INTRA- AND EXTRA-DIMENSIONAL LEARNING IN THE COMMON MARMOSET

MARIUM HUMAIRA ALVI

A THESIS SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF ARTS

GRADUATE PROGRAM IN PSYCHOLOGY
YORK UNIVERSITY
TORONTO, ONTARIO

SEPTEMBER 2025

Abstract

Cognitive flexibility is the brain's ability to suppress the current strategy in favour of a better alternative when the context changes. Impaired cognitive flexibility is a transdiagnostic deficit in several neuropsychiatric disorders. How the primate brain supports this process is not well understood. We trained 4 marmosets on a novel intra- and extra-dimensional shifting (IEDS) task involving 2 dimensions: shape and colour. They were pre-trained on each of the 6 features (3 shapes and 3 colours) separately, which were then combined into compound stimuli. Marmosets were required to select the compound stimulus with the target feature (e.g. red). After reaching criterion (8/10 correct), the target feature shifted within or across dimensions. Next, we implemented a feature reinforcement learning model, and found that the estimated learning rate, choice determinism and attentional parameters for marmosets fell within the range typically observed in healthy humans, underscoring translational value.

Acknowledgements

I am deeply grateful to my supervisor, Dr. Liya Ma, for creating an environment in which I could thrive in a niche field that was so new to me. Your consistent encouragement, support, and guidance helped me grow from a baby researcher to succeeding in this project. I continuously awe at your intelligence and understanding of the field and drive my passion to follow in suit. You made me into a Master and without your unwavering support I would not have had such a positive experience.

To my class and lab mates, thank you for the solidarity, the countless moments of venting during Friday seminars and the fun night outs that brought so much balance, joy and community along this process.

To my closest friend, Ansar Ali, my number one confidant and the person who has been my constant source of comfort throughout this journey. Thank you for always being the one I could lean on and cry to when things felt overwhelming. Your presence along this journey has been invaluable and I cannot even begin to express my gratitude for your love and calm through every storm I thought I may not make it out of.

To my family, Zaina, Bhaiya and Amma, thank you for your unwavering support and love throughout this journey. Supporting my crash outs and self-doubt by making me laugh or reinforcing that everything will work out has allowed me to succeed with a solid support system to catch me whenever I was down. Special mention to Shehryar Bhai and Meral Bhabhi (+Ariana and Amyrah) for not only being there as an ear to listen to my stresses but also as my biggest cheerleaders, always ready to celebrate my accomplishments. The amount of love I feel from your support cannot be matched.

Finally, to my dad, whose guidance and encouragement have been the foundation of my academic path – thank you for always cheering me on, reminding me of the importance of education and instilling in me the value of scientific thinking and rigor. You are the reason that I have any degrees at all. Your belief in me has made me continuously pursue my academic goals without hesitation and as a result, be successful within each step that I have taken. Thank you, Abba.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1.Cognitive Flexibility

Cognitive flexibility refers to the capacity to adjust behaviour and/or strategies of behaviour in response to changing task demands and contingencies (Miyake et al., 2000). It is a form of executive function and can be assessed and quantified using task switching and set-shifting paradigms (Dajani & Uddin, 2015). Cognitive flexibility comprises three components including shifting, updating and inhibition - which correspond to the three core subconstructs of the Research Domain Criteria Cognitive Control domain (NIMH RDoC Cognitive Control construct). Shifting refers to the ability to switch between rules/responses. Updating represents refreshing information within our working memory. Lastly, inhibition represents the countermanding of prepotent behavioural responses that have previously received reinforcement but no longer (Miyake et al., 2000). Deficits found within these three components result in inflexible adaptability to changing environments and are seen in some neuropsychiatric disorders (Miyake et al., 2000). Each component has been linked to specific regions of the frontal cortex (briefly reviewed in Section 1.3).

Reductions of cognitive flexibility have been found in the clinical realm with implications in disorders including schizophrenia, ADHD, autism spectrum disorder and mood disorders (Uddin, 2021). Cognitive inflexibility is seen within these disorders as individuals having difficulty shifting attention or strategy, and rigid behaviours (Oikonomidis et al., 2017). For instance, individuals with schizophrenia may experience perseveration characterized by the

continual application of the same rule/strategy despite receiving negative feedback (Pantelis et al., 1999). Similarly, some individuals with autism spectrum disorder find tasks that require the shifting of cognitive sets difficult (Hill, 2004). Furthering research studies to quantify and explore the disruptions within the neural circuitry underlying cognitive flexibility are crucial in understanding and evolving assessment and treatment of these disorders.

One common paradigm used in the clinical setting is known as the attentional set shifting task. In this context, a 'set' is defined as the property of the stimulus that is relevant on a given trial (Rushworth et al., 2005). For example, if colour is relevant, the set might be represented by the blue stimuli being rewarding. When feedback indicated that the current set no longer yields reward, the subject must detect that failure, update by identifying a new relevant stimulus property and then maintain that new set until performance stabilizes (Dajani & Uddin, 2015; Uddin, 2021). These paradigms distinguish two kinds of shift along stimulus dimensions, such as colour and shape. Here, a stimulus dimension refers to a feature category along which stimuli may vary with each dimension comprising a set of possible features (e.g. blue, red, yellow for colour; square, star, heart for shape) An intra-dimensional shift (IDS) involves changing from one rule to another within the same dimension (ex. the rewarding feature changing from blue to red, changing within the colour dimension). An extra-dimensional shift (EDS) requires the changing of rewarding feature to a different dimension (ex. blue to square, changing across dimensions from colour to shape). Because extra-dimensional shifts demand the reorientation of attention to a new feature class, they are expected to have a greater performance cost than intra-dimensional shifts. By comparing performance on these two shift types, researchers can isolate the cognitive and neural mechanisms that support flexible rule updating.

1.2. Set Shifting Across Species: Primate and Rodent Models

Primates

Set-shifting tasks originated from the Wisconsin Card Sorting Task (WCST). The WCST was initially used to demonstrate that frontal lobe damage in humans produced severe impairments in cognitive set-shifting ability (Milner, 1963). This led to the development of analogous tasks for non-human primates to investigate the underlying neural mechanisms. Early primate studies, such as by Settlage, Butler and Odoi (1956) examined the effects of lesions of the prefrontal cortex on a series of non-reversal shifts (Settlage et al., 1956), but were limited by their use of identical stimulus pairs across discriminations, potentially confounding true attentional set shifting with simple reversal (Dias et al., 1996).

In primate implementations of the CANTAB ID/ED task, compound stimuli typically vary along two dimensions – commonly shape (e.g. circle, square, triangle) and line orientation (e.g., horizontal vs. vertical) or shape and colour (e.g., blue vs red). Significant methodological advancements came through Roberts and colleagues' development of the CANTAB ID/ED task (Dias et al., 1996; Roberts et al., 1988), which provided a standardized platform for both humans and non-human primates. This paradigm has been applied extensively in macaques, using both touchscreen and joystick-controlled versions. In most macaque implementations, compound stimuli vary along shape and orientation, and each stage uses novel exemplars with a six-consecutive-correct rule per daily session (Dias et al., 1996; Mansouri et al., 2006). These task variations consistently have shown larger performance costs (more errors and trials to criterion) on extra-dimensional shifts compared to intra-dimensional shifts. It has been successfully implemented in marmoset monkeys using both manual and automated apparatus. Intra- and extra- dimensional set shifting tasks introduced the use of novel exemplars at each

testing stage, ensuring that performance reflected attentional shifting rather than stimulus specific learning.

This task progresses through a specific sequence of stages, each designed to probe a different cognitive process. Initially, simple and compound discriminations, establish a baseline associative learning and the ability to maintain a rule while ignoring irrelevant information. The paradigm then introduces an intra-dimensional shift (IDS), where new stimuli are presented but the relevant dimension remains the same. Successful performance on the IDS stage demonstrates the animal's ability to form and maintain an attentional set (Dias et al., 1996; Roberts et al., 1988). The core test for rule switching comes with the extradimensional shift (EDS), where the previously irrelevant dimension becomes the rewarding component, forcing the animal to release the established attentional set and shift to a new one. The performance cost on the EDS, typically an increase in errors or trials to criterion, serves as the primary measure of cognitive inflexibility. This is contrasted with discrimination reversal stages, where the stimulus reward contingencies are reversed within the same dimension, testing a different form of flexibility related to updating associations rather than shifting attentional dimensions (Dias et al., 1997). Detailed lesion and recording findings in non-human primates are reviewed in Section 1.3.

The primate attentional set shifting paradigm has provided valuable insights into the functional organization of the frontal lobes and offers a robust framework for investigating cognitive processes and deficits seen in human neuropsychiatric disorders.

Rodents

The translational power of this methodology is further evidenced by its successful adaptation in rodents, which confirmed that the medial prefrontal cortex in rats serves a function analogous to the primate lateral PFC in mediating set-shifting (Birrell & Brown, 2000). The

stimulus dimensions typically correspond to sensory modalities or features such as odour (e.g. mint vs. cinnamon) and texture of the digging medium (e.g. sand vs. gravel), in contrast to the primarily visual dimensions used in primate tasks. Early investigations of rat set-shifting used simple two-choice discriminations, demonstrating that rodents can form attentional sets and shift between stimulus dimensions with varying degrees of difficulty depending on the nature of the shift (Birrell & Brown, 2000; Ragozzino et al., 1999). Birrell and Brown (2000) introduced the bowl-digging paradigm as an analogue to the IDS/EDS task in CANTAB for primates. As in CANTAB, the task progresses through several stages including simple discrimination (SD), compound discrimination (CD), reversal learning, intra-dimensional shift (IDS) and extra-dimensional shift (EDS) (Birrell & Brown, 2000; Brown et al., 2007). This paradigm has proven particularly valuable as rats can learn novel discriminations based on odours and digging medium with short training durations compared to training rats on visual discriminations (Tait et al., 2014).

In parallel, maze-based procedures were developed where rats switched between spatial and response discriminations in a plus maze (Rich and Shapiro 2007, 2009; Floresco et al., 2008; Ragozzino et al., 1999). While these rule-switching tasks do not employ compound visual stimuli present in primate attentional set shifting tasks, the strategy-shifting variations had more similarity to the Wisconsin Card Sorting Task (Floresco et al., 2006). In the plus maze paradigm, the stimuli remain identical between 'spatial' (allocentric) and 'response' (egocentric) strategies.

More recently, automated operant procedures have been developed to address some limitations of manual testing methods. These procedures employ two choice discriminations in operant chambers, where rats must learn rules based on either visual cues (stimulus lights) or spatial locations (lever position) (Brady & Floresco, 2015; Floresco et al., 2008). The automated

approach offers several advantages such as a standardized stimulus presentation, minimal experimenter interaction and the ability to test multiple animals simultaneously. In these tasks, rats are typically trained to respond to either visual cues, such as pressing the lever with an illuminated stimulus light above it, or spatial locations, such as always pressing either the left or right lever regardless of stimulus light position. The implementation of automated tasks has allowed researchers to identify distinct types of errors: 1. perseverative errors (continued use of the old strategy), 2. regressive errors (returning to the old strategy after initially shifting) and 3. never reinforced errors (novel incorrect responses). This error analysis has revealed how different prefrontal regions make distinct contributions to flexibility (Floresco et al., 2008; Ghods-Sharifi et al., 2008).

A key difference between automated and traditional tasks lies in how the shifts are implemented. In the bowl digging paradigm, novel stimuli are presented during each shift, ensuring that performance deficits reflect difficulties in shifting attention rather than simple response inhibition. In contrast, automated tasks often use the same stimuli across different phases, making them more analogous to the Wisconsin Card Sorting Task.

Studies using the digging paradigm have revealed crucial dissociations between brain regions. Lesions of the medial prefrontal cortex (mPFC) specifically impair extra-dimensional shifts while leaving reversal learning intact, whereas orbitofrontal cortex (OFC) lesions produce the opposite pattern, selectively impairing reversal learning (McAlonan & Brown, 2003; Tait et al., 2014). Comparison of rodent and primate homologies appear in Section 1.3.

Research using the plus-maze strategy shifts has demonstrated that the inactivation of the mPFC selectively impairs shifting between strategies but not initial rule learning or reversal learning. These impairments are characterized by perseverative responding to the previously

relevant strategy (Block et al., 2007; Ragozzino & Rozman, 2007). Furthermore, studies revealed that distinct neural circuits contribute to different aspects of strategy shifting, such as the mediodorsal thalamic-PFC circuit mediating strategy suppression while the nucleus accumbens facilitating new strategy acquisition (Block et al., 2007; Floresco et al., 2006).

Investigations with operant chambers have shown that disruptions to corticostriatal loops, either via striatal inactivation or dopaminergic manipulation, increases perseverative errors on strategy shifts, while cholinergic modulation in the mPFC affects trials to criterion for EDS (Brady & Floresco, 2015; Ragozzino et al., 2002).

Lesion and inactivation work across digging, maze and operant tasks point to the same regions: the rodent medial PFC (prelimbic and ACC), mediodorsal thalamus and striatum. Each task tends to emphasize one or two of these areas. For example, ACC involvement in rule failure detection during digging tasks or striatal contributions to perseveration control in operant shifts and plus-maze tasks (Birrell & Brown, 2000; Block et al., 2007; Floresco et al., 2008). This highlights a shared cortico-thalamo-striatal circuit for flexibility. This overlapping network suggests that despite differences in sensory modality or task demands, the brain recruits a common pathway to break old rules and establish new ones.

1.3. Neural Mechanisms of Set Shifting

Research in both rodents and primates indicates that attentional set shifting depends on a coordinated network of cortical and subcortical regions. Within the frontal cortex, the dorsolateral prefrontal cortex (dlPFC) is critical for encoding and implementing new rules and is unique to primates with no clear anatomical homologue in rodents (Laubach et al., 2018). The anterior cingulate cortex (ACC) monitors conflict and signals when a current rule no longer applies, although it is a homologous structure across species and may contribute more directly to

set shifting in rodents, whereas its role is debated in primates (Heilbronner & Hayden, 2016; Laubach et al., 2018). The posterior parietal cortex (PPC) reallocates attention toward the newly relevant dimension and is part of a frontoparietal control network that is well developed in primates but not clearly represented in rodents (Scolari et al., 2015). These cortical mechanisms operate with corticostriatal loops, where dopaminergic prediction errors update rule representations in the striatum. The mediodorsal thalamus synchronizes rule related activity across the frontal regions and is essential for maintaining goal directed behaviour across shifts (Halassa & Kastner, 2017). Finally, neuromodulators including dopamine, acetylcholine and noradrenaline shape the speed and stability of rule updating by influencing prefrontal dynamics and cognitive control (Vijayraghavan et al., 2017).

1.3.1 Dorsolateral Prefrontal Cortex

Cognitive functions requiring flexibility and adaptability in non-human primates are critically linked to the dorsolateral region within the prefrontal cortex. Lesion studies in marmosets demonstrate that the dorsolateral prefrontal cortex (dlPFC) is essential for extra-dimensional set shifting (Dias et al., 1996). These findings suggest that the dlPFC plays a critical role in maintaining and shifting abstract attentional sets across dimensions. Rather than merely detecting when a shift is needed, the dlPFC may contribute to representing the currently relevant rule or dimension and flexibly updating its value when contingencies change.

Resting state fMRI in rhesus monkeys identified the dlPFC as a hub within a frontostriatal executive control network where connectivity strength correlates with individual set shifting performance (Grant et al., 2022). Electrophysiological recordings in macaques performing a WCST demonstrates that dlPFC neurons encode both the identity of the newly relevant rule and reward prediction errors immediately following an extra-dimensional shift, and

that the magnitude of these phasic signals correlates with trials to criterion (Mansouri et al., 2006).

1.3.2. Anterior Cingulate Cortex

The anterior cingulate cortex plays a role in integrating performance feedback, monitoring conflict between competing rule representations and adjusting cognitive control to support flexible set shifting (Hampshire & Owen, 2006). Dosenbach et al. (2006) showed that at rest the anterior cingulate cortex is connected with the anterior insula and thalamus as part of a 'cingulo-opercular" network that maintains task sets over time. This network also interacts with frontoparietal regions (dlPFC and posterior parietal cortex) to coordinate when rules must be held or updated during set shifting. In humans performing the CANTAB intra- and extra-dimensional paradigm, ACC activation peaks on extra-dimensional shift errors, and the magnitude of the error related signal predicts the shifts success, consistent with a conflict monitoring function (Hampshire & Owen, 2006). In macaques, single unit recordings show that anterior cingulate neurons encode feedback related signals and monitor performance during rule switching tasks, including increased activity in response to errors and rule conflict trials (Nakamura et al., 2005).

