary gain uniquely contribute to the feedback negativity. *Clinical Neurophysiology*, 126(7):1338–1347, 2015.

C. Fraley and A. E. Raftery. Mclust: Software for model-based cluster analysis. *Journal of classification*, 16(2):297–306, 1999.

R. Frey, R. Mata, and R. Hertwig. The role of cognitive abilities in decisions from experience: Age differences emerge as a function of choice set size. *Cognition*, 142: 60–80, 2015.

L. Frings, K. Wagner, A. Quiske, R. Schwarzwald, J. Spreer, U. Halsband, and A. Schulze-Bonhage. Precuneus is involved in allocentric spatial location encoding and recognition. *Experimental Brain Research*, 173(4):661–672, 2006.

M. V. Fuccillo. Striatal circuits as a common node for autism pathophysiology. *Frontiers in Neuroscience*, 10:27, 2016.

J. Fuster. The prefrontal cortex: Anatomy, physiology, and neuropsychology of the frontal lobe . new york. *Raven Press.(Original work published 1980). Dissertation Abstracts International: The Sciences and Engineering*, 60(11-B):255, 1989.

W. Gaetz, C. Liu, H. Zhu, L. Bloy, and T. P. Roberts. Evidence for a motor gamma-band network governing response interference. *Neuroimage*, 74:245–253, 2013.

C. R. Gerfen. The neostriatal mosaic: compartmentalization of corticostriatal input and striatonigral output systems. *Nature*, 311(5985):461, 1984.

J. M. Gilman, J. K. Kuster, S. Lee, M. J. Lee, B. W. Kim, N. Makris, A. van der Kouwe, A. J. Blood, and H. C. Breiter. Cannabis use is quantitatively associated with nucleus accumbens and amygdala abnormalities in young adult recreational users. *Journal of Neuroscience*, 34(16):5529–5538, 2014.

J. Gläscher, D. Tranel, L. K. Paul, D. Rudrauf, C. Rorden, A. Hornaday, T. Grabowski, H. Damasio, and R. Adolphs. Lesion mapping of cognitive abilities linked to intelligence. *Neuron*, 61(5):681–691, 2009.

P. W. Glimcher. Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. *Proceedings of the National Academy of Sciences*, 108(Supplement 3):15647–15654, 2011.

M. A. Goodale and A. D. Milner. Separate visual pathways for perception and action. *Trends in Neurosciences*, 15(1):20–25, 1992.

J. P. Goyer, M. G. Woldorff, and S. A. Huettel. Rapid electrophysiological brain responses are influenced by both valence and magnitude of monetary rewards. *Journal of cognitive neuroscience*, 20(11):2058–2069, 2008.

K. Gramann, H. Müller, B. Schönebeck, and G. Debus. The neural basis of ego-and allocentric reference frames in spatial navigation: Evidence from spatio-temporal coupled current density reconstruction. *Brain Research*, 1118(1):116–129, 2006.

M. S. Graziano and C. G. Gross. Multiple pathways for processing visual space. *Attention and performance XVI: Information integration in perception and communication*, pages 181–207, 1996.

M. S. Graziano and C. G. Gross. Spatial maps for the control of movement. *Current Opinion in Neurobiology*, 8(2):195–201, 1998.

H. J. Groenewegen. The basal ganglia and motor control. *Neural plasticity*, 10(1-2): 107–120, 2003.

G. M. Grospe, P. M. Baker, and M. E. Ragozzino. Cognitive flexibility deficits following 6-ohda lesions of the rat dorsomedial striatum. *Neuroscience*, 374:80–90, 2018.

A. J. Gruber and R. J. McDonald. Context, emotion, and the strategic pursuit of goals: interactions among multiple brain systems controlling motivated behavior. *Frontiers in Behavioral Neuroscience*, 6:50, 2012.

A. J. Gruber and R. Thapa. The memory trace supporting lose-shift responding decays rapidly after reward omission and is distinct from other learning mechanisms in rats. *Eneuro*, 3(6), 2016.

A. J. Gruber, R. Thapa, and S. H. Randolph. Feeder approach between trials is increased by uncertainty and affects subsequent choices. *eNeuro*, 4(6):ENEURO–0437, 2017.

T. Gruber, M. M. Müller, A. Keil, and T. Elbert. Selective visual-spatial attention alters induced gamma band responses in the human eeg. *Clinical neurophysiology*, 110(12):2074–2085, 1999.

G. Hajcak, J. D. Huppert, R. F. Simons, and E. B. Foa. Psychometric properties of the oci-r in a college sample. *Behaviour research and therapy*, 42(1):115–123, 2004.

A. HajiHosseini, A. Rodríguez-Fornells, and J. Marco-Pallarés. The role of beta-gamma oscillations in unexpected rewards processing. *Neuroimage*, 60(3):1678–1685, 2012.

A. Hampshire, S. R. Chamberlain, M. M. Monti, J. Duncan, and A. M. Owen. The role of the right inferior frontal gyrus: inhibition and attentional control. *Neuroimage*, 50(3):1313–1319, 2010.

J. Harclerode. Endocrine effects of marijuana in the male: preclinical studies. *NIDA Res. Monogr*, 44:46–64, 1984.

T. Harmony. The functional significance of delta oscillations in cognitive processing. *Frontiers in integrative neuroscience*, 7:83, 2013.

J. V. Haxby, C. L. Grady, B. Horwitz, L. G. Ungerleider, M. Mishkin, R. E. Carson, P. Herscovitch, M. B. Schapiro, and S. I. Rapoport. Dissociation of object and spatial visual processing pathways in human extrastriate cortex. *Proceedings of the National Academy of Sciences*, 88(5):1621–1625, 1991.

M. R. Hayatbakhsh, J. M. Najman, K. Jamrozik, A. A. Mamun, R. Alati, and W. Bor. Cannabis and anxiety and depression in young adults: a large prospective study. *Journal of the American Academy of Child & Adolescent Psychiatry*, 46(3):408–417, 2007.

D. B. Headley and D. Paré. Common oscillatory mechanisms across multiple memory systems. *NPJ science of learning*, 2(1):1, 2017.

M. Herkenham, A. B. Lynn, M. R. Johnson, L. S. Melvin, B. R. de Costa, and K. C. Rice. Characterization and localization of cannabinoid receptors in rat brain: a quantitative in vitro autoradiographic study. *Journal of Neuroscience*, 11(2):563–583, 1991.

N. A. Herweg, T. Apitz, G. Leicht, C. Mulert, L. Fuentemilla, and N. Bunzeck. Theta-alpha oscillations bind the hippocampus, prefrontal cortex, and striatum during recollection: evidence from simultaneous eeg–fmri. *Journal of Neuroscience*, 36 (12):3579–3587, 2016.

N. Hirai, S. Uchida, T. Maehara, Y. Okubo, and H. Shimizu. Enhanced gamma (30–150 hz) frequency in the human medial temporal lobe. *Neuroscience*, 90(4): 1149–1155, 1999.

C. B. Holroyd and M. G. Coles. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, 109(4):679, 2002.

C. B. Holroyd, S. Nieuwenhuis, N. Yeung, and J. D. Cohen. Errors in reward prediction are reflected in the event-related brain potential. *Neuroreport*, 14(18):2481–2484, 2003.

M. Ito and K. Doya. Validation of decision-making models and analysis of decision variables in the rat basal ganglia. *The Journal of neuroscience*, 29(31):9861–9874, 2009.

V. E. Ivan, P. J. Banks, K. Goodfellow, and A. J. Gruber. Lose-shift responding in humans is promoted by increased cognitive load. *Frontiers in integrative neuroscience*, 12:9, 2018.

G. Jager, R. I. Block, M. Luijten, and N. F. Ramsey. Tentative evidence for striatal hyperactivity in adolescent cannabis-using boys: a cross-sectional multicenter fmri study. *Journal of psychoactive drugs*, 45(2):156–167, 2013.

J. S. Janowsky, S. K. Oviatt, and E. S. Orwoll. Testosterone influences spatial cognition in older men. *Behavioral Neuroscience*, 108(2):325, 1994.

N. Jenkinson, A. A. Kühn, and P. Brown. Gamma oscillations in the human basal ganglia. *Experimental neurology*, 245:72–76, 2013.

J. D. Jentsch, A. Wise, Z. Katz, and R. H. Roth. $\alpha$-noradrenergic receptor modulation of the phencyclidine-and $\delta$9-tetrahydrocannabinol-induced increases in dopamine utilization in rat prefrontal cortex. *Synapse*, 28(1):21–26, 1998.

S. Johannes, T. Münte, H. Heinze, and G. R. Mangun. Luminance and spatial attention effects on early visual processing. *Cognitive Brain Research*, 2(3):189–205, 1995.

R. Johnson. A triarchic model of p300 amplitude. *Psychophysiology*, 1986.

K. Jordan, J. Schadow, T. Wuestenberg, H.-J. Heinze, and L. Jäncke. Different cortical activations for subjects using allocentric or egocentric strategies in a virtual navigation task. *Neuroreport*, 15(1):135–140, 2004.

M. J. Kahana. The cognitive correlates of human brain oscillations. *Journal of Neuroscience*, 26(6):1669–1672, 2006.

T. Kalenscher, C. S. Lansink, J. V. Lankelma, and C. M. Pennartz. Reward-associated gamma oscillations in ventral striatum are regionally differentiated and modulate local firing activity. *Journal of Neurophysiology*, 103(3):1658–1672, 2010.

A. C. Kamil and M. W. Hunter. Performance on object-discrimination learning set by the greater hill myna (gracula religiosa). *Journal of Comparative and Physiological Psychology*, 73(1):68, 1970.

J. Kayser and C. E. Tenke. Issues and considerations for using the scalp surface laplacian in eeg/erp research: A tutorial review. *International Journal of Psychophysiology*, 97(3):189–209, 2015.

A. Kelley, V. B. Domesick, and W. Nauta. The amygdalostriatal projection in the rat—an anatomical study by anterograde and retrograde tracing methods. *Neuroscience*, 7(3):615–630, 1982.

R. Kesner and B. DiMattia. Neurobiology of an attribute model of memory. *Progress in psychobiology and physiological psychology*, 12:207–277, 1987.

R. C. Kessler, L. Adler, M. Ames, O. Demler, S. Faraone, E. Hiripi, M. J. Howes, R. Jin, K. Secnik, T. Spencer, et al. The world health organization adult adhd self-report scale (asrs): a short screening scale for use in the general population. *Psychological medicine*, 35(2):245, 2005.

L. A. Khibnik, N. X. Tritsch, and B. L. Sabatini. A direct projection from mouse primary visual cortex to dorsomedial striatum. *PloS one*, 9(8), 2014.

H.-y. Kim, Y. Shin, and S. Han. The reconstruction of choice value in the brain: A look into the size of consideration sets and their affective consequences. *Journal of cognitive neuroscience*, 26(4):810–824, 2014.

A. Kok. On the utility of p3 amplitude as a measure of processing capacity. *Psychophysiology*, 38(3):557–577, 2001.

R. C. Kolodny, W. H. Masters, R. M. Kolodner, and G. Toro. Depression of plasma testosterone levels after chronic intensive marihuana use. *New England Journal of Medicine*, 290(16):872–874, 1974.