In rodents, functionally homologous regions in the medial frontal cortex, including the ACC, prelimbic and infralimbic cortices, also contribute to flexible behaviour though their functions are not directly equivalent to those in primates. Reversible inactivation of the ACC selectively impairs performance on extra-dimensional set shift without affecting intra-dimensional shifts. This confirms the anterior cingulate cortex's necessity for detecting rule failure and triggering set shifts (Ragozzino & Rozman, 2007). Excitotoxic lesions of medial frontal regions including ACC disrupts the retention of a newly shifted response but do not

impair its initial acquisition, further indicating the ACC's importance in maintaining the updated rule (Rich & Shapiro, 2007). Inactivation of prelimbic and infralimbic cortex in rats shows that these medial frontal subregions contribute differentially to extra-dimensional shifts (Floresco et al., 2008). Prelimbic inactivation produced a pronounced impairment when rats had to shift to the more demanding extra-dimensional shift, requiring more trials to reach criterion. Infralimbic inactivation also impaired performance on the more difficult shift but to a significantly lesser extent than prelimbic inactivation. Infralimbic and prelimbic lesions showed impairments in shifting ability but not in learning discriminations (Oualian & Gisquet-Verrier, 2010).

Together these findings demonstrate that the ACC can help detect when a rule no longer applies and signals executive circuits such as the dlPFC to release prior rules and establish new attentional sets (Ragozzino & Rozman, 2007; Seamans & Floresco, 2022). Its role within this network ensures that performance feedback is integrated into behavioural adjustments, making the ACC essential for flexible set shifting (Dosenbach et al., 2006a; Hampshire & Owen, 2006).

1.3.3. Posterior Parietal Cortex (PPC)

Resting state fMRI in humans places the intraparietal sulcus (IPS) within the frontoparietal control network alongside the dlPFC (Cole et al., 2013). In task-based fMRI, IPS BOLD signal increases during extra-dimensional shifts, and the magnitude of this response predicts both error rates and reaction time costs. This implicated the PPC in reorienting attention to newly relevant stimulus dimensions (Kim et al., 2012). In macaques, single-unit recordings in the posterior parietal cortex reveal that neurons show transient increases in firing before EDS shifts that predict successful rule reconfiguration, supporting its role in shifting cognitive sets (Kamigaki et al., 2009). Although direct lesion or inactivation work in primates is sparse, rodent studies show that excitotoxic or mechanical lesions of the PPC impair performance on extra-

dimensional set shifting tasks in digging paradigms while sparing intra-dimensional shifts. This confirms a conserved role for the PPC in attentional reallocation during set shifting (Bucci, 2009). These findings indicate that the PPC contribute to the set shifting process by enabling efficient switching of the attentional focus to the new rule dimensions.

1.3.4. Basal Ganglia & Thalamus

Cognitive flexibility depends not only on frontoparietal cortical circuits but also on subcortical structures such as the basal ganglia and thalamus, which form integrated cortico-striato-thalamo-cortical loops. These circuits rely on excitatory (glutaminergic) and inhibitory (GABAergic) transmission to support the selection and maintenance of rule representations in prefrontal cortex.

In parallel, dopaminergic neurons in the midbrain broadcast reward prediction error signals that modulate plasticity in the striatum and frontal cortex, influencing the updating of rule values and guiding adaptive behaviour (Schultz et al., 1997). Pharmacological inactivation of the striatum increases perseverative errors during strategy switching demonstrating its importance in interrupting response patterns and supporting behavioural flexibility (Block et al., 2007).

The mediodorsal thalamus plays a central role in synchronizing activity across prefrontal areas during rule shifts and maintaining task relevant information over time (Mitchell & Chakraborty, 2013). Lesions to the mediodorsal thalamus impair performance on extra-dimensional shifts and disrupt flexible behaviour (Ouhaz et al., 2022).

1.3.5. Neuromodulation

Dopamine, acetylcholine and noradrenaline each contribute to flexible cognitive control by modulating prefrontal circuits during set shifting. In the rodent mPFC , mesocortical dopamine supports the flexible updating of stimulus reward contingencies. Blocking D1

receptors in the prelimbic cortex increases perseverative errors on extra-dimensional shifts while sparing intra-dimensional shifts and initial discrimination, suggesting D1 signalling is critical for interrupting prior rules and reconfiguring attention (Ragozzino et al., 2002). Complementary evidence shows that impairments in set shifting also emerge among D2 antagonism or D4 antagonism in medial prefrontal cortex, indicating that multiple dopamine receptor subtypes contribute to the modulation of rule switching (Floresco et al., 2006).

Cholinergic modulation of the prelimbic cortex also shapes attentional flexibility. In rodent studies, selective removal of cholinergic input from the basal forebrain to the mPFC using 192-IgG-saporin increases trials to criterion during ED shifts, while sparing reversal learning (Birrell & Brown, 2000; Chudasama et al., 2004; Dalley et al., 2004). This pattern aligns with a reduction in learning rate observed in feature-based RL models, suggesting acetylcholine may regulate the speed at which feature value associations are updated.

Similarly, noradrenergic contributes to set shifting by facilitating control when unexpected changes arise. Targeted depletion of noradrenaline from the mPFC using dopamine-β-hydroxylase-saporin selectively impairs ED shift performance, indicating that phasic noradrenergic signalling supports cognitive flexibility under changing task demands (McGaughy et al., 2008; Newman et al., 2008), whereas enhancing noradrenergic tone with atomoxetine improves EDS performance (Totah et al., 2015).

These finding suggest that dopamine, acetylcholine and noradrenaline converge on prefrontal circuits to support set shifting. At a computational level they may regulate key reinforcement learning parameters, such as theta mediated updating of feature value estimates (learning rate) and beta-mediated control over choice determinism.

1.4. Marmoset Model

The use of the common marmoset (*Callithrix jacchus*) as an ideal model to decipher the mechanisms through which inter-areal communication supports cognitive flexibility. As a diurnal primate, marmosets rely on their vision and have cingulo-insular and frontoparietal functional networks homologous to those of macaques and humans (Schaeffer et al., 2019). Their quick maturation, high fecundity and short lifespan offer unique potentials not found in other primates for creating genetic and developmental models for neuropsychiatric diseases and bridging the gap between preclinical research in rodents and clinical studies (Johnston et al., 2019; Oikonomidis et al., 2017; Prins et al., 2017).

Unlike other primates, the marmoset's smooth cerebral cortex permits laminar analysis of neural activity across all LPFC and PPC areas. Recently, laminar differences in the marmoset LPFC have been reported in task-related spiking and oscillatory activities and in rule coding through inter-laminar interactions in oscillations (Johnston et al 2019). Furthermore, deep structured such as the mediodorsal (MD) thalamus can be readily accessed with high-density probes due to the small size of the marmoset brain, unlike the larger macaque.

1.5. Reinforcement-learning accounts of rule switching

While lesion and recording studies illuminate where set shifting occurs, computational models clarify how subjects learn new rules from feedback. Reinforcement learning (RL) frameworks provide an approach to understanding how subjects detect and adapt to changing reward contingencies, making them particularly relevant for modelling rule-switching behaviour. At the core of RL theory is the concept of a reward prediction error (δ), defined as the difference between the reward received on a given trial and the reward that was expected based on prior

experience (Schultz et al., 1997). Early neurophysiological recordings in non-human primates demonstrated that phasic activity of midbrain dopamine neurons can encode this prediction error. Unexpected rewards elicit transient bursts of dopaminergic firing, whereas absence of expected rewards lead to brief dips in firing rate (Bayer & Glimcher, 2005). These findings support phasic dopamine as a neural correlate for δ that drives trial-by-trial adjustments of outcome expectation.

More recent computational work has extended these principles to explain how animals and humans adapt to shifting rules in multi-dimensional environments. Hierarchical and feature-based RL models show that learning involves not only updating action values but also allocating attention across competing stimulus dimensions (Farashahi et al., 2017; Niv et al., 2015; Wilson & Niv, 2011). In these accounts, flexible behaviour emerges when attention is dynamically biased towards features that better predict reward. Neuroimaging evidence aligns with this view, showing that RL signals guide attentional selection processes (Leong et al., 2017). Interactions between reinforcement driven learning and attentional control provide the basis for adapting in uncertainty (Radulescu et al., 2019). These findings indicate that rule switching reflects a combination of reinforcement learning principles and attentionally mediated computations offering an account that links behaviour with the neural mechanisms that support cognitive flexibility.

1.6. Current Study

Previous literature has found that the prefrontal cortex and executive function critically manages rule switching cognition and behaviour in humans and animal models. FMRI studies have confirmed that the Wisconsin Card Sorting Task for assessments of cognitive flexibility implicate regions of the prefrontal and orbitofrontal cortices regarding set shifting and reversal learning (Hampshire & Owen, 2006). This neural coding seen within these studies supports the

ability to switch between cognitive ‘sets’, an integral component of cognitive flexibility (Hampshire & Owen, 2006). However, despite extensive research in rodents and humans, the mechanistic understanding of how cognitive flexibility is supported by inter-areal interactions in the primate brain remains limited. To learn how neuronal dynamics in the brain network support flexibility, rule switches must be achieved more than once within a recording session, so confounding factors such as the fatigue, practice effect, and the presence/anticipation of reward can be controlled for. This is not possible using classical paradigms such as the CANTAB intra-/extra-dimensional shift task.

To bridge this gap in knowledge, this study had two objectives: (1) Training marmosets on a novel, touchscreen-based Wisconsin Card Sorting Task (WCST) to successfully shift rules multiple times within session, and (2) Implement a reinforcement learning model to simulate rule switching behaviour. Combined with such a paradigm, future recordings from large, simultaneously active ensembles of neurons will then reveal the flexible updating of rule representations on a trial-by-trial basis.

1.7. Structure of this thesis

Following the introduction above, the remainder of this thesis is organized as follows:

Chapter 2 describes the development of an intra- and extra-dimensional learning paradigm for the common marmoset. This chapter presents the methods used to train animals on multiple feature-based rule switches and reports behavioural results demonstrating their ability to perform both intra-and extra-dimensional shifts.

Chapter 3 introduces reinforcement learning models as a framework for understanding rule switching behaviour. Here, I outline the specifications of different feature-based RL models,

test their ability to simulate marmoset behaviour and compare estimated parameter values to those reported in humans. This chapter evaluated the extent to which model-based analyses can capture both observed errors and underlying mechanisms of cognitive flexibility.

Chapter 4 provides a general discussion of the findings. I summarize the main results from both aims, interpret them in the context of existing primate literature and highlight implications for understanding the neural and computational mechanisms of shifting. This chapter also considers limitations of the present work and suggests directions for future studies that could bridge behavioural and electrophysiological approaches in marmosets.

Chapter 2

Aim 1: Establishing an intra- and extra-dimensional learning task in the marmoset for electrophysiological studies

2.1. Background

In the current study, we designed and trained marmosets on a new intra- and extra-dimensional shift (IEDS) task which is optimized for electrophysiological studies. This task is analogous to the Wisconsin Card Sorting Task (WCST), a neuropsychological test which has been used reliably to measure cognitive flexibility through testing rule acquisition and reversal (Grant & Berg, 1948). The WCST entails the sorting of cards picturing geometric figures based on shape, colour or number, the correct sorting rule is found through trial and error. Once established, the rule is referred to as a cognitive set. The sorting rule is then changed, characterizing a cognitive set-shift. The set-shift is achieved based on feedback received from

correct and incorrect sorting. The reinforcement learning framework depicts how behaviour is adapted in response to positive and negative feedback (Steinke et al., 2020). Intra-dimensional shift is between different features within the same category, such as shifting from a black square to a black heart to a black circle with shape being the overarching category. By contrast, an extra-dimensional shift crosses categories. For instance, the ability to switch from square to red, as the shift includes a dimensional strategic change from shape to colour. This study employs a simplified version of the WCST that allows for multiple target shifts within a session, both intra- and extra-dimensionally, thereby paves the way for the analysis of neurophysiological underpinnings of cognitive flexibility in non-human primates.

2.2. Hypothesis

Aim 1 evaluates how shift type, intra- vs. extra-dimensional, influences marmoset set shifting performance. We formalize two primary questions and their associated null and alternative hypotheses:

1. Effect of shift type on trials to criterion

   We predict that extra-dimensional shifts will require more trials to reach criterion than intra-dimensional shifts.

   H_0: There is no difference in mean trials to criterion following an intra- and extra-dimensional shift.

   H_1: Marmosets take significantly more trials to reach criterion following extra-dimensional shifts than intra-dimensional shifts.

2. Transfer effect on extra-dimensional cost

We predict that the classic extra-dimensional shift cost (more trials to criterion after EDS than IDS) will emerge in the New stimulus set after being absent in the original stimuli set.

H_0: In both the old and new stimuli sets, there is no difference in mean trials to criterion following EDS versus IDS.

H_1: In the new stimuli set, mean trials to criterion after EDS will be significantly greater than IDS, whereas no such difference will be observed in the original set.

We tested these hypotheses by calculating mean trials to criterion for IDS and EDS separately in the old and new stimulus sets. For hypothesis 1, we compared IDS versus EDS trials to criterion across all sessions. For hypothesis 2, we conducted paired comparisons of EDS versus IDS within each stimulus set to determine whether an extra-dimensional cost emerges in the new set but remains absent in the old set.

2.3. Methods

2.3.1. Animals

Four male common marmosets, Callithrix jacchus, ages 2 to 3 years, were used in this study. Marmosets M and T were housed together in one cage and Marmosets K and F were housed together in one cage within view of one another. All procedures were carried out in accordance with the Canadian Council for Animal Care policies on care and use of laboratory animals, following Animal Use Protocol 2023-01 as approved by the Animal Care Committee of York University.

2.3.2 Apparatus

A custom touchscreen box was built for this study. This custom-built box was equipped with an Elo 10" screen, controlled by a Raspberry Pi 4 Model B computer that was built to be attached to the home cage during training. The box contained an infusion pump to deliver viscous reward of the animal's preference (diluted acacia gum or diluted marshmallow fluff). Two high-definition Pi cameras with resolution 1280 * 720 and a frame rate of 60 were installed within the box and recorded marmoset behaviour from two different angles (above and side viewpoints). The administration of the task was controlled by software custom-written in Python, using PsychoPy, PyGame and other Python libraries.

2.3.3. Task Training

The custom touchscreen box was attached to the home cage. The door of the cage leading into the box is opened acting as an invitation for the marmoset to enter the transfer box area to enter voluntarily, one at a time and the door is closed upon entrance. Another door separating the box from the cage was used as an additional transparent barrier to prevent physical interaction during training but maintained each marmoset's ability to see and hear their cage mate.

Step 1 - Pretraining: marmosets were trained to touch a large, filled circle on the screen that covered almost the whole width of the screen. A small amount of reward was placed on the screen to motivate marmosets to touch the screen. The stimulus was presented for 10s per trial followed by an 1-s inter-trial interval (ITI).

Step 2 - Varying locations: Once marmosets successfully learned to touch the filled circle, they learned to touch a new target stimulus, a 'flower' pattern consisting of black and white geometric shapes, that may appear at middle, left and right locations on the touchscreen.

Step 3 - ITI training: With the same design as in Step 2, marmosets learned to refrain from touching the screen during the 2s ITI, with each erroneous touch lengthening the ITI by 0.1s.

Step 4a: Reinforcement of shapes: marmosets learned to touch the correct one of two circular colour patches presented to the left and right side of the screen. In the first round of training, the shapes were star, square and heart. In a daily session, 2 of the 3 were presented with locations randomized, and 1 of the 2 was rewarded. Once animals able to choose the correct target in 8 of 10 trials in 3 consecutive 10-trial blocks, on the following day, they were presented with a new pair of stimuli. The previous target now served as the foil, and the third shape was presented as the correct target. Across several sessions, animals learned that each of the 3 could be either the target or the foil. Initially, marmosets were allowed correction through no reward being delivered for the incorrect trial and a repeat of the same trial until the correct response was made . Once criterion was established, 1 or more additional training sessions were given in which no correction was allowed.

Step 4b: Reinforcement of colours: Similar to Step 4a, animals learned to choose 1 target from 2 stimuli and learned that each of 3 colours could be target or foil across several daily sessions. In the first round of training, the colours were red, yellow and blue. The order of Step 4a and 4b were counterbalanced across subjects. Steps 4a and 4b were only conducted for the first set of features; for the second set, animals were directly given the intra- and extra-dimensional shifting task.