P. Lakatos, N. Szilágyi, Z. Pincze, C. Rajkai, I. Ulbert, and G. Karmos. Attention and arousal related modulation of spontaneous gamma-activity in the auditory cortex of the cat. *Cognitive brain research*, 19(1):1–9, 2004.

S. Lee and S. R. Jones. Distinguishing mechanisms of gamma frequency oscillations in human current source signals using a computational model of a laminar neocortical network. *Frontiers in Human Neuroscience*, 7:869, 2013.

S. Legleye, F. Beck, M. Khlat, P. Peretti-Watel, and N. Chau. The influence of socioeconomic status on cannabis use among french adolescents. *Journal of Adolescent Health*, 50(4):395–402, 2012.

T. Lejarraga and R. Hertwig. How the threat of losses makes people explore more than the promise of gains. *Psychonomic bulletin & review*, 24(3):708–720, 2017.

L.-W. Leung and J. Borst. Electrical activity of the cingulate cortex. i. generating mechanisms and relations to behavior. *Brain Research*, 407(1):68–80, 1987.

S. D. Lichenstein, S. Musselman, D. S. Shaw, S. Sitnick, and E. E. Forbes. Nucleus accumbens functional connectivity at age 20 is associated with trajectory of adolescent cannabis use and predicts psychosocial functioning in young adulthood. *Addiction*, 112(11):1961–1970, 2017.

D. E. Linden. The p300: where in the brain is it produced and what does it tell us? *The Neuroscientist*, 11(6):563–576, 2005.

I. Liste, G. Rozas, M. Guerra, and J. Labandeira-Garcia. Cortical stimulation induces fos expression in striatal neurons via nmda glutamate and dopamine receptors. *Brain Research*, 700(1-2):1–12, 1995.

V. Litvak, A. Eusebio, A. Jha, R. Oostenveld, G. Barnes, T. Foltynie, P. Limousin, L. Zrinzo, M. I. Hariz, K. Friston, et al. Movement-related changes in local and long-range synchronization in parkinson's disease revealed by simultaneous magnetoencephalography and intracranial recordings. *Journal of Neuroscience*, 32(31): 10541–10553, 2012.

S. H. Lovibond and P. F. Lovibond. *Manual for the depression anxiety stress scales*. Psychology Foundation of Australia, 1996.

D. I. Lubman, A. Cheetham, and M. Yücel. Cannabis and adolescent brain development. *Pharmacology & therapeutics*, 148:1–16, 2015.

F. Lucantonio, D. Caprioli, and G. Schoenbaum. Transition from 'model-based'to 'model-free'behavioral control in addiction: involvement of the orbitofrontal cortex and dorsolateral striatum. *Neuropharmacology*, 76:407–415, 2014.

C. Lucchiari and G. Pravettoni. Feedback related brain activity in a gambling task: a temporal analysis of eeg correlates. *Scandinavian journal of psychology*, 51(6): 449–454, 2010.

Q. Luo, D. Mitchell, X. Cheng, K. Mondillo, D. Mccaffrey, T. Holroyd, F. Carver, R. Coppola, and J. Blair. Visual awareness, emotion, and gamma band synchronization. *Cerebral Cortex*, 19(8):1896–1904, 2008.

G. Luppino, M. Matelli, R. Camarda, V. Gallese, and G. Rizzolatti. Multiple representations of body movements in mesial area 6 and the adjacent cingulate cortex: an intracortical microstimulation study in the macaque monkey. *Journal of Comparative Neurology*, 311(4):463–482, 1991.

M. T. Lynskey, A. L. Glowinski, A. A. Todorov, K. K. Bucholz, P. A. Madden, E. C. Nelson, D. J. Statham, N. G. Martin, and A. C. Heath. Major depressive disorder, suicidal ideation, and suicide attempt intwins discordant for cannabis dependence and early-onset cannabis use. *Archives of general psychiatry*, 61(10):1026–1032, 2004.

S. Makeig, M. Westerfield, T.-P. Jung, S. Enghoff, J. Townsend, E. Courchesne, and T. J. Sejnowski. Dynamic brain sources of visual evoked responses. *Science*, 295 (5555):690–694, 2002.

S. Makeig, S. Debener, J. Onton, and A. Delorme. Mining event-related brain dynamics. *Trends in cognitive sciences*, 8(5):204–210, 2004.

P. Makela, J. Wakeley, H. Gijsman, P. J. Robson, Z. Bhagwagar, and R. D. Rogers. Low doses of $\delta$-9 tetrahydrocannabinol (thc) have divergent effects on short-term spatial memory in young, healthy adults. *Neuropsychopharmacology*, 31(2):462–470, 2006.

R. Malach and A. M. Graybiel. Mosaic architecture of the somatic sensory-recipient sector of the cat's striatum. *Journal of Neuroscience*, 6(12):3436–3458, 1986.

D. A. Malone Jr, D. D. Dougherty, A. R. Rezai, L. L. Carpenter, G. M. Friehs, E. N. Eskandar, S. L. Rauch, S. A. Rasmussen, A. G. Machado, C. S. Kubu, et al. Deep brain stimulation of the ventral capsule/ventral striatum for treatment-resistant depression. *Biological Psychiatry*, 65(4):267–275, 2009.

W. R. Marchand and D. Yurgelun-Todd. Striatal structure and function in mood disorders: a comprehensive review. *Bipolar Disorders*, 12(8):764–785, 2010.

S. Markett, G. Heeren, C. Montag, B. Weber, and M. Reuter. Loss aversion is associated with bilateral insula volume. a voxel based morphometry study. *Neuroscience letters*, 619:172–176, 2016.

A. Marschner, T. Mell, I. Wartenburger, A. Villringer, F. M. Reischies, and H. R.

Heekeren. Reward-based decision-making and aging. *Brain research bulletin*, 67 (5):382–390, 2005.

D. Martinez, R. Gil, M. Slifstein, D.-R. Hwang, Y. Huang, A. Perez, L. Kegeles, P. Talbot, S. Evans, J. Krystal, et al. Alcohol dependence is associated with blunted dopamine transmission in the ventral striatum. *Biological Psychiatry*, 58 (10):779–786, 2005.

L. Marzetti, G. Nolte, M. G. Perrucci, G. L. Romani, and C. Del Gratta. The use of standardized infinity reference in eeg coherency studies. *Neuroimage*, 36(1):48–63, 2007.

A. Mashhoori, S. Hashemnia, B. L. McNaughton, D. R. Euston, and A. J. Gruber. Rat anterior cingulate cortex recalls features of remote reward locations after disfavoured reinforcements. *ELife*, 7:e29793, 2018.

B. Masimore, N. C. Schmitzer-Torbert, J. Kakalios, and A. D. Redish. Transient striatal $\gamma$ local field potentials signal movement initiation in rats. *Neuroreport*, 16 (18):2021–2024, 2005.

M. Maskarinec, G. Shipley, M. Novotny, D. Brown, and R. Forney. Endocrine effects of cannabis in male rats. *Toxicology and applied pharmacology*, 45(2):617–628, 1978.

J. B. Mattingley, M. Husain, C. Rorden, C. Kennard, and J. Driver. Motor role of human inferior parietal lobe revealed in unilateral neglect patients. *Nature*, 392 (6672):179–182, 1998.

R. J. McDonald and N. M. White. A triple dissociation of memory systems: hippocampus, amygdala, and dorsal striatum. *Behavioral Neuroscience*, 2013.

R. J. McDonald, A. L. King, and N. S. Hong. Neurotoxic damage to the dorsomedial striatum exaggerates the behavioral influence of a context-specific inhibitory association mediated by the ventral hippocampus. *Behavioral Neuroscience*, 122(1):27, 2008.

J. T. McGuire, M. R. Nassar, J. I. Gold, and J. W. Kable. Functionally dissociable influences on learning rate in a dynamic environment. *Neuron*, 84(4):870–881, 2014.

L. W. Means and R. D. Holsten. Individual aged rats are impaired on repeated reversal due to loss of different behavioral patterns. *Physiology & behavior*, 52(5): 959–963, 1992.

M. H. Meier, A. Caspi, A. Ambler, H. Harrington, R. Houts, R. S. Keefe, K. McDonald, A. Ward, R. Poulton, and T. E. Moffitt. Persistent cannabis users show neuropsychological decline from childhood to midlife. *Proceedings of the National Academy of Sciences*, 109(40):E2657–E2664, 2012.

J. W. Mink. The basal ganglia: focused selection and inhibition of competing motor programs. *Progress in neurobiology*, 50(4):381–425, 1996.

M. Mishkin, L. G. Ungerleider, and K. A. Macko. Object vision and spatial vision: two cortical pathways. *Trends in Neurosciences*, 6:414–417, 1983.

J. T. Mitchell, M. M. Sweitzer, A. M. Tunno, S. H. Kollins, and F. J. McClernon. "i use weed for my adhd": a qualitative analysis of online forum discussions on cannabis use and adhd. *PloS one*, 11(5):e0156614, 2016.

C. Mokrysz, R. Landy, S. H. Gage, M. R. Munafò, J. P. Roiser, and H. V. Curran.

Are iq and educational outcomes in teenagers related to their cannabis use? a prospective cohort study. *Journal of psychopharmacology*, 30(2):159–168, 2016.

T. Mullen. Nitrc: Cleanline: Tool/resource info, 2012. Available online at: http://www.nitrc.org/projects/cleanline.

T. Mullen, C. Kothe, Y. M. Chi, A. Ojeda, T. Kerth, S. Makeig, G. Cauwenberghs, and T.-P. Jung. Real-time modeling and 3d visualization of source dynamics and connectivity using wearable eeg. In *2013 35th annual international conference of the IEEE engineering in medicine and biology society (EMBC)*, pages 2184–2187. IEEE, 2013.

C. A. Munro, M. E. McCaul, D. F. Wong, L. M. Oswald, Y. Zhou, J. Brasic, H. Kuwabara, A. Kumar, M. Alexander, W. Ye, et al. Sex differences in striatal dopamine release in healthy adults. *Biological Psychiatry*, 59(10):966–974, 2006.

E. A. Murray, S. P. Wise, and S. E. Rhodes. What can different brains do with reward. *Neurobiology of sensation and reward*, pages 61–98, 2011.

M. R. Nassar, R. C. Wilson, B. Heasly, and J. I. Gold. An approximately bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *Journal of Neuroscience*, 30(37):12366–12378, 2010.

L. Nestor, R. Hester, and H. Garavan. Increased ventral striatal bold activity during non-drug reward anticipation in cannabis users. *Neuroimage*, 49(1):1133–1143, 2010.

S. Nieuwenhuis, K. R. Ridderinkhof, J. Blom, G. P. Band, and A. Kok. Error-related

brain potentials are differentially related to awareness of response errors: Evidence from an antisaccade task. *Psychophysiology*, 38(5):752–760, 2001.

R. G. Northcutt. Forebrain evolution in bony fishes. *Brain research bulletin*, 75(2-4): 191–205, 2008.

M. Nowak and K. Sigmund. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game. *Nature*, 364(6432):56–58, 1993.

M. Nowak, C. Zich, and C. J. Stagg. Motor cortical gamma oscillations: What have we learnt and where are we headed? *Current behavioral neuroscience reports*, 5(2): 136–142, 2018.

S. A. Oberg, G. J. Christie, and M. S. Tata. Problem gamblers exhibit reward hypersensitivity in medial frontal cortex during gambling. *Neuropsychologia*, 49 (13):3768–3775, 2011.