Intra- and Extra- Dimensional Shifting

Following Steps 4a-4b, animals completed a within-session shifting task in which the rewarded rule was a single target feature (e.g., heart or yellow) that changes after criterion is

reached (Figure 1B, 1C). On each trial (Figure 1C), two compound stimuli, a coloured shape on the left and right) were presented. One stimulus contained the current target feature on the relevant dimension; the alternative (foil) contained a different feature from that same dimension. Animals indicated their choice by touching one stimulus; correct choices triggered reward delivery via the infusion pump, whereas incorrect choices resulted in no reward and the next trial began after ITI.

Criterion for each rule was defined as 80% correct over the most recent 10 trials, subject to a minimum of 20 trials completed under that rule before a switch could occur. When criterion was met and the 20-trial minimum satisfied, the target feature changed within the session. Intra-dimensional shifts (IDS) kept the same relevant dimension while changing the target feature (e.g., heart to star; Figure 1B). Extra dimensional shifts (EDS) changed the relevant dimension (e.g., heart to yellow; Figure 1B). The sequence of rules within a session was predetermined before testing to include both IDS and EDS.

For the original stimulus set (Figure 1A), shapes were star, square and heart; colours were red, yellow and blue. A second stimulus set (Figure 1D) comprised three novel shapes (pentagon-like, stapler-like, and pacman-like) and three novel colours (orange, cyan, purple); these were introduced directly as compound stimuli for shifting without additional pre-training. The novel shapes were generated in Python by creating polygons through manipulation of vertex coordinates. Starting from simple shapes (e.g., squares and triangles), vertices were repositioned to create unique polygons that were easily distinguishable from one another. The final shapes were selected to maximize perceptual distinctiveness while maintaining a level of concave-convex variation comparable to the first set of stimuli. All colours were protanopia-friendly, as male marmosets are red-green colorblind and distinct from the first set of colours. All other task

parameters (trial structure, criterion, and rule-switch logic) were identical for the second set. Session length and the number of rules completed varied across animals and days. All trials were recorded with synchronized overhead and side-view cameras.

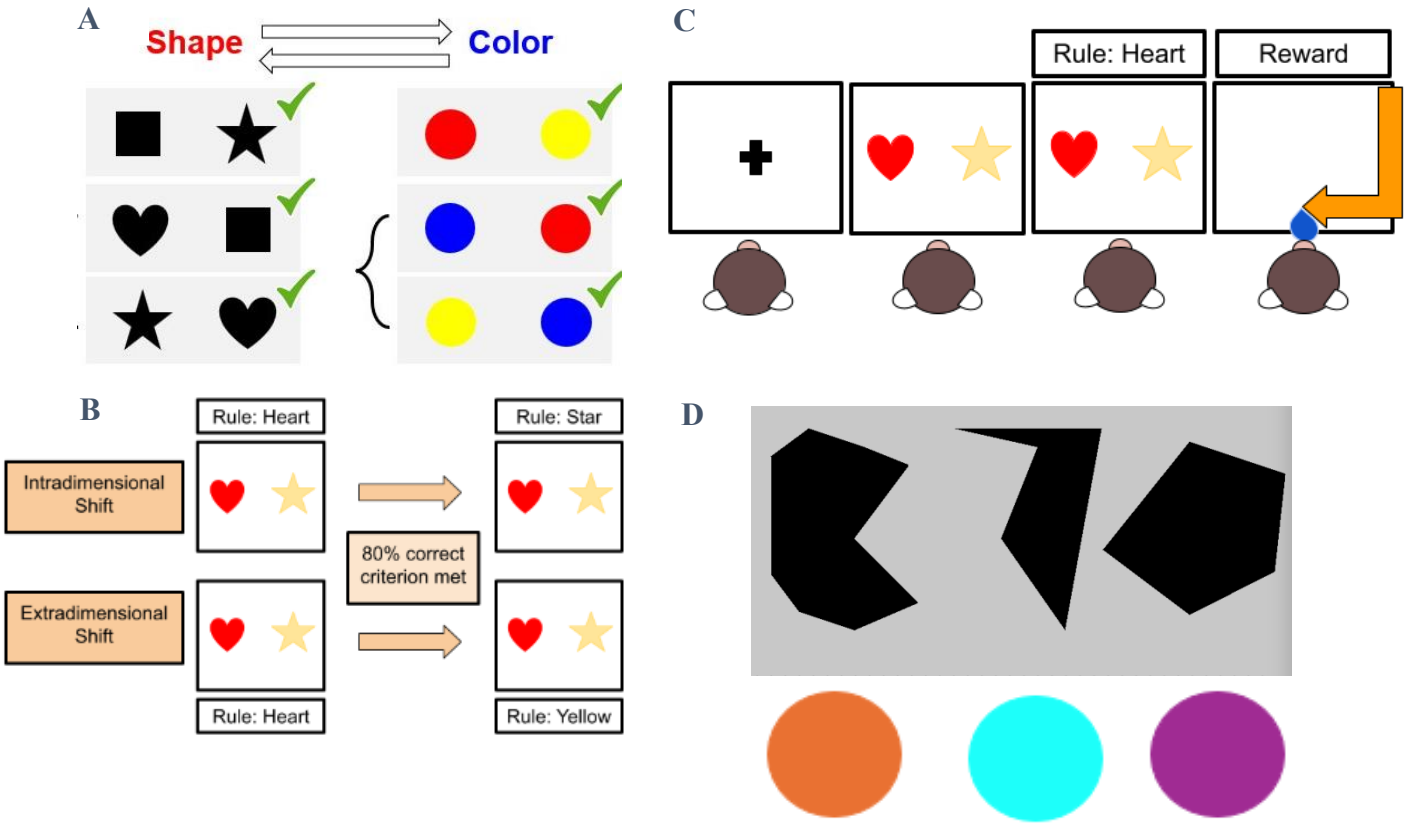

Figure 1: Task Design and Training. A: Initial shape and colour training for the original stimulus set. Each shape and colour acted as a target and a foil. B: Intra-dimensional shifts are when the target feature changes within dimension (e.g. heart to star) upon meeting the 80% correct response criterion. An extra-dimensional shift is when the target feature changes across dimensions (e.g. heart to yellow). C: Rule switching touchscreen task. D: New Stimuli set shapes and colours.

2.4.Results

2.4.1. Marmosets can successfully perform multiple switches within a single session

When presented with pairs of compound stimuli, all subjects learned to focus on one feature at a time and then switch to a new feature when the rule changed. For the first set of

features, this was achieved with <1 week of training once they became familiar with each shape and colour. Across sessions for the first stimulus set (73 sessions), the mean number of switches per session was 3.60 (median = 4, range = 0-8) (Figure 2A). For the second set of features, marmosets were able to successfully switch without initial shape and colour training. In the second set (90 sessions), marmosets completed an average of 3.61 switches per session (median = 4, range = 1-8) (Figure 2B).

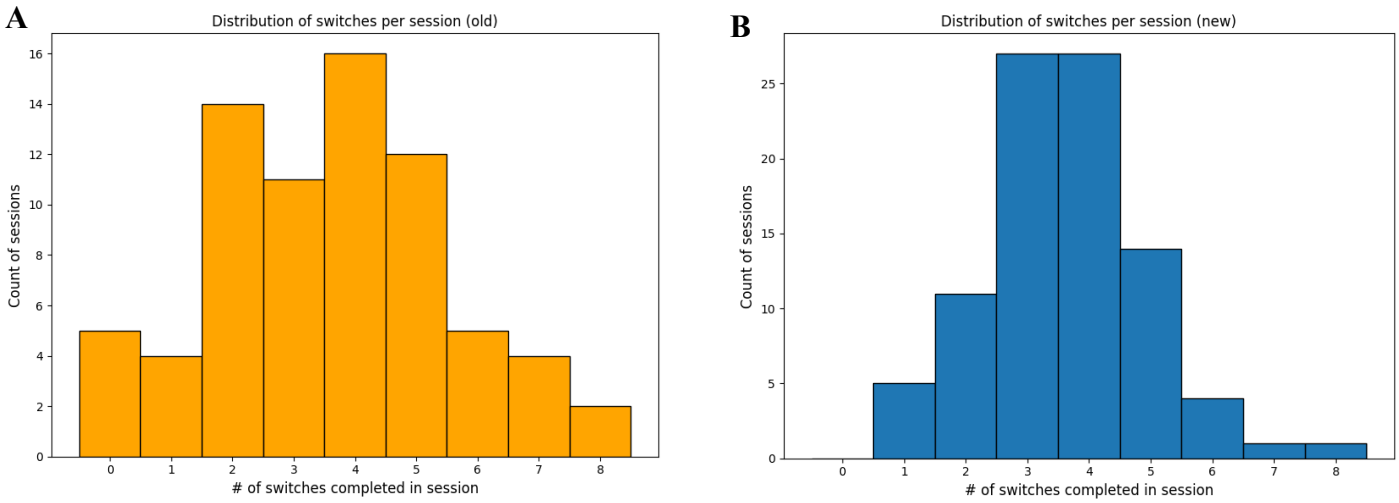

Figure 2. Distribution of the switches per session. (A) Sessions from first set of stimuli. (B) Session from the second set of stimuli. Histograms show the frequency of sessions as a function of the number of rule switches achieved within a session.

Notably, they were able to perform multiple switches both within the colour or shape dimensions and across dimensions like humans and macaques (Figure 3). In an example session (Figure 3), Marmoset M started with 'heart' as the target feature (orange shade). He had to discover the correct feature by trial and error and choose heart-shaped stimuli while ignoring its colour or location. By the end of approximately 50 trials, he made 8 correct choices (blue dots) in 10 consecutive trials, as highlighted by the red horizontal line. The target then switched cross-dimensionally to 'red' (light blue shade), and Marmoset M now had to learn to ignore the stimuli's shape instead of colour. It took approximately 20 trials to establish a consistent choice

of this new target feature. In this session, Marmoset M reached criterion following 6 switches, including colour-colour, shape-shape, and cross-dimensional switches .

Figure 3 displays an example session from each of the other 3 subjects.

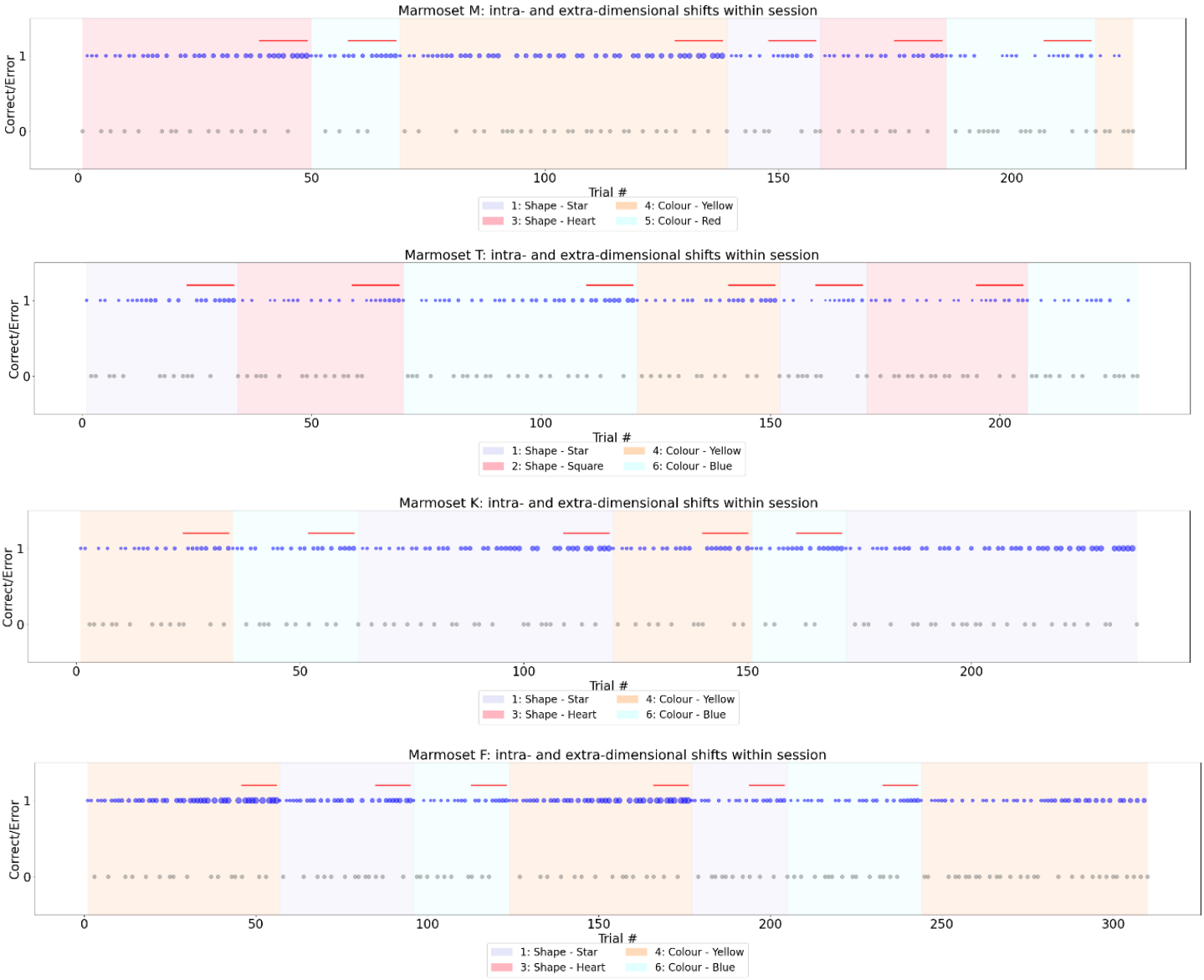

Figure 3: Intra- and extra-dimensional shifts within a single session. Each horizontal panel shows one animal (M, K, F, T from top to bottom) plotted by trial number and response outcome. Shaded background regions denote the actively rewarding rule. Circle size scales with the length

of consecutive correct responses, highlighting the build-up of learning. Red horizontal bars above the trials indicate the trials at which criterion was reached.

2.4.2. Performance accuracy changed significantly when a rule switches

Averaged across animals, mean accuracy dropped from 80 - 82 % in the 10 trials immediately before the switch to chance level on the 10 trials post switch for the original stimuli dataset, and from 80-82% pre-switch to 61-67% post-switch for the new stimuli dataset.

Target switching with the first set of features

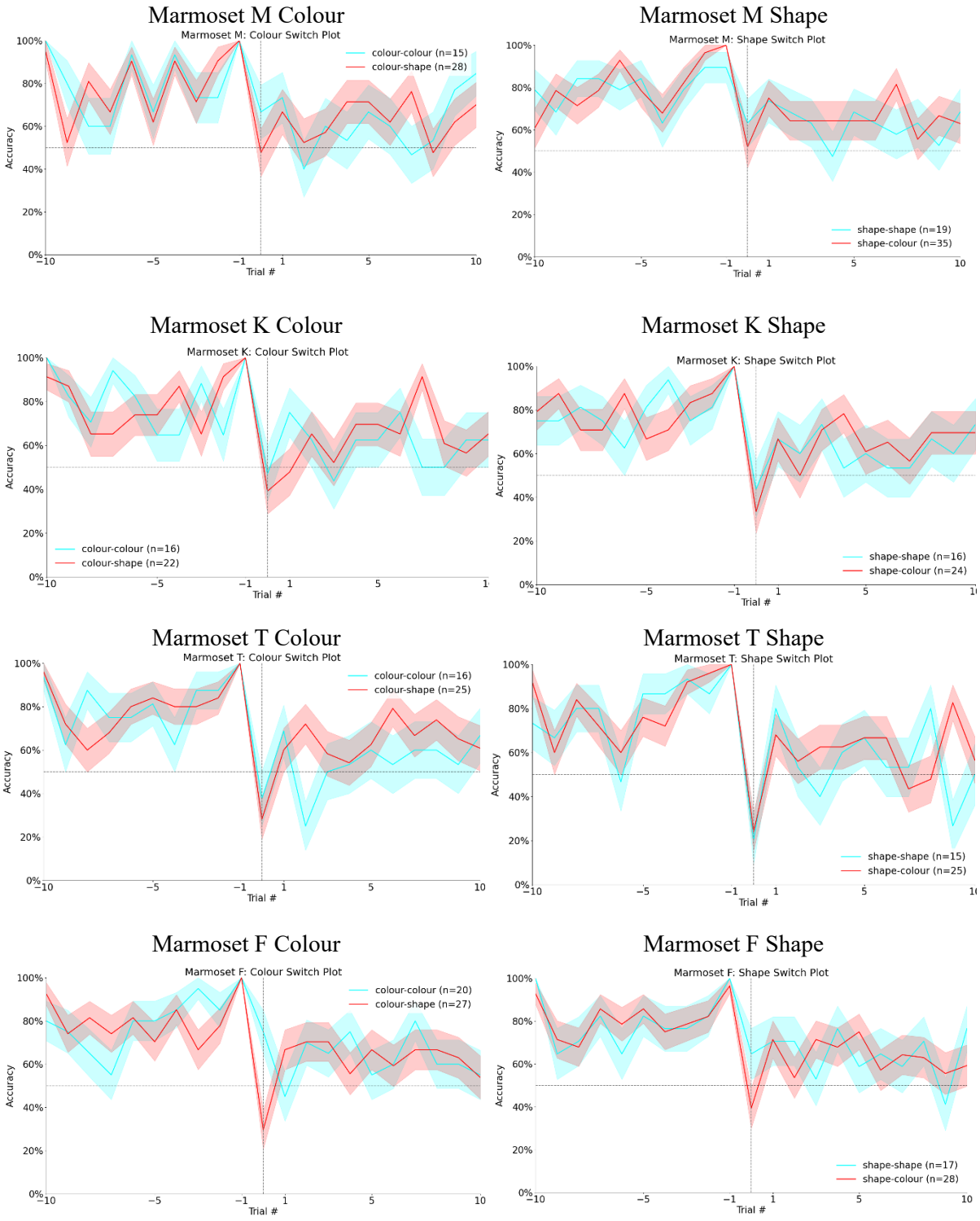

Figure 4: Original Stimulus Set Percentage accuracy across 10 trials before and after rule switches for intra-dimensional (cyan) and extra-dimensional (red) shifts. Accuracy dropped immediately following rule switches in both conditions.

Table 1. *Statistics for the change in performance by the type of switch, using the* ***first*** *set of features.*

	Intra: Shape-Shape			Intra: Colour-Colour			Extra: Shape-Colour			Extra: Colour-Shape		
Statistics	t-stat	df	p	t-stat	df	p	t-stat	df	p	t-stat	df	p
Marmoset T	6.18	13	<.001	6.50	14	<.001	6.91	22	<.001	5.72	22	<.001
Marmoset M	4.92	18	<.001	9.50	20	<.001	5.92	21	<.001	7.82	24	<.001
Marmoset F	4.04	16	<.001	4.65	19	<.001	7.78	26	<.001	8.38	25	<.001
Marmoset K	6.08	14	<.001	4.77	15	<.001	6.31	22	<.001	5.00	22	<.001

All four marmosets showed high levels of accuracy in the 10 trials preceding each rule switch, followed by a significant decline in performance immediately after the switch (Figure 5)

. This pattern was consistent across animals and switch types, confirming that marmosets were reliably following the current rule and responded to rule changes with a measurable drop in accuracy.

Post-switch accuracy did not significantly differ between intra-dimensional and extra-dimensional shifts for any individual subject, nor did one-way ANOVAs reveal group-level differences in post-switch accuracy by switch type (Table 1). These results suggest that, as expected, rule changes disrupted performance and animals required trial-and-error learning to identify the new target feature regardless of the dimension being switched.

Target switching with the second set of features

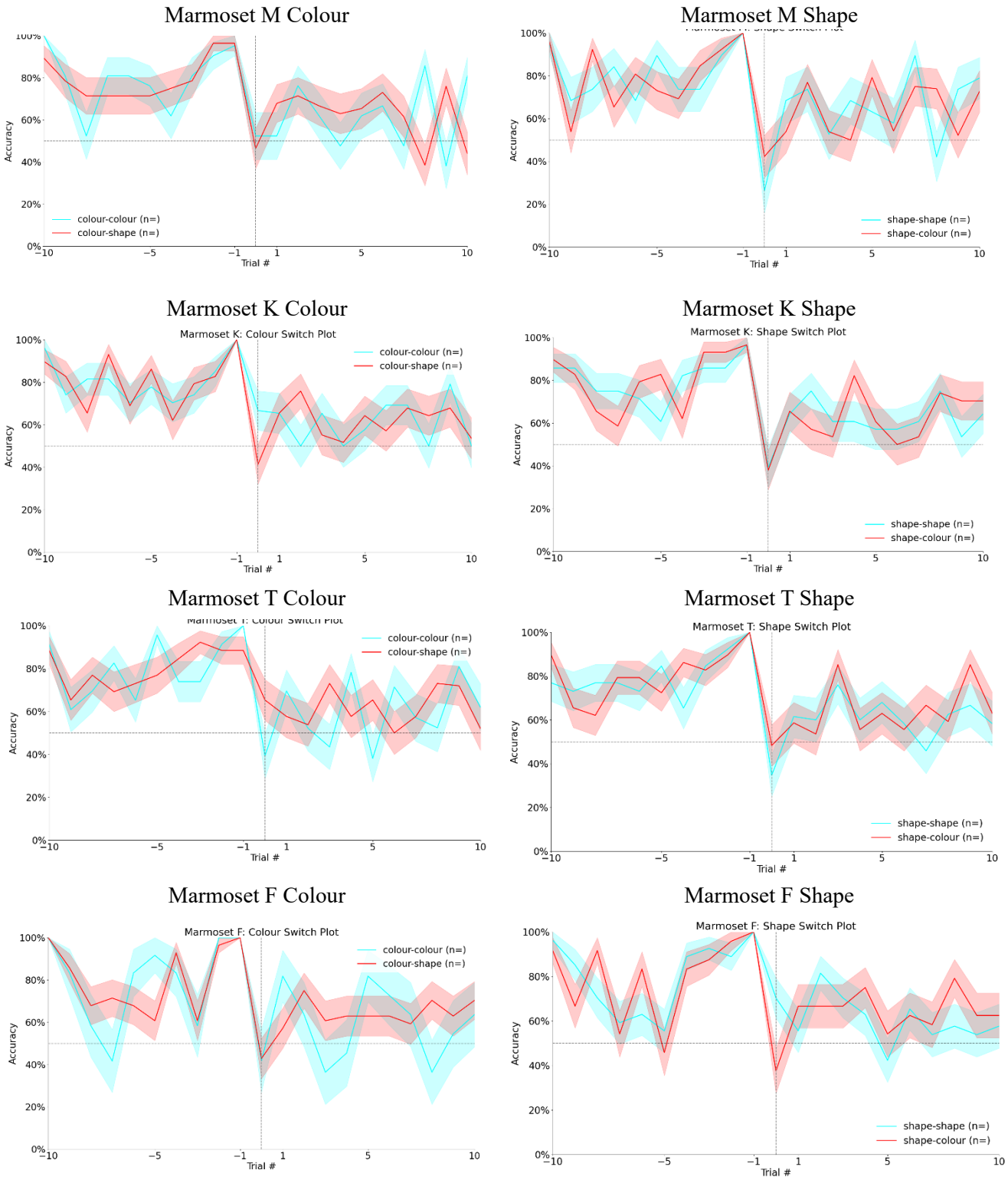

Figure 5: New Stimulus Set Percentage accuracy across 10 trials before and after rule switches for intra-dimensional (cyan) and extra-dimensional (red) shifts. Accuracy dropped immediately following rule switches in both conditions.

Table 2. Statistics for the change in performance by the type of switch, using the second set of features.

	Intra: Shape-Shape			Intra: Colour-Colour			Extra: Shape-Colour			Extra: Colour-Shape		
Statistics	t-stat	df	p	t-stat	df	p	t-stat	df	p	t-stat	df	p
Marmoset T	6.40	23	<.001	8.74	20	<.001	6.80	26	<.001	9.25	24	<.001
Marmoset M	4.92	18	<.001	9.50	20	<.001	5.92	21	<.001	7.82	24	<.001
Marmoset F	4.04	16	<.001	4.65	19	<.001	7.78	26	<.001	8.38	25	<.001
Marmoset K	8.9	27	<.001	8.88	23	<.001	8.52	26	<.001	7.72	27	<.001

As with the first stimulus set, all four marmosets showed high pre-switch accuracy across all switch types, followed by near-chance performance immediately after the rule changed, with a significant drop in accuracy in each case (Table 2, Figure 5). As before, this confirmed that animals consistently tracked the current target feature and needed to discover the new rule through trial and error. Importantly, this performance was observed using an entirely new set of compound stimuli with novel colours and shapes. No stimulus discrimination or reversal training was provided prior to testing. Instead, the animals were reintroduced to the intra- and extra-dimensional switching task using feature set 2, approximately five months after training on feature set 1. Remarkably, marmosets reached the same level of task engagement and pre-switch accuracy as they had with the original features, suggesting strong generalization of the task structure.

2.4.3. No significant differences in reaction time were observed across switch types

As illustrated in Figure 6, mean reaction time (RT) in the ten trials immediately before versus after a rule switch showed insignificant changes across switch types. One-way ANOVAs on post-switch RT revealed no significant differences among switch-type conditions for any animal (Table 6)

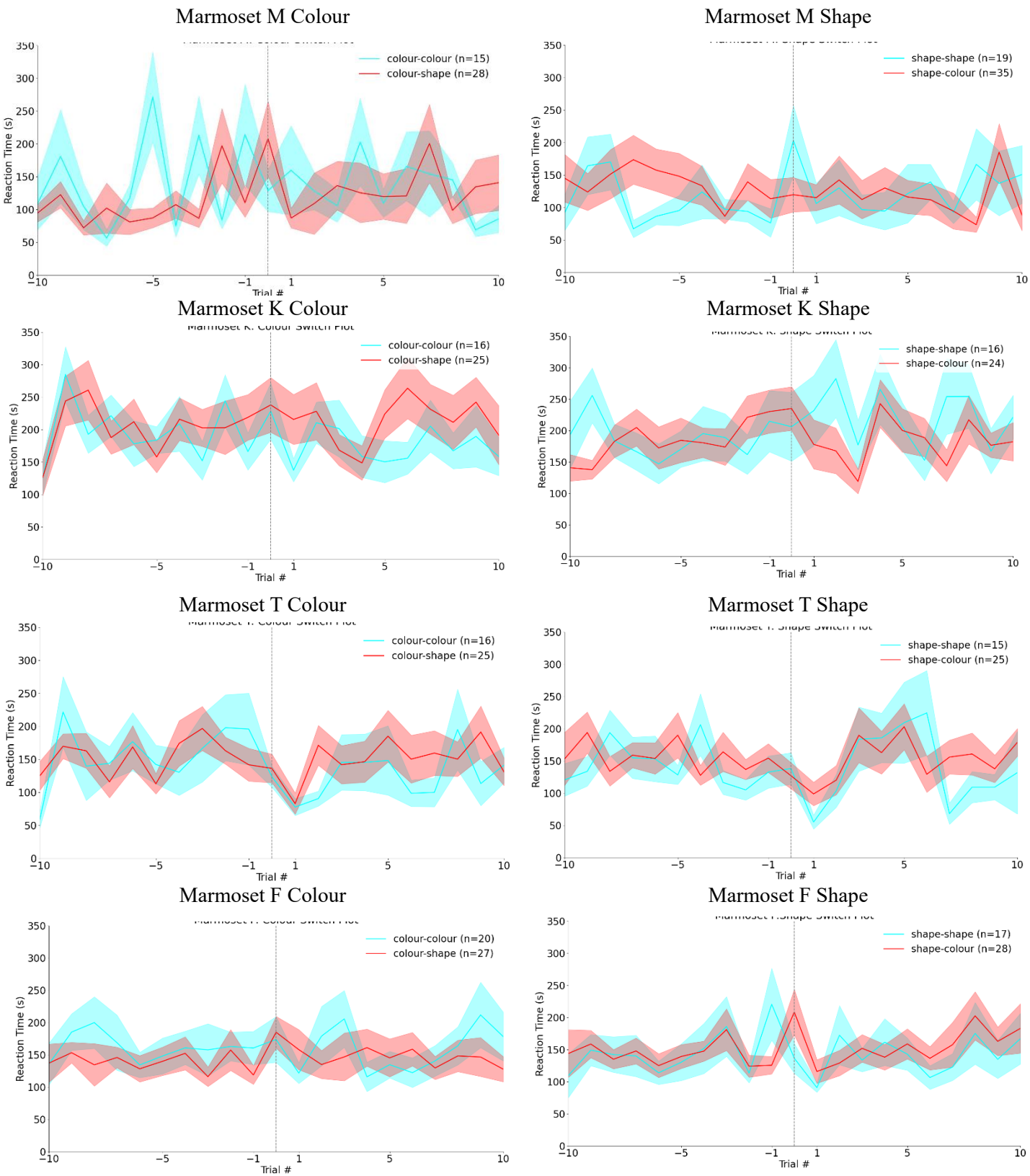

Figure 6: Original stimulus set reaction time, 10 trials before and after rule switches for intra-dimensional (cyan) and extra-dimensional (red)

Table 3. Statistics for the change in reaction time by the type of switch, using the first set of features.

	Intra: Shape-Shape			Intra: Colour-Colour			Extra: Shape-Colour			Extra: Colour-Shape		
Statistics	t-stat	df	p	t-stat	df	p	t-stat	df	p	t-stat	df	p
Marmoset T	0.34	13	.738	1.78	14	.096	-0.07	22	.946	-0.35	22	.732
Marmoset M	-1.4	18	.177	0.02	12	.981	1.57	26	.129	-0.92	18	.370
Marmoset F	-0.02	18	.986	-0.67	19	.512	-0.79	23	.432	-1.15	25	.257
Marmoset K	-1.99	14	.066	2.64	15	.018	-0.47	22	.645	-0.52	22	.611

As shown in Figure 8, mean reaction times before and after switches were largely unchanged across all four switch types. In individual comparisons (Table 3), only Marmoset K showed a statistically significant speeding after colour–colour switches ($t(15)=2.64$, $p=.018$), while Marmoset T trended toward faster responses in the same condition ($t(14)=1.78$, $p=.096$). No other switch type produced reliable RT changes in any animal. Group-level one-way ANOVAs on post-switch RTs confirmed no differences among switch types for any marmoset. These findings suggest that although accuracy reliably dipped immediately after a rule change, marmosets did not trade speed for accuracy. Their response times remained consistent, indicating they stayed engaged with the task rather than responding impulsively or hesitantly once the rule shifted.

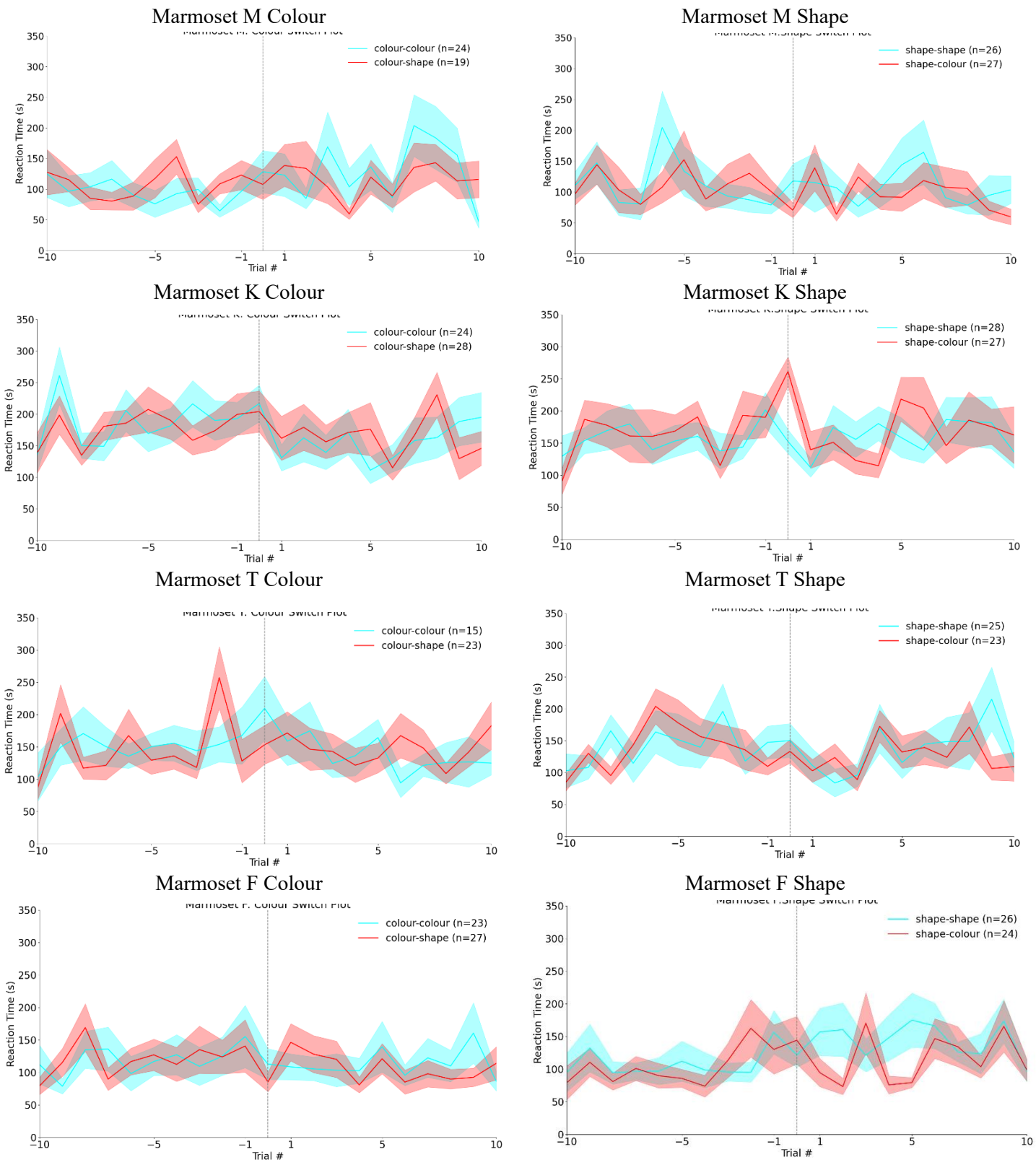

Figure 7: New stimulus set, reaction time, 10 trials before and after rule switches for intra-dimensional (cyan) and extra-dimensional (red) shifts.