R. C. Oldfield et al. The assessment and analysis of handedness: the edinburgh inventory. *Neuropsychologia*, 9(1):97–113, 1971.

J. Olds. Hypothalamic substrates of reward. *Physiological reviews*, 42(4):554–604, 1962.

D. S. Olton, C. Collison, and M. A. Werz. Spatial memory and radial arm maze performance of rats. *Learning and motivation*, 8(3):289–314, 1977.

J. R. S. Onge, H. Abhari, and S. B. Floresco. Dissociable contributions by prefrontal d1 and d2 receptors to risk-based decision making. *Journal of Neuroscience*, 31 (23):8625–8633, 2011.

A. M. Owen, K. M. McMillan, A. R. Laird, and E. Bullmore. N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Human Brain Mapping*, 25(1):46–59, 2005.

M. G. Packard and N. M. White. Dissociation of hippocampus and caudate nucleus memory systems by posttraining intracerebral injection of dopamine agonists. *Behavioral Neuroscience*, 105(2):295, 1991.

M. G. Packard, R. Hirsh, and N. M. White. Differential effects of fornix and caudate nucleus lesions on two radial maze tasks: evidence for multiple memory systems. *Journal of Neuroscience*, 9(5):1465–1472, 1989.

C. A. Palencia and M. E. Ragozzino. The contribution of nmda receptors in the dorsolateral striatum to egocentric response learning. *Behavioral Neuroscience*, 119(4):953, 2005.

J. A. Palmer, K. Kreutz-Delgado, and S. Makeig. Amica: An adaptive mixture of independent component analyzers with shared components. *Swartz Center for Computatonal Neursoscience, University of California San Diego, Tech. Rep*, 2012.

W. X. Pan, T. Mao, and J. T. Dudman. Inputs to the dorsal striatum of the mouse reflect the parallel circuit architecture of the forebrain. *Frontiers in neuroanatomy*, 4:147, 2010.

A.-A. Pape and M. Siegel. Motor cortex activity predicts response alternation during sensorimotor decisions. *Nature communications*, 7:13098, 2016.

A. Parent and L.-N. Hazrati. Functional anatomy of the basal ganglia. i. the cortico-basal ganglia-thalamo-cortical loop. *Brain Research Reviews*, 20(1):91–127, 1995.

D. M. Parslow, D. Rose, B. Brooks, S. Fleminger, J. A. Gray, V. Giampietro, M. J. Brammer, S. Williams, D. Gasston, C. Andrew, et al. Allocentric spatial memory activation of the hippocampal formation measured with fmri. *Neuropsychology*, 18 (3):450, 2004.

R. D. Pascual-Marqui. Discrete, 3d distributed, linear imaging methods of electric neuronal activity. part 1: exact, zero error localization. *arXiv preprint arXiv:0710.3341*, 2007.

R. D. Pascual-Marqui, C. M. Michel, and D. Lehmann. Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. *International Journal of Psychophysiology*, 18(1):49–65, 1994.

R. D. Pascual-Marqui, P. L. Faber, T. Kinoshita, K. Kochi, P. Milz, K. Nishida, and M. Yoshimura. Comparing eeg/meg neuroimaging methods based on localization error, false positive activity, and false positive connectivity. *bioRxiv*, page 269753, 2018.

R. D. Pascual-Marqui et al. Standardized low-resolution brain electromagnetic tomography (sloreta): technical details. *Methods Find Exp Clin Pharmacol*, 24(Suppl D):5–12, 2002.

R. E. Passingham and S. P. Wise. *The neurobiology of the prefrontal cortex: anatomy, evolution, and the origin of insight*. Number 50 in Oxford Psychology Series. Oxford University Press, 2012.

G. C. Patton, C. Coffey, J. B. Carlin, L. Degenhardt, M. Lynskey, and W. Hall.

Cannabis use and mental health in young people: cohort study. *Bmj*, 325(7374): 1195–1198, 2002.

M. P. Paulus, N. Hozack, L. Frank, and G. G. Brown. Error rate and outcome predictability affect neural activation in prefrontal cortex and anterior cingulate during decision-making. *Neuroimage*, 15(4):836–846, 2002a.

M. P. Paulus, N. E. Hozack, B. E. Zauscher, L. Frank, G. G. Brown, D. L. Braff, and M. A. Schuckit. Behavioral and functional neuroimaging evidence for prefrontal dysfunction in methamphetamine-dependent subjects. *Neuropsychopharmacology*, 26(1):53, 2002b.

M. P. Paulus, J. S. Feinstein, D. Leland, and A. N. Simmons. Superior temporal gyrus and insula provide response and outcome-dependent information during assessment and action selection in a decision-making situation. *Neuroimage*, 25(2):607–615, 2005.

D. M. Pfabigan, J. Alexopoulos, H. Bauer, and U. Sailer. Manipulation of feedback expectancy and valence induces negative and positive reward prediction error signals manifest in event-related brain potentials. *Psychophysiology*, 48(5):656–664, 2011.

D. M. Pfabigan, E.-M. Seidel, R. Sladky, A. Hahn, K. Paul, A. Grahl, M. Küblböck, C. Kraus, A. Hummer, G. S. Kranz, et al. P300 amplitude variation is related to ventral striatum bold response during gain and loss anticipation: an eeg and fmri experiment. *NeuroImage*, 96:12–21, 2014.

L. Pion-Tonachini, K. Kreutz-Delgado, and S. Makeig. Iclabel: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage*, 198:181–197, 2019.

E.-M. Pool, A. K. Rehme, G. R. Fink, S. B. Eickhoff, and C. Grefkes. Handedness and effective connectivity of the motor system. *Neuroimage*, 99:451–460, 2014.