Table 4. Statistics for the change in reaction time by the type of switch, using the second set of features.

	Intra: Shape-Shape			Intra: Colour-Colour			Extra: Shape-Colour			Extra: Colour-Shape		
Statistics	t-stat	df	p	t-stat	df	p	t-stat	df	p	t-stat	df	p
Marmoset T	0.06	24	.955	0.85	14	.405	0.60	22	.552	0	22	.996
Marmoset M	0.30	25	.769	-2.00	23	.059	1.15	26	.262	-0.53	18	.598
Marmoset F	-2.67	25	.013	0.39	22	.698	-0.79	23	.432	1.90	26	.069
Marmoset K	-0.26	27	.793	1.67	23	.108	-0.01	26	.989	1.13	27	.270

Reaction time patterns in the second stimulus set were generally stable across switch types, with no reliable effects of rule changes on speed of responding (Figure 7). For all four marmosets, pre- and post-switch RTs remained consistent across both intra-dimensional (colour to colour, shape to shape) and extra-dimensional (colour to shape, shape to colour) switches. As shown in 4, paired *t*-tests revealed no significant changes in RT following the switch in nearly all conditions. The only exception was Marmoset F, who showed a significant RT increase following shape to shape switches, $t\ (25) = –2.67$, $p = .013$. Marmoset M showed a marginal trend toward slower RTs after colour-to-colour switches ($p = .059$), but this did not reach significance. No other switch types produced significant changes in response speed. One-way ANOVAs on post-switch RTs confirmed that reaction time did not significantly differ by switch type in any animal. In contrast to the reliable drop in performance accuracy following rule changes, these findings indicate that RTs were unaffected by the type of rule switch.

2.4.4. The difficulty of establishing the new rule depended on the type of switch in some animals

Figure 8 (panels A–H) shows the number of trials each marmoset required to reach criterion after intra-dimensional and extra-dimensional shifts, separately for shape and colour rules across the original (white bars) and new (black bars) stimulus sets. For most animals, the pattern of results was consistent across datasets, with no reliable differences in trials to criterion between intra- and extra-dimensional shifts or between rule dimensions. However, as shown in Figure 9F, Marmoset K showed a notable exception, requiring significantly more trials to reach criterion following extra-dimensional shifts in the new dataset compared to intra-dimensional shifts, t (111) = 2.14, p = .034, with this effect also supported by an ANOVA main effect of Switch Type, F (1,109) = 4.57, p = .035. No such effect was observed in the original set, and there was no difference for shape rules (Figure 9E). This suggests that under certain conditions, the cognitive demand of shifting between dimensions may differentially impact performance depending on the stimulus set. No other marmoset showed a significant switch cost.

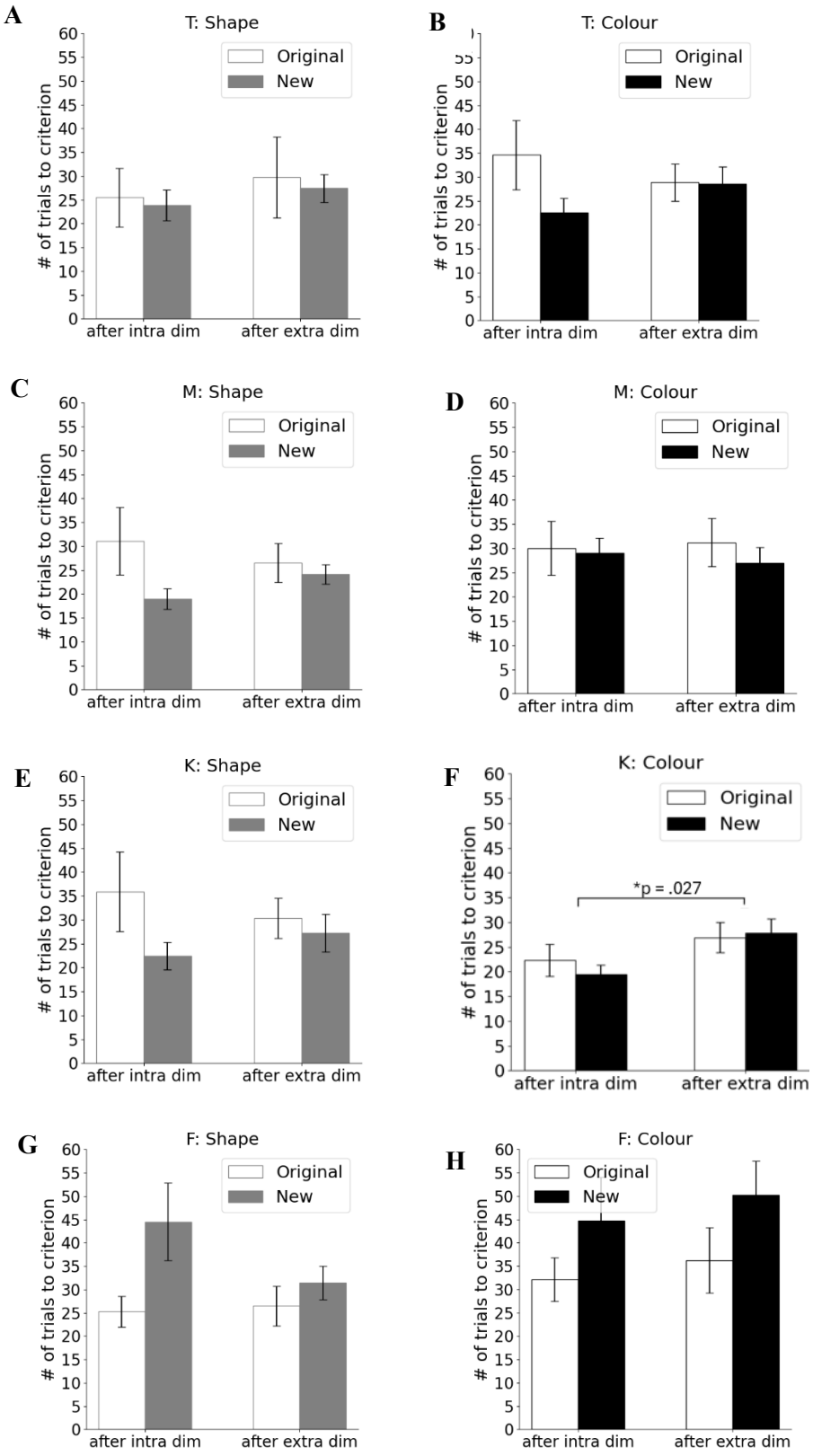
A
T: Shape
Original
New
# of trials to criterion
after intra dim
after extra dim
B
T: Colour
C
M: Shape
D
M: Colour
E
K: Shape
F
K: Colour
*p = .027
G
F: Shape
H
F: Colour

*Figure 8: Mean number of trials required to reach criterion for each marmoset (T, M, K, and F) following intra-dimensional and extra-dimensional shifts. Left panels show trials spent on shape rules, right panels show colour rules. White bars represent data from the original dataset and shaded bars represent data from the new dataset. Marmoset K required significantly more trials to reach criterion after extra-dimensional colour shifts in the new dataset (*p = .027).*

The behavioural findings presented in this chapter demonstrate that marmosets could perform our intra- and extra-dimensional shifting task. Additionally, all subjects successfully completed multiple switches within a session including both intra- and extra-dimensional shifts. A consistent drop in performance following each switch, followed by a recovery in accuracy confirms that the animals were attending to the relevant stimulus feature and updating their behaviour based on feedback.

Importantly, this behavioural flexibility generalized to a second, novel set of stimuli introduced after a five-month delay and without any pre-training or retraining. Marmosets performed comparably to their performance on the original stimuli set, suggesting that they had acquired an understanding of the task structure itself. While an extra-dimensional shift cost was not observed uniformly across animals, one subject (Figure 8F) showed significantly more trials to criterion following extra-dimensional shifts in the new stimulus set, providing preliminary evidence for a dimension-based set formation.

These results support the feasibility and validity of the current task as a tool for assessing attentional set shifting in marmosets. This task provides a strong behavioural foundation for future electrophysiological investigations into the neural mechanisms of cognitive flexibility.

Chapter 3

Aim 2: Implement a reinforcement learning model to simulate rule switching behaviour

3.1. Background

Reinforcement Learning (RL) is a computational framework for modeling how agents—biological or artificial—learn to make decisions by interacting with their often-dynamic environment. RL aims at optimizing the cumulative outcome in the long run and involves learning from trial and error through rewards and/or punishments (Sutton & Barto, 2018). Therefore, performance on set-shifting and other flexibility tasks are well-suited for RL modeling.

In addition to updating the values associated with stimuli or response, the RL model for set-shifting tasks must capture some unique parameters. First, set-shifting studies in humans and primate models typically involve compound visual stimuli, the value for each of 2-3 features within a stimulus must be updated separately, and more than one learning rate may be required (Collins & Frank, 2013). Second, set-shifting studies involve the switching of target both within and across dimensions, with the latter typically taking more trial and error than the former, the RL model will need to account for this effect of dimension.

To meet these unique requirements, Steinke et al. (2020) incorporated a dynamic attentional weight that modulates the degree to which recent feedback influences feature updating. In their model, positive prediction errors in one dimension increase an attentional weight for that dimension, leading to disproportionately larger updates when the same dimension

is relevant on subsequent trials (Steinke et al., 2020). Such models have improved fit to human set-shifting data by accounting for the reduction in perseverative errors and more rapid reversals (Steinke et al., 2020). More recently, Talwar et al. (2024) developed RL models that included three free parameters: 1). learning rate (α); 2). choice determinism (β), which quantifies the extent to which the choice is determined by the value of the stimulus or was random, and 3). 'dimension primacy' (θ) which measures the extent to which the initially relevant dimension is attended to.

An RL model fine-tuned for the set shifting tasks will be helpful for identifying the neural substrates. Mansouri et al. (2006) recorded single unit activity in the dorsolateral prefrontal cortex (dlPFC) of macaque monkeys performing a set-shifting task involving the shape and colour rules. They observed neurons with activities correlated with prediction error in a rule-specific manner, suggesting that primate prefrontal circuits implement some aspects of RL computations underlying rule switching behaviour. It remains unclear how other aspects of the RL-like computation is supported by dlPFC neurons, or neurons elsewhere in the extended control network. Collectively these findings support the use of feature-based RL models and their extensions to quantify learning parameters such as learning rate, choice determinism, dimension primacy. In applying these models to the marmoset WCST performance, we can test whether the common marmoset exhibits computational profiles like those of humans and macaques, thereby extending our understanding of cognitive flexibility across species.

3.2. Hypothesis

Aim 2 evaluates whether reinforcement learning models can capture marmoset rule switching behaviour and whether models that include attentional modulation provide a superior fit to the data. We specify two primary questions and their corresponding null and alternative hypotheses:

1. Model comparison of fit quality

   We predict that models incorporating attention weights or separate dimension specific learning rates will fit the behavioural data better than a basic feature reinforcement learning model.

   H_0: There is no difference in goodness of fit when using Bayesian Information Criterion (BIC) to compare among the feature reinforcement learning model, combined attention modulated FRL (CaFRL) and separate attention modulated FRL (SaFRL) models.

   H_1: The CaFRL or SaFRL models that account for dimensional effects, achieve significantly better fit than the basic FRL model.

2. Parameter correspondence with human data

   We predict that key parameter estimates will fall within healthy human ranges as found by Talwar et al. (2024).

   H_0: Marmoset parameter estimated are not different from the human parameter values reported by Talwar et al. (2024).

   H_1: Marmoset parameter estimates are significantly different from healthy human parameter estimate ranges.

To test these hypotheses, each model will be fit to individual marmoset choice sequences by maximum likelihood estimation. Model fits will be compared and marmoset parameter estimates will be compared to the human values and ranges reported in Talwar et al. (2024).

3.3. Methods

3.3.1. Model specifications

To capture trial-by-trial learning and attentional dynamics in our marmoset intra- and extra-dimensional set-shifting task, we implemented three reinforcement learning models originally described by Talwar et al. (2024). Each model computes stimulus values based on their features, uses a softmax function to produce the choice, and updates feature and attention weights via prediction errors. Compared to the IED task in CANTAB, our task involves continuous within-session rule switches embedded in a trial-based format rather than distinct phases. We adapted their algorithm to reflect this small difference without changing the underlying mathematics and applied maximum-likelihood estimation to obtain parameter estimates.

3.3.1.1 Feature Reinforcement Learning (fRL)

The simplest model (fRL) assumes that the value V(S) of a compound stimulus, S, is the sum of its feature weights:

$$V(S) = \sum_{f \in F(S)} W(f) \qquad \text{Equation 1}$$

Where F(S) is the set of two features (one colour, one shape). Choices among the two on-screen stimuli were governed by a softmax rule:

$$p(S_i) = \frac{e^{(\beta V(S_i))}}{e^{(\beta V(S_i))} + e^{(\beta V(S_j))}}, \qquad \text{Equation 2}$$

with β (choice determinism), the inverse temperature parameter controlling exploitative versus exploratory choices. On each trial, *t*, the weights of the chosen stimulus features are updated via

$$W(f)_{t-1} = W(f)_{t-1} + \alpha(R_t - V(S)_{t-1}) \qquad \text{Equation 3}$$

Where α is the learning rate and R_t is the feedback or outcome of their response (correct = 1, incorrect = 0). This model does not include dimensional considerations or separations.

In fRL, each stimulus features, i.e. a particular shape or colour, is associated with a value V(f). After choosing a stimulus composed of two features, f1 and f2, the prediction error is computed as

$$\delta_t = R_t - [V(f_1)_{t-1} + V(f_2)_{t-1}] \quad \text{Equation 4}$$

Where R_t is the reward received. The feature values are then updated according to

$$V(f_i)_t = V(f_i)_{t-1} + \alpha\delta_t,$$

Equation 5

3.3.1.2 Combined Attention-Modulated Feature Reinforcement Learning (Ca-fRL)

In Ca-fRL, stimulus value is computed as a weighted sum of its colour and shape features, where attention biases the contribution of each dimension. Only the features belonging to the currently relevant dimension are updated on a given trial, thereby capturing selective attention to the rewarded dimension. Updates are applied to features of the chosen stimulus and are scaled by the current attention weight of their dimension; features from the unchosen stimulus are not updated. Specifically, let $\alpha_t(\text{colour}) = \sigma(\theta_t)$ and $\alpha_t(\text{shape}) = 1 - \sigma(\theta_t)$; then

$$V(S) = \alpha_t(\text{colour}) \sum_{f \in F_{colour}(S)} W(f) + (\alpha_t \text{shape}) \sum_{f \in F_{shape}(S)} W(f) \quad \text{Equation 6}$$

The dimension primacy parameter, θ_0 is estimated and initial attention weights follow deterministically via the sigmoid; $\alpha_0(\text{colour}) = \sigma(\theta_t)$, $\alpha_0(\text{shape}) = 1 - \sigma(\theta_t)$. On each trial, both feature weights $W(f)$ and the attention weight are updated on each trial through backpropagation of the squared-error loss, using the same learning rate (Equation 3). Here, α

denotes the learning rate; the attention parameter, θ is updated with the same α in Ca-fRL, so attention weights can change via θ_t. This allows attention to dynamically shift toward the dimension that best predicts reward over time.