H. G. Pope, A. Jacobs, J.-P. Mialet, D. Yurgelun-Todd, and S. Gruber. Evidence for a sex-specific residual effect of cannabis on visuospatial memory. *Psychotherapy and psychosomatics*, 66(4):179–184, 1997.

K. L. Possin, H. Kim, M. D. Geschwind, T. Moskowitz, E. T. Johnson, J. S. Sharon, A. Apple, D. Xu, B. L. Miller, S. Finkbeiner, et al. Egocentric and allocentric visuospatial working memory in premotor huntington's disease: A double dissociation with caudate and hippocampal volumes. *Neuropsychologia*, 101:57–64, 2017.

B. R. Postle and M. D'Esposito. Spatial working memory activity of the caudate nucleus is sensitive to frame of reference. *Cognitive, Affective, & Behavioral Neuroscience*, 3(2):133–144, 2003.

V. Purohit. Can alcohol promote aromatization of androgens to estrogens? a review. *Alcohol*, 22(3):123–127, 2000.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2019. URL `https://www.R-project.org/`.

M. E. Ragozzino. The contribution of the medial prefrontal cortex, orbitofrontal

cortex, and dorsomedial striatum to behavioral flexibility. *Annals of the New York academy of sciences*, 1121(1):355–375, 2007.

M. E. Ragozzino, K. E. Ragozzino, S. J. Mizumori, and R. P. Kesner. Role of the dorsomedial striatum in behavioral flexibility for response and visual cue discrimination learning. *Behavioral Neuroscience*, 116(1):105, 2002.

G. E. Reckless, O. T. Ousdal, A. Server, H. Walter, O. A. Andreassen, and J. Jensen. The left inferior frontal gyrus is involved in adjusting response bias during a perceptual decision-making task. *Brain and behavior*, 4(3):398–407, 2014.

R. Reig and G. Silberberg. Multisensory integration in the mouse striatum. *Neuron*, 83(5):1200–1212, 2014.

A. Reiner, L. Medina, and C. L. Veenman. Structural and functional evolution of the basal ganglia in vertebrates. *Brain Research Reviews*, 28(3):235–285, 1998.

B. Reuderink, C. Mühl, and M. Poel. Valence, arousal and dominance in the eeg during game play. *International journal of autonomous and adaptive communications systems*, 6(1):45–62, 2013.

E. Reutskaja and R. M. Hogarth. Satisfaction in choice as a function of the number of alternatives: When "goods satiate". *Psychology & Marketing*, 26(3):197–203, 2009.

E. Reutskaja, A. Lindner, R. Nagel, R. A. Andersen, and C. F. Camerer. Choice overload reduces neural signatures of choice set value in dorsal striatum and anterior cingulate cortex. *Nature human behaviour*, 2(12):925–935, 2018.

T. E. Robinson. Stimulant drugs and stress: factors influencing individual differences in the susceptibility to sensitization. *Sensitization of the nervous system*, pages 145–173, 1988.

O. Rogeberg. Correlations between cannabis use and iq change in the dunedin cohort are consistent with confounding from socioeconomic status. *Proceedings of the National Academy of Sciences*, 110(11):4251–4254, 2013.

H. Roitblat, W. Tham, and L. Golub. Performance ofbetta splendens in a radial arm maze. *Animal Learning & Behavior*, 10(1):108–114, 1982.

M. Rotermann. Analysis of trends in the prevalence of cannabis use and related metrics in canada. *Health reports*, 30(6):3, 2019.

F. Ruotolo, G. Ruggiero, M. Raemaekers, T. Iachini, I. van der Ham, A. Fracasso, and A. Postma. Neural correlates of egocentric and allocentric frames of reference combined with metric and non-metric spatial relations. *Neuroscience*, 409:235–252, 2019.

Y. Sakurai-Yamashita, Y. Kataoka, M. Fujiwara, K. Mine, and S. Ueki. $\delta$9-tetrahydrocannabinol facilitates striatal dopaminergic transmission. *Pharmacology Biochemistry and Behavior*, 33(2):397–400, 1989.

G. R. Samanez-Larkin and B. Knutson. Reward processing and risky decision making in the aging brain. In *The neuroscience of risky decision making.*, pages 123–142. American Psychological Association, 2014. doi: 10.1037/14322-006. URL `https://doi.org/10.1037%2F14322-006`.

R. San Martín. Event-related potential studies of outcome processing and feedback-guided learning. *Frontiers in Human Neuroscience*, 6:304, 2012.

R. San Martín, F. Manes, E. Hurtado, P. Isla, and A. Ibañez. Size and probability of rewards modulate the feedback error-related negativity associated with wins but not losses in a monetarily rewarded gambling task. *Neuroimage*, 51(3):1194–1204, 2010.

M. Sárvári, L. Deli, P. Kocsis, L. Márk, G. Maász, E. Hrabovszky, I. Kalló, D. Gajári, C. Vastagh, B. Sümegi, et al. Estradiol and isotype-selective estrogen receptor agonists modulate the mesocortical dopaminergic system in gonadectomized female rats. *Brain Research*, 1583:1–11, 2014.

M. K. Scheffers and M. G. Coles. Performance monitoring in a confusing world: error-related brain activity, judgments of response accuracy, and types of errors. *Journal of Experimental Psychology: Human Perception and Performance*, 26(1):141, 2000.

B. H. Schott, L. Niehaus, B. C. Wittmann, H. Schütze, C. I. Seidenbecher, H.-J. Heinze, and E. Düzel. Ageing and early-stage parkinson's disease affect separable neural mechanisms of mesolimbic reward processing. *Brain*, 130(9):2412–2424, 2007.

W. Schultz, P. Dayan, and P. R. Montague. A neural substrate of prediction and reward. *Science*, 275(5306):1593–1599, 1997.

B. Series. Studio encoding parameters of digital television for standard 4: 3 and wide-screen 16: 9 aspect ratios. *International Telecommunication Union, Radiocommunication Sector*, 2011.

R. Shansky, C. Glavis-Bloom, D. Lerman, P. McRae, C. Benson, K. Miller, L. Cosand, T. Horvath, and A. Arnsten. Estrogen mediates sex differences in stress-induced prefrontal cortex dysfunction. *Molecular psychiatry*, 9(5):531, 2004.

D. Sharp, V. Bonnelle, X. De Boissezon, C. Beckmann, S. James, M. Patel, and M. A. Mehta. Distinct frontal systems for response inhibition, attentional capture, and error processing. *Proceedings of the National Academy of Sciences*, 107(13): 6106–6111, 2010.

I. Skelin, R. Hakstol, J. VanOyen, D. Mudiayi, L. A. Molina, V. Holec, N. S. Hong, D. R. Euston, R. J. McDonald, and A. J. Gruber. Lesions of dorsal striatum eliminate lose-switch responding but not mixed-response strategies in rats. *European Journal of Neuroscience*, 39(10):1655–1663, 2014.

M. Smolkina, K. Morley, F. Rijsdijk, A. Agrawal, J. Bergin, E. Nelson, D. Statham, N. Martin, and M. Lynskey. Cannabis and depression: a twin model approach to co-morbidity. *Behavior genetics*, 47(4):394–404, 2017.

Y. H. Sohn and M. Hallett. Surround inhibition in human motor system. *Experimental brain research*, 158(4):397–404, 2004.

P. Sokol-Hessner, C. F. Camerer, and E. A. Phelps. Emotion regulation reduces loss aversion and decreases amygdala responses to losses. *Social cognitive and affective neuroscience*, 8(3):341–350, 2013.

D. Struik, F. Sanna, and L. Fattore. The modulating role of sex and anabolic-androgenic steroid hormones in cannabinoid sensitivity. *Frontiers in Behavioral Neuroscience*, 12:249, 2018.

C. Summerfield and K. Tsetsos. Do humans make good decisions? *Trends in cognitive sciences*, 19(1):27–34, 2015.

R. S. Sutton and A. G. Barto. *Reinforcement learning: An introduction*. MIT press, 1998.

R. S. Sutton and A. G. Barto. *Reinforcement learning: An introduction*. MIT press, 2018.

D. Swick, V. Ashley, and U. Turken. Left inferior frontal gyrus is critical for response inhibition. *BMC neuroscience*, 9(1):102, 2008.

A. Syed, P. M. Baker, and M. E. Ragozzino. Pedunculopontine tegmental nucleus lesions impair probabilistic reversal learning by reducing sensitivity to positive reward feedback. *Neurobiology of Learning and Memory*, 131:1–8, 2016.

C. Tallon-Baudry, O. Bertrand, C. Delpuech, and J. Pernier. Oscillatory $\gamma$-band (30–70 hz) activity induced by a visual search task in humans. *Journal of Neuroscience*, 17(2):722–734, 1997.

G. Tanda, F. E. Pontieri, and G. Di Chiara. Cannabinoid and heroin activation of mesolimbic dopamine transmission by a common $\mu$1 opioid receptor mechanism. *Science*, 276(5321):2048–2050, 1997.

R. Thapa and A. J. Gruber. Lesions of ventrolateral striatum eliminate lose-shift but not win-stay behaviour in rats. *Neurobiology of Learning and Memory*, 2018.

R. Thapa, C. H. Donovan, S. A. Wong, R. J. Sutherland, and A. J. Gruber. Lesions of lateral habenula attenuate win-stay but not lose-shift responses in a competitive choice task. *Neuroscience letters*, 692:159–166, 2019.

E. Thorndike. Animal intelligence: An experimental study of the associative processes in animals. new york, ny, us, 1898.

E. Thorndike. *Animal intelligence: Experimental studies*. New York: Macmillan, June 1911. doi: https://doi.org/10.5962/bhl.title.55072.

K. Uemoto, M. Yoshioka, H. Liang, and C. Zhu. Effect of motor intensity on motion imagery with electroencephalogram signal analysis in mirror neuron system. *Journal of Neuroscience and Neuroengineering*, 4(1):38–43, 2017.

E. van de Giessen, J. J. Weinstein, C. M. Cassidy, M. Haney, Z. Dong, R. Ghazzaoui, N. Ojeil, L. S. Kegeles, X. Xu, N. P. Vadhan, et al. Deficits in striatal dopamine release in cannabis dependence. *Molecular psychiatry*, 22(1):68–75, 2017.

M. A. Van Der Meer and A. D. Redish. Low and high gamma oscillations in rat ventral striatum have distinct relationships to behavior, reward, and spiking activity on a learned spatial decision task. *Frontiers in integrative neuroscience*, 3:9, 2009.

M. A. van der Meer and A. D. Redish. Theta phase precession in rat ventral striatum links place and reward information. *Journal of Neuroscience*, 31(8):2843–2854, 2011.

M. A. van der Meer, A. Johnson, N. C. Schmitzer-Torbert, and A. D. Redish. Triple dissociation of information processing in dorsal striatum, ventral striatum, and hippocampus on a learned spatial decision task. *Neuron*, 67(1):25–32, 2010.

H. H. van Hell, M. Vink, L. Ossewaarde, G. Jager, R. S. Kahn, and N. F. Ramsey. Chronic effects of cannabis use on the human reward system: an fmri study. *European neuropsychopharmacology*, 20(3):153–163, 2010.

A. Varjačić, D. Mantini, J. Levenstein, E. D. Slavkova, N. Demeyere, and C. R. Gillebert. The role of left insula in executive set-switching: Lesion evidence from an acute stroke cohort. *cortex*, 107:92–101, 2018.