3.3.1.3 Separate Attention-Modulated Feature Reinforcement Learning (Sa-fRL)

Sa-fRL extends Ca-fRL by introducing a second learning rate, ε parameter, which is a separate learning rate that governs the updates to the dimension attentional weights, while α remains the learning rate for the feature weights. Talwar et al, (2024) describe Sa-fRL as identical to Ca-fRL except that, instead of using one α to update both feature and attention weights, dimension, θ, weights are updated with their own learning rate, ε:

$$\theta_t = \theta_{t-1} - \varepsilon \frac{\partial L}{\partial \theta}, \qquad \text{Equation 7}$$

Whereas feature weight still updates via

$$W(f)_t = W(f)_{t-1} - \alpha \frac{\partial L}{\partial W(f)} \qquad \text{Equation 8}$$

Therefore, ε indexes how quickly attention shifts between dimensions, independently of how fast the feature values are learned.

3.4. Results

3.4.1. Observed vs Simulated Errors

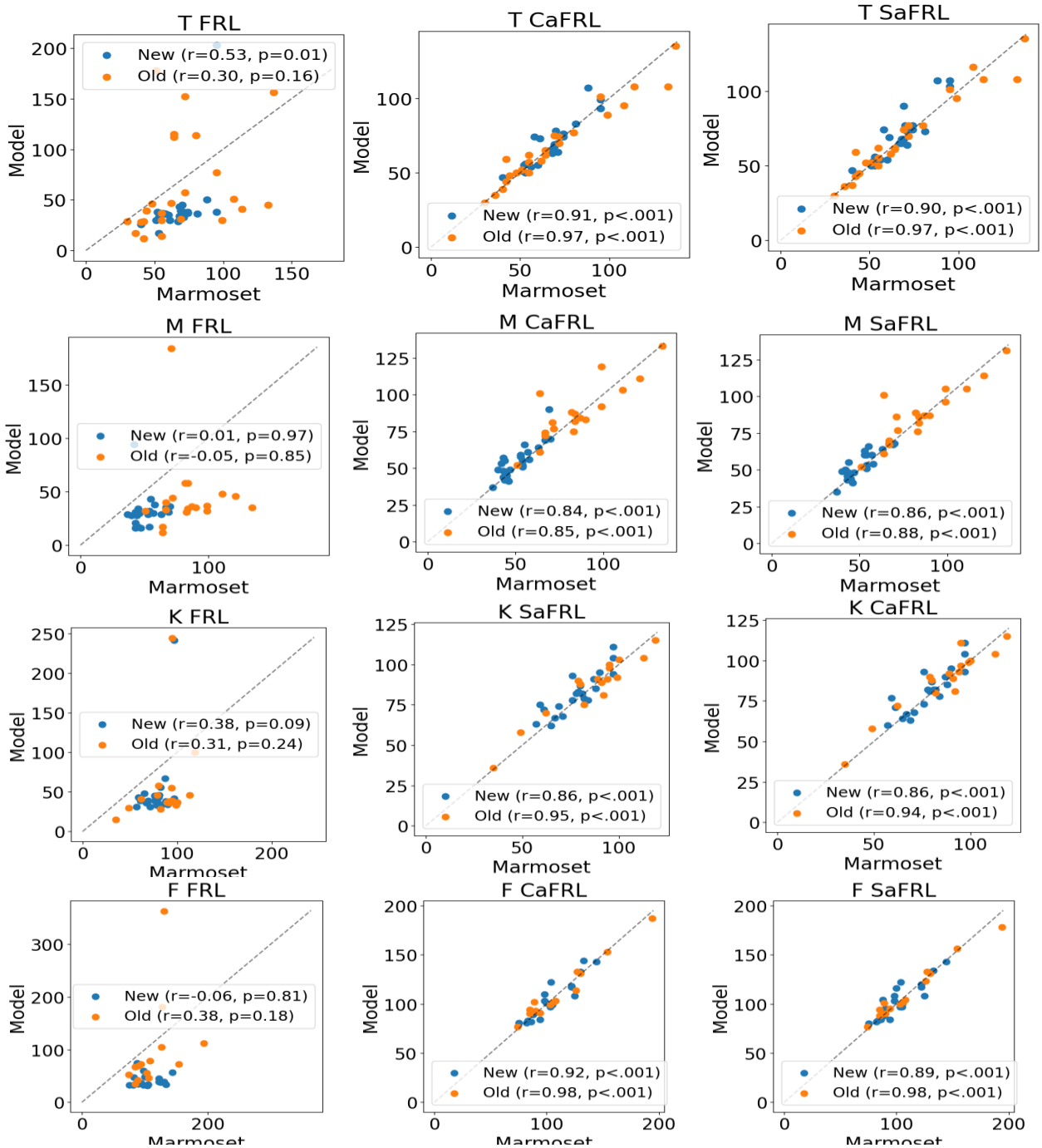

Figure 9: Trial by trial error correlations for FRL, CaFRL and SaFRL models across four marmosets. Each panel plots observed marmoset error counts against model predicted errors for the new stimuli set trials (blue circles) and the original stimuli set (orange circles). Rows correspond to individual animals (Marmoset T, M, K and F) and columns to the three models (FRL, CaFRL, SaFRL). Dashed diagonal lines indicate perfect prediction. Pearson's r and p-values for each condition are displayed within the legends of each panel.

Figure 9 presents for each marmoset and each model, scatter plots of marmoset versus model predicted error rates with original stimuli trials in blue and new stimuli trials in orange. Pearson correlation coefficients and significance levels are overlaid in each panel.

FRL model did not correlate with animal performance in three of four animals (Marmoset M, F and K, $p > 0.09$ for both first and second set of stimuli). For Marmoset T, FRL correlated with behavioural performance with the New set of stimuli ($r = .53$, $p = .02$) but not for the Original ($r = .30$, $p=.16$).This result indicates that the single learning rate FRL model did not capture trial-level error dynamics.

CaFRL produced uniformly high and significant correlations across all animals and both stimulus sets (all $r \geq .84$, $p< .001$). SaFRL exhibited similarly excellent performance (all $r \geq .86$, $p<.001$) with new stimuli set correlations ranging from .86 - .92 and original stimuli set correlations .88-.98 across animals. These results confirm that models that account for dimensional effects capture marmoset choice errors at the trial level far better than the simple FRL.

3.4.2. CaFRL provides the best overall fit

Across all four marmosets and both stimulus sets, CaFRL yielded the lowest BIC values (Table 5), indicating the best balance of fit quality and model complexity. SaFRL was consistently the second-best model and FRL the poorest. This uniform pattern (CaFRL performed the best in 8/8 cases) provides strong evidence that explicit dimensional parameter is critical for accurately capturing marmoset set-shifting behaviour.

Table 5: Bayesian Information Criterion (BIC) Values by Model and Stimulus Set

Model	**FRL**		**CaFRL**		**SaFRL**	
Stim Set	**Old**	**New**	**Old**	**New**	**Old**	**New**
Mouku	492.91	308.42	466.73	279.07	475.32	289.07
Tippy	382.53	396.78	353.8	366.92	363.98	367.16
Kipper	481.94	455.37	450.47	422.33	457.24	428.95
Fuego	640.82	599.7	583.1	552.51	599.52	560.72

3.4.2. Marmoset CaFRL Parameter values are not different from those of healthy humans

Tables 6-9 summarize the parameter estimates (mean ± SD) for each reinforcement learning model, animal and stimulus set. We report learning rates (α), choice determinism (β), dimension primacy (θ) and dimensional learning rates (ε).

To assess whether marmoset learning dynamics resemble those of humans, we compared our parameter estimates against the distributions reported by Talwar et al. (2024). In their study, healthy adults' FRL parameters averaged a learning rate of .62 (SD = .04) and a choice determinism of .91 (SD=.33). Under the combined-attention CaFRL model, human participants exhibited a learning rate of .92 (SD = .14), choice determinism of 1.34 (SD = 0.43), and a dimension primacy of 1.78 (SD = .97) . Finally, the separate attention-modulated, SaFRL model yielded human means for learning rate (features) of 0.91 (SD =0.14), learning rate (dimension) of 0.85 (SD = 0.24), choice determinism of 1.34 (SD= 0.42) and dimension primacy of 1.73 (SD = 0.87).

Table 6: Average Learning Rate by Model and Stimulus Set

Model	**FRL**		**CaFRL**		**SaFRL**	
Stim Set	**Old**	**New**	**Old**	**New**	**Old**	**New**
Marmoset T	0.67 ± 0.29	0.52 ± 0.25	0.85 ± 0.18	0.86 ± 0.13	0.86 ± 0.19	0.67 ± 0.29
Marmoset M	0.52 ± 0.19	0.56 ± 0.23	0.95 ± 0.09	0.87 ± 0.11	0.93 ± 0.11	0.52 ± 0.19
Marmoset K	0.59 ± 0.19	0.59 ± 0.20	0.83 ± 0.19	0.86 ± 0.13	0.91 ± 0.19	0.59 ± 0.19
Marmoset F	0.49 ± 0.30	0.42 ± 0.23	0.79 ± 0.20	0.90 ± 0.14	0.83 ± 0.29	0.49 ± 0.30
Human	0.62 ± 0.04		0.92 ± 0.14		0.91 ± 0.14	

Note: Human Parameter estimates from Table 1, Talwar et al. (2024).

Table 7: Average Choice Determinism by Model and Stimulus Set

Model	**FRL**		**CaFRL**		**SaFRL**	
Stim Set	**Old**	**New**	**Old**	**New**	**Old**	**New**
Marmoset T	0.26 ± 0.12	0.31 ± 0.08	0.27 ± 0.12	0.36 ± 0.09	0.29 ± 0.12	0.35 ± 0.07
Marmoset M	0.26 ± 0.06	0.33 ± 0.07	0.27 ± 0.05	0.36 ± 0.08	0.28 ± 0.05	0.36 ± 0.08
Marmoset K	0.25 ± 0.07	0.30 ± 0.05	0.28 ± 0.10	0.32 ± 0.05	0.29 ± 0.09	0.32 ± 0.05
Marmoset F	0.28 ± 0.06	0.27 ± 0.05	0.30 ± 0.07	0.29 ± 0.05	0.32 ± 0.08	0.31 ± 0.05
Human	0.91 ± 0.33		1.34 ± 0.43		1.34 ± 0.24	

Note: Human Parameter estimates from Table 1, Talwar et al. (2024).

Table 8: Average Dimension Primacy by Model and Stimulus Set

Model	**CaFRL**		**SaFRL**	
Stim Set	**Old**	**New**	**Old**	**New**
Marmoset T	-0.34 ± 11.1	2.22 ± 3.45	-0.74 ± 14.9	1.81 ± 3.99
Marmoset M	3.78± 5.41	1.87 ± 3.55	3.35 ± 5.37	8.58 ± 3.13
Marmoset K	-0.33 ± 6.68	2.79 ± 4.79	2.76 ± 5.85	2.71 ± 3.17
Marmoset F	-0.03 ± 8.67	0.21 ± 8.51	0.21 ± 8.51	4.89 ± 9.06
Human	1.78 ± 0.97		1.73 ± 0.87	

Note: Human Parameter estimates from Table 1, Talwar et al. (2024).

Table 9: Average Dimension Learning Rate by Model and Stimulus Set

Model	**SaFRL**	
Stim Set	**Old**	**New**
Marmoset T	0.93 ± 0.12	0.82 ± 0.34
Marmoset M	0.79 ± 0.35	0.88 ± 0.28
Marmoset K	0.77 ± 0.38	0.80 ± 0.36
Marmoset F	0.73 ± 0.36	0.71± 0.36
Human	0.85 ± 0.24	

Note: Human Parameter estimate from Table 1, Talwar et al. (2024).

In marmosets (Tables 6-9), FRL learning rates ($\alpha = 0.420 – 0.668$) overlapped the human range (0.58 – 0.66) for three of four subjects, though Marmoset F's values (.420 - .486) fell slightly below. FRL choice determinism ($\beta = 0.258 – 0.328$) was lower than the human range indicating generally more exploratory choices by marmosets. Under CaFRL, marmoset α (0.788-0.950) fell squarely within the human 0.78 – 1.06 interval and θ (-0.340 – 2.990) spanned the human mean ± SD (0.81-2.75). However, CaFRL β (0.274 – 0.362) remained below the human 0.91 – 1.77 range. Finally SaFRL feature learning rate (0.820-0.932), dimension learning rate (0.705-0.931), and θ (-0.735 – 8.581) all overlapped their respective human ranges, whereas SaFRLβ (0.288-0.357) again fell below the human 0.92-1.76 window. Thus, marmoset learning rates, and dimensional effect parameters generally mirror those of healthy adults, but marmosets exhibit systematically lower choice determinism values across all models.

Figure 10 provides a graphical comparison of CaFRL parameter distributions between humans and marmosets. Panel A depicts the human cluster centered means for α, β, and θ as reported by Talwar et al. (2024, Fig 3D). Panel B shows individual marmoset estimates (Original stimuli set in blue; New stimuli set in orange). Across all three parameters, marmoset points lie

within the human clusters (Figure 10), visually reinforcing the numerical overlap observed in Tables 6-9.

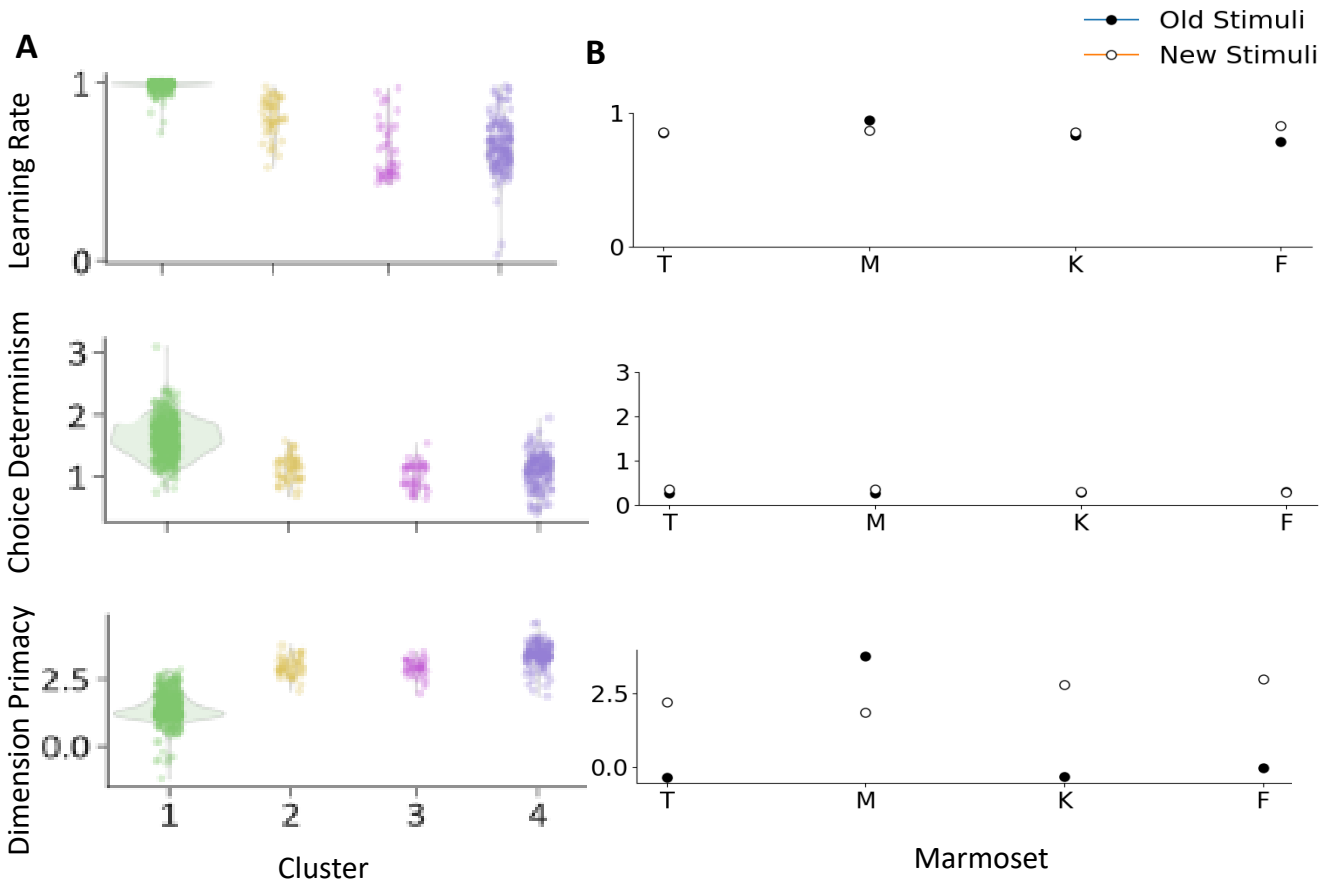

Figure 10: Comparison of CaFRL Parameters in Humans and Marmosets: Panel A presents human cluster means for learning rate, choice determinism and dimension primacy from Talwar et al. (2024,Fig3D). Panel B overlays individual marmoset parameter estimated for the original stimuli set (filled black circles) and new stimuli set (empty black outlined circles).