M. Vavrečka and L. Lhotská. The eeg correlates of the allocentric and the egocentric spatial reference frames processing. In *World Congress on Medical Physics and Biomedical Engineering, September 7-12, 2009, Munich, Germany*, pages 2012–2015. Springer, 2009.

A. A. Vena, R. Mangieri, and R. A. Gonzales. Regional analysis of the pharmacological effects of acute ethanol on extracellular striatal dopamine activity. *Alcoholism: Clinical and Experimental Research*, 40(12):2528–2536, 2016.

J. Villares. Chronic use of marijuana decreases cannabinoid receptor binding and mrna expression in the human brain. *Neuroscience*, 145(1):323–334, 2007.

N. D. Volkow, G.-J. Wang, J. S. Fowler, D. Tomasi, F. Telang, and R. Baler. Addiction: decreased reward sensitivity and increased expectation sensitivity conspire to overwhelm the brain's control circuit. *Bioessays*, 32(9):748–755, 2010.

P. Voorn, L. J. Vanderschuren, H. J. Groenewegen, T. W. Robbins, and C. M. Pennartz. Putting a spin on the dorsal–ventral divide of the striatum. *Trends in Neurosciences*, 27(8):468–474, 2004.

J. Wacker, D. G. Dillon, and D. A. Pizzagalli. The role of the nucleus accumbens and rostral anterior cingulate cortex in anhedonia: integration of resting eeg, fmri, and volumetric techniques. *Neuroimage*, 46(1):327–337, 2009.

A. A. Wakley, J. L. Wiley, and R. M. Craft. Sex differences in antinociceptive tolerance to delta-9-tetrahydrocannabinol in the rat. *Drug and Alcohol Dependence*, 143:22–28, 2014.

Q. Wang, X. Xu, and M. Zhang. Normal aging in the basal ganglia evaluated by eigenvalues of diffusion tensor imaging. *American journal of neuroradiology*, 31(3): 516–520, 2010.

R. T. Watson, K. M. Heilman, J. C. Cauthen, and F. A. King. Neglect after cingulectomy. *Neurology*, 1973.

J. A. Weller, I. P. Levin, B. Shiv, and A. Bechara. The effects of insula damage on decision-making for risky gains and losses. *Social neuroscience*, 4(4):347–358, 2009.

S. A. Wong, S. H. Randolph, V. E. Ivan, and A. J. Gruber. Acute $\delta$-9-tetrahydrocannabinol administration in female rats attenuates immediate responses following losses but not multi-trial reinforcement learning from wins. *Behavioural Brain Research*, 335:136–144, 2017a.

S. A. Wong, R. Thapa, C. A. Badenhorst, A. R. Briggs, J. A. Sawada, and A. J. Gruber. Opposing effects of acute and chronic d-amphetamine on decision-making in rats. *Neuroscience*, 2017b.

D. A. Worthy, M. J. Hawthorne, and A. R. Otto. Heterogeneity of strategy use in the iowa gambling task: a comparison of win-stay/lose-shift and reinforcement learning models. *Psychonomic bulletin & review*, 20(2):364–371, 2013.

J. Yang, S. Zhang, Y. Lou, Q. Long, Y. Liang, S. Xie, and J. Yuan. The increased

sex differences in susceptibility to emotional stimuli during adolescence: an event-related potential study. *Frontiers in Human Neuroscience*, 11:660, 2018.

X.-h. Yang, J. Huang, Y. Lan, C.-y. Zhu, X.-q. Liu, Y.-f. Wang, E. F. Cheung, G.-r. Xie, and R. C. Chan. Diminished caudate and superior temporal gyrus responses to effort-based decision making in patients with first-episode major depressive disorder. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 64:52–59, 2016.

D. Yao. A method to standardize a reference of scalp eeg recordings to a point at infinity. *Physiological measurement*, 22(4):693, 2001.

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

H. H. Yin, B. J. Knowlton, and B. W. Balleine. Blockade of nmda receptors in the dorsomedial striatum prevents action–outcome learning in instrumental conditioning. *European Journal of Neuroscience*, 22(2):505–512, 2005.

S. W. Yip, E. E. DeVito, H. Kober, P. D. Worhunsky, K. M. Carroll, and M. N. Potenza. Pretreatment measures of brain structure and reward-processing brain function in cannabis dependence: an exploratory study of relationships with abstinence during behavioral treatment. *Drug and Alcohol Dependence*, 140:33–41, 2014.

J. E. Yonker, L.-G. Nilsson, A. Herlitz, and R. Anthenelli. Sex differences in spatial visualization and episodic memory as a function of alcohol consumption. *Alcohol and Alcoholism*, 40(3):201–207, 2005.

R. Yu, D. Mobbs, B. Seymour, and A. J. Calder. Insula and striatum mediate the default bias. *The Journal of Neuroscience*, 30(44):14702–14707, 2010.

T. Zaehle, K. Jordan, T. Wüstenberg, J. Baudewig, P. Dechent, and F. W. Mast. The neural basis of the egocentric and allocentric spatial frame of reference. *Brain Research*, 1137:92–103, 2007.

M. Zaepffel, R. Trachel, B. E. Kilavik, and T. Brochier. Modulations of eeg beta power during planning and execution of grasping movements. *PloS one*, 8(3):e60060, 2013.

G. Zheng, X. Qi, Y. Li, W. Zhang, and Y. Yu. A comparative study of standardized infinity reference and average reference for eeg of three typical brain states. *Frontiers in neuroscience*, 12:158, 2018.

M. L. Zilberman, H. Tavares, S. B. Blume, and N. El-Guebaly. Substance use disorders: sex differences and psychiatric comorbidities. *The Canadian Journal of Psychiatry*, 48(1):5–13, 2003.