This chapter examined whether RL models could capture rule-switching behaviour in marmosets performing an IED shifting task. We implemented three models adapted from Talwar et al. (2024): a basic feature reinforcement learning model, a combined attention-modulated model, and a separate attention-modulated model. Consistent with our prediction, CaFRL provided the best overall fit to marmoset data across all subjects and stimulus sets, with SaFRL

closely following. These results suggest that including an explicit representation of dimensional attention is critical for accurately modeling marmoset performance in shifting tasks.

We also compared key parameter estimates from our marmosets to those previously reported in healthy human participants. Marmoset learning rates and dimension primacy values generally overlapped within human ranges across both CaFRL and SaFRL models. However, marmosets consistently exhibited lower choice determinism parameters, indicating more exploratory choice behaviour relative to humans. Notably, model-based error simulations for CaFRL and SaFRL correlated strongly with actual marmoset performance, whereas the simpler fRL model failed to account for trial-level error variability.

Together, these findings demonstrate that marmoset rule-switching behaviour can be meaningfully characterized using feature-based RL models with attentional modulation and that core learning mechanisms appear conserved across species. The consistency of dimensional attention effects and parameter comparability with human data further supporting the translational potential of the marmoset model for investigating the neural basis of cognitive flexibility.

Chapter 4

Discussion

4.1. Main Findings

In our touchscreen-based rule-switching task, marmosets quickly learned and reliably performed multiple rule switches in a single session. These switches were marked by clear drops in accuracy immediately post switch for both intra-dimensional shifts and extra-dimensional shifts. There were no significant intra- vs extra-dimensional shift differences in accuracy or reaction time using the original set of features. With the new set of features, one of four animals showed significantly greater trials to criterion for a novel colour rule following an extra-dimensional shift, while two others showed a similar trend. Thus, effects of dimension may become better established through exposure to more sets of features.

Did the marmosets show any sign of forming attentional sets? Solely based on the comparison of trials to criterion, they did not and instead learned the features independently without grouping them into dimension. However, of the three reinforcement learning models tested the combines-attention, CaFRL—a model that takes into account the currently relevant dimension—achieved the lowest BIC in all eight comparisons (4 animals x 2 stimulus sets). Moreover, CaFRL accounted for trial-by-trial error patterns with high correlations across subjects ($r \geq .84$, $p < .001$), whereas FRL—the model that did not consider the effect of dimension—predictions were weak and non-significant. These findings clearly suggest a significant effect of dimensions.

In summary, our observation that marmosets can execute multiple rule shifts within a single session successfully confirms that, like macaque monkeys (Goudar et al 2024), they are

capable of rapid switching between task rules. The length of training required was relatively short compared to similar macaque studies. Additionally, key parameters (learning rate, choice determinism and dimension primacy) fell within ranges reported in healthy humans (Talwar et al., 2024), though marmosets consistently show lower β values indicating more exploratory behaviour. These together suggest that the marmosets can serve as a reliable translational non-human primate model for human cognitive flexibility.

4.2. Task structure

While the present task is structurally simpler than multi-stage set shifting paradigms such as the CANTAB intra-/extra-dimensional (ID/ED) task developed by Roberts and colleagues, it is not equivalent to a simple reversal learning paradigm (Dias et al., 1996). In traditional reversal tasks, animals typically learn to associate specific, repeated stimuli with reward and must reverse these associations when contingencies shift. In contrast, our task required marmosets to track a relevant feature (e.g., red or star) in compound stimuli. On each trial, different colour-shape combinations were presented (e.g., red heart vs. yellow star, followed by yellow heart vs. red star), meaning that animals could not rely on object-level memory or stimulus repetition. Instead, successful performance required them to abstract a feature-based rule (e.g., choose yellow) and ignore the irrelevant dimension. These task demands are consistent with attentional set-shifting rather than simple reversal. Thus, albeit not presenting novel exemplars after each switch, our task retains essential aspects of attentional shifting, including rule abstraction, flexible updating and dimensional control, while offering a more compact design that allows multiple rule switches within a singular session. The consistent use of the same set of features within session will also be crucial for neurophysiological analyses. Similar task design with a constant feature set have been used in macaque rule-switching studies, which required the animals to shift

attention among features within or across dimensions within session (Goudar et al., 2024; Mansouri et al., 2006).

In short, because the target stimuli change across trials and the animal must generalize the rule across multiple exemplars, our task encourages the formation of attentional set using a fixed set of features across sessions. These qualities allow us to track rule switching within a single session, so neuronal correlates could be identified for features and dimensions, for expected and surprising outcome, and for rule exploration and exploitation.

Our task design also promoted the voluntary engagement of marmosets and minimized training duration. Sessions were run in the familiar home-cage environment immediately before the animals' first scheduled meal, eliminating the need for food restriction and reducing stress from transport or restraint of any kind. For each ~40-minute session, we attached a portable touchscreen transfer box to the cage and allowed on marmoset to freely interact with the task. Additionally, each cage mate could observe the task through the closed off, yet not visually blocked area, fostering social facilitation and observational learning as they saw that correct responses produced palatable rewards. Across sessions, individual marmosets typically completed three to nine rule switches within a single session, demonstrating both a sustained level of engagement and the feasibility of obtaining multiple switches in a single session. Importantly, this approach eliminated the need for chair training, preventing additional stressors associated with restraint and separation from cage mates, and preserved social dynamics that support welfare and reliable performance.

4.3.Extra-Dimensional Shifting

In the original stimulus set, marmosets did not show a performance cost following extra-dimension (ED) shift, relative to intra-dimensional (ID) shifts. One interpretation is that this phase reflected early acquisition of the task structure, during which animals had not yet formed strong attentional sets. In contrast, in the new stimulus set, one marmoset showed a clear ED cost, requiring significantly more trials to reach criterion following an ED switch. Notably, this occurred even though the stimuli themselves (new shapes and colours) were novel. The marmosets had only been pre-trained on the original stimuli set features and did not receive initial shape and colour training for the second set of stimuli. Yet, marmosets were able to engage with the full task structure and switch rules within a few days. This suggests that they had generalized the underlying demands of the task, including the need to monitor for rule switches and track relevant dimensions, beyond the specific stimuli.

This transfer across feature sets reflects the possible emergence of a consistent shifting strategy. However, another plausible explanation is that the limited number of exemplars used in each feature dimension during the original set may have encouraged feature-specific responding rather than true dimensional abstraction. As Roberts and colleagues have argued, the use of more exemplars helps promote generalization and the formation of attentional set (Roberts et al., 2004). The emergence of an ED cost in the new stimulus set, where novel exemplars were used, provides empirical support for this idea, suggesting that this dimensional abstraction may have emerged once more exemplars were shown. The delayed appearance of an ED cost from the original to new stimulus set may mark a transition from task learning to strategic performance, as has been observed in human and non-human primate set shifting paradigms. Previous studies in both rodents and primates have shown that extra-dimensional costs typically emerge only after

subjects experienced a consistent series of intra-dimensional shifts, allowing a bias towards a specific dimension to form (Birrell & Brown, 2000; Dias et al., 1996; Tait et al., 2018).

This interpretation is further supported by the model derived estimates of dimension primacy, which quantify the extent to which attention was biased towards one stimulus dimension. In the first stimulus set, three of four animals showed near-zero-dimension primacy values, consistent with the absence of an ED cost and suggesting that an attentional set had not formed, yet trial by trial attentional modulation was still present as captured by models accounting for dimensional effects (CaFRL and SaFRL). In contrast, in the new stimulus set, all four animals exhibited a strong dimension bias, mirroring the behavioural shift cost observed in at least one animal. This pattern supports prior findings that ED costs emerge only after a consistent attentional set has developed through the task (Dias et al., 1996). Together, the modelling and behavioural findings suggest that animals transitioned from a learning phase to a strategy-based phase in which attentional control guided performance. Interestingly, while all four marmosets showed an increased dimension primacy estimates in the second stimulus set, suggesting the development of an attentional bias, only one animal demonstrated a statistically significant ED cost at the behavioural level. It is possible that the remaining animals also experienced increased difficulty disengaging from the previously relevant dimension, but this did not manifest in the number of trials to criterion. It may require more training with more feature sets to establish an attentional set.

4.4. Reinforcement Learning Mechanisms and Dimensional Effects

Aim 2 tested whether models that explicitly represent dimensional parameters capture marmoset behaviour better. Both the CaFRL and SaFRL, which allow for differential updating for colour vs. shape, yielded lower BIC scores and stronger error behaviour correlations than

simplest fRL model. The inclusion of θ in CaFRL yielded higher learning rates (α = 0.79 – 0.95) and θ values between-0.34 and 2.99 indicated attentional biases at onset. SaFRL further refined these effects by fitting separate learning rates (ε = 0.705-0.931), capturing subtle shifts in attentional allocation. This aligns with human findings, suggesting that marmosets like humans allocate attention unevenly across feature dimensions when learning new rules (Collins & Frank, 2013; Talwar et al., 2024). The fitted parameter ranges fell within human ranges reported by Talwar et al. (2024), further supporting cross-species similarities in computational underpinnings of set shifting.

4.5. Neural mechanisms

Although this study lacks direct neural measurements, our computational approach can inform hypotheses about underlying neural mechanisms. Here I attempt to interpret model parameters as potential proxies for cognitive processes supported by specific neural systems.

Learning rate parameters may reflect processes linked to cholinergic modulation in the prefrontal cortex and its associated network (Hasselmo & Sarter, 2011), while prediction errors conceptually align with dopaminergic signaling in corticostriatal circuits (Schultz et al., 1997). Additionally, posterior parietal cortex activity spikes at shift onset, consistent with the attentional reorientation captured by θ (Kim et al., 2012). Lower β values may be supported by ACC, in which the detection of a conflict can help reduce decision thresholds and promote exploration upon rule change (Hampshire & Owen, 2006). Finally, the mediodorsal thalamus likely facilitates these cognitive processes across frontal regions during rule updating (Ouhaz et al., 2022).

These neural hypotheses are speculative but testable. Future studies could combine electrophysiological and causal approaches to further probe the network mechanisms underlying set shifting. For instance, since θ has been associated with PPC-mediated attentional control, then single-unit recordings in PPC could reveal how theta dynamics relate to behavioural modulation during ED shifts (Yang et al., 2017). Likewise, recordings in ACC could examine how beta-related dynamics are linked to performance monitoring and updating (Floresco, 2013; Mansouri et al., 2009). Recordings in dlPFC could identify rule-specific prediction error signals that support flexibility (Buckley et al., 2009; Mansouri et al., 2006). Future work should also incorporate causal manipulations such as reversible inactivation to directly test whether these regions are necessary for flexible behaviour.

4.6. Limitations & Future directions

A primary limitation of this study is that all four marmosets were male, which may restrict the generalizability of our results. Future research should include female marmosets to assess potential sex differences in attentional parameters, learning rate and choice determinism. (Burman et al., 2011). However, meaningful assessment of sex differences will require larger sample sizes or pooled data across studies, as single studies with small cohorts unlikely to provide sufficient statistical power. Sex differences in set shifting for non-human primates are scarce, further validating the importance for future work to include female marmosets to determine whether parameters such as learning rate or dimension primacy differ by sex.

Beyond sampling, direct neurophysiology remains an important next step. Combining our paradigm with electrophysiological recording and reversible inactivations in dlPFC, ACC, PPC and key subcortical structures would test the circuit mechanisms inferred by our CaFRL model. For instance, recordings in the dlPFC during rule switches could reveal neural correlates of the

dimension primacy parameter, indicating how attentional priorities shift (Kamigaki et al., 2009; Mansouri et al., 2009). Similarly, ACC recording could clarify how conflict-related activities influences choice determinism and exploration (Botvinick et al., 2001; Mansouri et al., 2017), while PPC and mediodorsal thalamus recordings could capture network-level dynamics underlying strategy transitions (Bissonette & Roesch, 2015; Schmitt et al., 2017). Complementary reversible inactivation studies across these regions could then test whether they are necessary for flexible behaviour. These approaches could provide a rigorous test of this model's neural predictions and advance the understanding of cognitive flexibility.

Marmosets offer unique advantages for studying cognitive flexibility. Their smooth cerebral cortex facilitates neural access while retaining key primate features of prefrontal organization. Their short life span is amenable to longitudinal studies, which found that their age related declines in working memory resembled to those observed in humans (Glavis-Bloom et al., 2022). These findings underscore the potential of marmosets to serve as an effective model for investigating cognitive aging processes relevant to humans. Marmosets also serve as robust models for anxiety disorders. In approach-avoidance paradigms, lesions to the ventrolateral prefrontal cortex produces heightened anxiety-like behaviours when marmosets face predator stimuli (Shiba et al., 2015). Collectively, these studies validate marmosets as translational models that bridge cognitive domains, making them ideal for integrated behaviour, modelling and neural investigations.

References

Bayer, H. M., & Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron*, *47*(1), Article 1. https://doi.org/10.1016/j.neuron.2005.05.020

Birrell, J. M., & Brown, V. J. (2000). Medial Frontal Cortex Mediates Perceptual Attentional Set Shifting in the Rat. *Journal of Neuroscience*, *20*(11), Article 11. https://doi.org/10.1523/JNEUROSCI.20-11-04320.2000

Bissonette, G. B., & Roesch, M. R. (2015). Rule encoding in dorsal striatum impacts action selection. *European Journal of Neuroscience*, *42*(8), Article 8. https://doi.org/10.1111/ejn.13042

Block, A. E., Dhanji, H., Thompson-Tardif, S. F., & Floresco, S. B. (2007). Thalamic-prefrontal cortical-ventral striatal circuitry mediates dissociable components of strategy set shifting. *Cerebral Cortex (New York, N.Y.: 1991)*, *17*(7), Article 7. https://doi.org/10.1093/cercor/bhl073

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), Article 3.

Brady, A. M., & Floresco, S. B. (2015). Operant Procedures for Assessing Behavioral Flexibility in Rats. *Journal of Visualized Experiments : JoVE*, *96*, 52387. https://doi.org/10.3791/52387

Brown, M. R. G., Vilis, T., & Everling, S. (2007). Frontoparietal activation with preparation for antisaccades. *Journal of Neurophysiology*, *98*(3), Article 3. https://doi.org/10.1152/jn.00460.2007

Bucci, D. J. (2009). Posterior parietal cortex: An interface between attention and learning? *Neurobiology of Learning and Memory*, *91*(2), Article 2. https://doi.org/10.1016/j.nlm.2008.07.004

Buckley, M. J., Mansouri, F. A., Hoda, H., Mahboubi, M., Browning, P. G. F., Kwok, S. C., Phillips, A., & Tanaka, K. (2009). Dissociable Components of Rule-Guided Behavior Depend on Distinct Medial and Prefrontal Regions. *Science*, *325*(5936), Article 5936. https://doi.org/10.1126/science.1172377

Chudasama, Y., Dalley, J. W., Nathwani, F., Bouger, P., & Robbins, T. W. (2004). Cholinergic Modulation of Visual Attention and Working Memory: Dissociable Effects of Basal Forebrain 192-IgG-saporin Lesions and Intraprefrontal Infusions of Scopolamine. *Learning & Memory*, *11*(1), 78–86. https://doi.org/10.1101/lm.70904

Cole, M. W., Reynolds, J. R., Power, J. D., Repovs, G., Anticevic, A., & Braver, T. S. (2013). Multi-task connectivity reveals flexible hubs for adaptive task control. *Nature Neuroscience*, *16*(9), 1348–1355. https://doi.org/10.1038/nn.3470

Collins, A. G. E., & Frank, M. J. (2013). Cognitive control over learning: Creating, clustering, and generalizing task-set structure. *Psychological Review*, *120*(1), 190–229. https://doi.org/10.1037/a0030852

Dajani, D. R., & Uddin, L. Q. (2015). Demystifying cognitive flexibility: Implications for clinical and developmental neuroscience. *Trends in Neurosciences*, *38*(9), Article 9. https://doi.org/10.1016/j.tins.2015.07.003

Dalley, J. W., Cardinal, R. N., & Robbins, T. W. (2004). Prefrontal executive and cognitive functions in rodents: Neural and neurochemical substrates. *Neuroscience & Biobehavioral Reviews*, *28*(7), Article 7. https://doi.org/10.1016/j.neubiorev.2004.09.006

Dias, R., Robbins, T. W., & Roberts, A. C. (1996). Primate analogue of the Wisconsin Card Sorting Test: Effects of excitotoxic lesions of the prefrontal cortex in the marmoset. *Behavioral Neuroscience*, *110*(5), Article 5. https://doi.org/10.1037//0735-7044.110.5.872

Dias, R., Robbins, T. W., & Roberts, A. C. (1997). Dissociable forms of inhibitory control within prefrontal cortex with an analog of the Wisconsin Card Sort Test: Restriction to novel situations and independence from "on-line" processing. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *17*(23), Article 23.

Dosenbach, N. U. F., Visscher, K. M., Palmer, E. D., Miezin, F. M., Wenger, K. K., Kang, H. C., Burgund, E. D., Grimes, A. L., Schlaggar, B. L., & Petersen, S. E. (2006a). A Core System for the Implementation of Task Sets. *Neuron*, *50*(5), Article 5. https://doi.org/10.1016/j.neuron.2006.04.031

Dosenbach, N. U. F., Visscher, K. M., Palmer, E. D., Miezin, F. M., Wenger, K. K., Kang, H. C., Burgund, E. D., Grimes, A. L., Schlaggar, B. L., & Petersen, S. E. (2006b). A Core System for the Implementation of Task Sets. *Neuron*, *50*(5), Article 5. https://doi.org/10.1016/j.neuron.2006.04.031

Farashahi, S., Donahue, C. H., Khorsand, P., Seo, H., Lee, D., & Soltani, A. (2017). Metaplasticity as a Neural Substrate for Adaptive Learning and Choice under Uncertainty. *Neuron*, *94*(2), Article 2. https://doi.org/10.1016/j.neuron.2017.03.044

Floresco, S. B. (2013). Prefrontal dopamine and behavioral flexibility: Shifting from an "inverted-U" toward a family of functions. *Frontiers in Neuroscience*, *7*, 62. https://doi.org/10.3389/fnins.2013.00062

Floresco, S. B., Block, A. E., & Tse, M. T. L. (2008). Inactivation of the medial prefrontal cortex of the rat impairs strategy set-shifting, but not reversal learning, using a novel, automated procedure. *Behavioural Brain Research*, *190*(1), Article 1. https://doi.org/10.1016/j.bbr.2008.02.008

Floresco, S. B., Ghods-Sharifi, S., Vexelman, C., & Magyar, O. (2006). Dissociable Roles for the Nucleus Accumbens Core and Shell in Regulating Set Shifting. *The Journal of Neuroscience*, *26*(9), Article 9. https://doi.org/10.1523/JNEUROSCI.4431-05.2006

Ghods-Sharifi, S., Haluk, D. M., & Floresco, S. B. (2008). Differential effects of inactivation of the orbitofrontal cortex on strategy set-shifting and reversal learning. *Neurobiology of Learning and Memory*, *89*(4), Article 4. https://doi.org/10.1016/j.nlm.2007.10.007

Glavis-Bloom, C., Vanderlip, C. R., & Reynolds, J. H. (2022). Age-related learning and working memory impairment in the common marmoset. *The Journal of Neuroscience*, JN-RM-0985-22. https://doi.org/10.1523/JNEUROSCI.0985-22.2022

Goudar, V., Kim, J.-W., Liu, Y., Dede, A. J. O., Jutras, M. J., Skelin, I., Ruvalcaba, M., Chang, W., Fairhall, A. L., Lin, J. J., Knight, R. T., Buffalo, E. A., & Wang, X.-J. (2023). *Comparing rapid rule-learning strategies in humans and monkeys* (p. 2023.01.10.523416). bioRxiv. https://doi.org/10.1101/2023.01.10.523416

Grant, K. A., Newman, N., Lynn, C., Davenport, C., Gonzales, S., Cuzon Carlson, V. C., & Kroenke, C. D. (2022). Brain Functional Connectivity Mapping of Behavioral Flexibility in Rhesus Monkeys. *The Journal of Neuroscience*, *42*(24), 4867–4878. https://doi.org/10.1523/JNEUROSCI.0816-21.2022

Halassa, M. M., & Kastner, S. (2017). Thalamic functions in distributed cognitive control. *Nature Neuroscience*, *20*(12), Article 12. https://doi.org/10.1038/s41593-017-0020-1

Hampshire, A., & Owen, A. M. (2006). Fractionating attentional control using event-related fMRI. *Cerebral Cortex (New York, N.Y.: 1991)*, *16*(12), 1679–1689. https://doi.org/10.1093/cercor/bhj116

Hasselmo, M. E., & Sarter, M. (2011). Modes and models of forebrain cholinergic neuromodulation of cognition. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, *36*(1), 52–73. https://doi.org/10.1038/npp.2010.104

Heilbronner, S. R., & Hayden, B. Y. (2016). Dorsal Anterior Cingulate Cortex: A Bottom-Up View. *Annu. Rev. Neurosci*, *39*, 149–170. https://doi.org/10.1146/annurev-neuro-070815-013952

Hill, E. L. (2004). Executive dysfunction in autism. *Trends in Cognitive Sciences*, *8*(1), 26–32. https://doi.org/10.1016/j.tics.2003.11.003

Johnston, K. D., Ma, L., Schaeffer, L., & Everling, S. (2019). Alpha Oscillations Modulate Preparatory Activity in Marmoset Area 8Ad. *The Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.2703-18.2019

Kamigaki, T., Fukushima, T., & Miyashita, Y. (2009). Cognitive Set Reconfiguration Signaled by Macaque Posterior Parietal Neurons. *Neuron*, *61*(6), Article 6. https://doi.org/10.1016/j.neuron.2009.01.028

Kim, C., Cilles, S. E., Johnson, N. F., & Gold, B. T. (2012). Domain general and domain preferential brain regions associated with different types of task switching: A meta-analysis. *Human Brain Mapping*, *33*(1), 130–142. https://doi.org/10.1002/hbm.21199

Laubach, M., Amarante, L. M., Swanson, K., & White, S. R. (2018). What, If Anything, Is Rodent Prefrontal Cortex? *Eneuro*, *5*(5), Article 5. https://doi.org/10.1523/eneuro.0315-18.2018

Leong, Y. C., Radulescu, A., Daniel, R., DeWoskin, V., & Niv, Y. (2017). Dynamic Interaction between Reinforcement Learning and Attention in Multidimensional Environments. *Neuron*, *93*(2), 451–463. https://doi.org/10.1016/j.neuron.2016.12.040

Mansouri, F. A., Egner, T., & Buckley, M. J. (2017). Monitoring Demands for Executive Control: Shared Functions between Human and Nonhuman Primates. *Trends in Neurosciences*, *40*(1), Article 1. https://doi.org/10.1016/j.tins.2016.11.001

Mansouri, F. A., Matsumoto, K., & Tanaka, K. (2006). Prefrontal cell activities related to monkeys' success and failure in adapting to rule changes in a Wisconsin Card Sorting Test analog. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *26*(10), Article 10. https://doi.org/10.1523/JNEUROSCI.5238-05.2006

Mansouri, F. A., Tanaka, K., & Buckley, M. J. (2009). Conflict-induced behavioural adjustment: A clue to the executive functions of the prefrontal cortex. *Nature Reviews. Neuroscience*, *10*(2), Article 2. https://doi.org/10.1038/nrn2538

McAlonan, K., & Brown, V. J. (2003). Orbital prefrontal cortex mediates reversal learning and not attentional set shifting in the rat. *Behavioural Brain Research*, *146*(1–2), Article 1–2. https://doi.org/10.1016/j.bbr.2003.09.019

McGaughy, J., Ross, R. S., & Eichenbaum, H. (2008). Noradrenergic, but not cholinergic, deafferentation of prefrontal cortex impairs attentional set-shifting. *Neuroscience*, *153*(1), Article 1. https://doi.org/10.1016/j.neuroscience.2008.01.064

MILNER, B. (1963). Effects of Different Brain Lesions on Card Sorting: The Role of the Frontal Lobes. *Archives of Neurology*, *9*(1), Article 1. https://doi.org/10.1001/archneur.1963.00460070100010

Mitchell, A. S., & Chakraborty, S. (2013). What does the mediodorsal thalamus do? *Frontiers in Systems Neuroscience*, *7*, 37. https://doi.org/10.3389/fnsys.2013.00037

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Nakamura, K., Roesch, M. R., & Olson, C. R. (2005). Neuronal activity in macaque SEF and ACC during performance of tasks involving conflict. *Journal of Neurophysiology*, *93*(2), 884–908. https://doi.org/10.1152/jn.00305.2004

Niv, Y., Daniel, R., Geana, A., Gershman, S. J., Leong, Y. C., Radulescu, A., & Wilson, R. C. (2015). Reinforcement learning in multidimensional environments relies on attention mechanisms. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *35*(21), 8145–8157. https://doi.org/10.1523/JNEUROSCI.2978-14.2015

Oikonomidis, L., Santangelo, A. M., Shiba, Y., Clarke, F. H., Robbins, T. W., & Roberts, A. C. (2017). A dimensional approach to modeling symptoms of neuropsychiatric disorders in the marmoset monkey. *Developmental Neurobiology*, *77*(3), Article 3. https://doi.org/10.1002/dneu.22446

Oualian, C., & Gisquet-Verrier, P. (2010). The differential involvement of the prelimbic and infralimbic cortices in response conflict affects behavioral flexibility in rats trained in a new automated strategy-switching task. *Learning & Memory (Cold Spring Harbor, N.Y.)*, *17*(12), 654–668. https://doi.org/10.1101/lm.1858010

Ouhaz, Z., Perry, B. A. L., Nakamura, K., & Mitchell, A. S. (2022). Mediodorsal Thalamus Is Critical for Updating during Extradimensional Shifts But Not Reversals in the Attentional Set-Shifting Task. *eNeuro*, *9*(2), ENEURO.0162-21.2022. https://doi.org/10.1523/ENEURO.0162-21.2022

Pantelis, C., Barber, F. Z., Barnes, T. R. E., Nelson, H. E., Owen, A. M., & Robbins, T. W. (1999). Comparison of set-shifting ability in patients with chronic schizophrenia and frontal lobe damage. *Schizophrenia Research*, *37*(3), 251–270. https://doi.org/10.1016/S0920-9964(98)00156-X

Prins, N. W., Pohlmeyer, E., Debnath, S., Mylavarapu, R., Geng, S., Sanchez, J. C., Rothen, D., & Prasad, A. (2017). Common Marmoset (Callithrix jacchus) as a Primate Model for Behavioral Neuroscience Studies. *Journal of Neuroscience Methods*, *284*, 35–46. https://doi.org/10.1016/j.jneumeth.2017.04.004

Radulescu, A., Niv, Y., & Ballard, I. (2019). Holistic Reinforcement Learning: The Role of Structure and Attention. *Trends in Cognitive Sciences*, *23*(4), 278–292. https://doi.org/10.1016/j.tics.2019.01.010

Ragozzino, M. E., Detrick, S., & Kesner, R. P. (1999). Involvement of the Prelimbic–Infralimbic Areas of the Rodent Prefrontal Cortex in Behavioral Flexibility for Place and Response Learning. *The Journal of Neuroscience*, *19*(11), 4585–4594. https://doi.org/10.1523/JNEUROSCI.19-11-04585.1999

Ragozzino, M. E., Ragozzino, K. E., Mizumori, S. J. Y., & Kesner, R. P. (2002). Role of the dorsomedial striatum in behavioral flexibility for response and visual cue discrimination learning. *Behavioral Neuroscience*, *116*(1), 105–115. https://doi.org/10.1037//0735-7044.116.1.105

Ragozzino, M. E., & Rozman, S. (2007). The effect of rat anterior cingulate inactivation on cognitive flexibility. *Behavioral Neuroscience*, *121*(4), Article 4. https://doi.org/10.1037/0735-7044.121.4.698

Rich, E. L., & Shapiro, M. L. (2007). Prelimbic/Infralimbic Inactivation Impairs Memory for Multiple Task Switches, But Not Flexible Selection of Familiar Tasks. *The Journal of Neuroscience*, *27*(17), Article 17. https://doi.org/10.1523/JNEUROSCI.0369-07.2007

Roberts, A. C., Robbins, T. W., & Everitt, B. J. (1988). The effects of intradimensional and extradimensional shifts on visual discrimination learning in humans and non-human primates. *The Quarterly Journal of Experimental Psychology Section B*, *40*(4), Article 4. https://doi.org/10.1080/14640748808402328

Rushworth, M. F. S., Kennerley, S. W., & Walton, M. E. (2005). Cognitive neuroscience: Resolving conflict in and over the medial frontal cortex. *Current Biology : CB*, *15*(2), Article 2. https://doi.org/10.1016/j.cub.2004.12.054

Schaeffer, D. J., Gilbert, K. M., Gati, J. S., Menon, R. S., & Everling, S. (2019). Intrinsic Functional Boundaries of Lateral Frontal Cortex in the Common Marmoset Monkey. *The Journal of Neuroscience*, *39*(6), Article 6. https://doi.org/10.1523/jneurosci.2595-18.2018

Schmitt, L. I., Wimmer, R. D., Nakajima, M., Happ, M., Mofakham, S., & Halassa, M. M. (2017). Thalamic amplification of cortical connectivity sustains attentional control. *Nature*, *545*(7653), Article 7653. https://doi.org/10.1038/nature22073

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

Scolari, M., Seidl-Rathkopf, K. N., & Kastner, S. (2015). Functions of the human frontoparietal attention network: Evidence from neuroimaging. *Current Opinion in Behavioral Sciences*, *1*, 32–39. https://doi.org/10.1016/j.cobeha.2014.08.003

Seamans, J. K., & Floresco, S. B. (2022). Event-based control of autonomic and emotional states by the anterior cingulate cortex. *Neuroscience & Biobehavioral Reviews*, *133*, 104503. https://doi.org/10.1016/j.neubiorev.2021.12.026

Settlage, P., Butler, R., & Odoi, H. (1956). Perseverative Interference in Monkeys Following Bilateral Removal of the Prefrontal Areas. *The Journal of General Psychology*, *54*(2), 255–262. https://doi.org/10.1080/00221309.1956.9920281

Shiba, Y., Kim, C., Santangelo, A. M., & Roberts, A. C. (2015). Lesions of either anterior orbitofrontal cortex or ventrolateral prefrontal cortex in marmoset monkeys heighten innate fear and attenuate active coping behaviors to predator threat. *Frontiers in Systems Neuroscience*, *8*, 250. https://doi.org/10.3389/fnsys.2014.00250

Steinke, A., Lange, F., & Kopp, B. (2020). Parallel model-based and model-free reinforcement learning for card sorting performance. *Scientific Reports*, *10*(1), 15464. https://doi.org/10.1038/s41598-020-72407-7

Sutton, R. S., & Barto, A. G. (n.d.). *Reinforcement Learning: An Introduction Second edition, in progress*. https://web.stanford.edu/class/psych209/Readings/SuttonBartoIPRLBook2ndEd.pdf

Tait, D. S., Chase, E. A., & Brown, V. J. (2014). Attentional set-shifting in rodents: A review of behavioural methods and pharmacological results. *Current Pharmaceutical Design*, *20*(31), 5046–5059. https://doi.org/10.2174/1381612819666131216115802

Talwar, A., Cormack, F., Huys, Q. J. M., & Roiser, J. P. (2024). A hierarchical reinforcement learning model explains individual differences in attentional set shifting. *Cognitive, Affective & Behavioral Neuroscience*, *24*(6), 1008–1022. https://doi.org/10.3758/s13415-024-01223-7

Totah, N. K., Logothetis, N. K., & Eschenko, O. (2015). Atomoxetine accelerates attentional set shifting without affecting learning rate in the rat. *Psychopharmacology*, *232*(20), Article 20. https://doi.org/10.1007/s00213-015-4028-5

Uddin, L. Q. (2021). Cognitive and behavioural flexibility: Neural mechanisms and clinical considerations. *Nature Reviews Neuroscience*, *22*(3), 167–179. https://doi.org/10.1038/s41583-021-00428-w

Vijayraghavan, S., Major, A. J., & Everling, S. (2017). Neuromodulation of Prefrontal Cortex in Non-Human Primates by Dopaminergic Receptors during Rule-Guided Flexible Behavior and Cognitive Control. *Frontiers in Neural Circuits*, *11*, 91–91. https://doi.org/10.3389/fncir.2017.00091

Wilson, R. C., & Niv, Y. (2011). Inferring relevance in a changing world. *Frontiers in Human Neuroscience*, *5*, 189. https://doi.org/10.3389/fnhum.2011.00189

Yang, F.-C., Jacobson, T. K., & Burwell, R. D. (2017). Single neuron activity and theta modulation in the posterior parietal cortex in a visuospatial attention task. *Hippocampus*, *27*(3), 263–273. https://doi.org/10.1002/hipo.22691