7-13-2021 2:15 PM

Neural substrates of reward, error, and effort processing underlying adaptive motor behaviour

Dimitrios J. Palidis, *The University of Western Ontario*

Supervisor: Paul, Gribble L., *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Neuroscience

Abstract

Human motor control is highly adaptive to new tasks and changing environments. Motor adaptation relies on multiple dissociable processes that function to increase attainment of reward and to reduce sensory error and physical effort as costs. This thesis tests the hypothesis that fronto-striatal and dopaminergic neural systems contribute to specific aspects of motor adaptation that occur through reinforcement of rewarding actions.

Behavioral tasks were designed to isolate learning in response to feedback conveying information about reward, error, and physical effort. We also measured behavioral effects of savings and anterograde interference, by which memories from previous motor learning can facilitate or impair subsequent learning. Electroencephalography (EEG) was used to record neural event-related potentials (ERPs) elicited by task-related feedback. We measured the feedback-related negativity/ reward positivity (FRN/RP), a midfrontal component of ERP responses to feedback stimuli that correlates with neural activity throughout fronto-striatal circuits. Levodopa, a dopamine precursor, was used to manipulate dopamine release in healthy volunteers, as it has been shown to impair reward-based learning in various cognitive tasks.

We first determined that medial frontal feedback processing indexed by the FRN/RP is a specific neural correlate of reward prediction error during motor adaptation, and that the FRN/RP is not elicited by sensory error. Next, we found that levodopa did not affect either the FRN/RP, reward-based motor adaptation, savings, or anterograde interference. Finally, we determined that medial frontal activity indexed by the FRN/RP does not respond to physical effort as a cost that discounts the value of reward. However, effort increased neural sensitivity to reinforcement outcomes in activity measured by a midfrontal ERP component that was spatially and temporally distinct from the FRN/RP.

These findings suggest that mid-frontal feedback processing measured by the FRN/RP may play a specific role in reward-based motor learning that is distinct from error- and effort-based learning processes. Our findings also indicate that reinforcement learning

mechanisms that contribute to motor adaptation do not depend on the same dopaminergic processes that are impaired by levodopa in cognitive learning tasks.

Keywords

Human; reaching; motor learning; motor adaptation; reward; reinforcement; effort; EEG; feedback-related negativity; reward positivity; P300; dopamine; levodopa; savings; anterograde interference; visuomotor rotation; force field

Summary for Lay Audience

People have a remarkable ability to adapt their movements to changing conditions. Most research about motor adaptation has studied how people adjust movements according to spatial errors. For example, when the wind begins to push a tennis player's serve in an unexpected direction, the player sees the placement of each serve and adapts to compensate for the errors they experience visually. Movements can also be adapted and refined through reinforcement learning. In reinforcement-based motor learning, variable movements are produced during repeated practice, and we learn to repeat the movements that result in successful or rewarding outcomes. For example, a tennis player might learn to adjust their stance in a way that provides more power and wins more points. People also adapt their movements to be more efficient and require less effort. One hypothesis is that the brain treats effort as a cost that discounts the value of rewards.

When decisions result in rewarding outcomes, a chemical called dopamine is released in the brain. Dopamine is thought to cause changes in reward-processing brain areas that reinforce successful decisions. We tested whether similar mechanisms also contribute to reinforcement-based motor learning. We performed experiments in which people learned to adapt reaching arm movements based on rewards and errors. Some people took a drug called levodopa that impairs reinforcement-based learning in cognitive decision making tasks by overstimulating dopamine release. We placed electrodes on participants' heads to measure electrical activity from the brain in response to reward and error during learning. We also performed an experiment to test whether physical effort affects brain responses to reward as a cost.

We identified electrical signals from frontal areas in the brain that may reflect a specific mechanism for reinforcement-based motor learning. This brain mechanism was sensitive only to reward or success, while signals from different brain areas reflected effort and error. Levodopa did not affect these specific brain responses to reward, nor did it affect motor learning across a variety of tasks. These results indicate that motor learning does

not depend on the same dopamine learning mechanism that is impaired by levodopa in cognitive decision-making tasks.

Co-Authorship Statement

Dimitrios J. Palidis wrote chapter 1, Paul L. Gribble assisted with editing the text.

A version of chapter 2 has been published: Palidis DJ, Cashaback JGA, Gribble PL (2019). Neural signatures of reward and sensory error feedback processing in motor learning. Journal of neurophysiology, 121(4), 1561–1574. **Contributions:** DJP, JGAC, and PLG conceived and designed research; DJP performed experiments; DJP analyzed data; DJP, JGAC, and PLG interpreted results of experiments; DJP prepared figures; DJP wrote the manuscript; DJP, JGAC, and PLG edited the manuscript. Approximate total contribution by DJP: 80%.

The studies contained in chapters 3 and 4 are in press by the Journal of Neurophysiology as a single manuscript. The manuscript is available online in preprint form. The preprint web address will contain a link to the final peer reviewed article. Palidis DJ, McGregor HR, Vo A, MacDonald PA, Gribble PL (2021). Null effects of levodopa on reward- and error-based motor adaptation, savings, and anterograde interference. BioRxiv, doi: https://doi.org/10.1101/2020.11.19.390302. **Chapter 3 contributions:** DJP, PLG, PAM designed research. DJP performed experiments. DJP analyzed data. DJP wrote the manuscript. DJP, HRM, AV, PAM, PLG edited the manuscript. Approximate total contribution: 80%. **Chapter 4 contributions:** HRM, PLG, PAM designed research. AV performed experiments. DJP analyzed data. DJP wrote the manuscript. DJP, HRM, AV, PAM, PLG edited the manuscript. Approximate total contribution by DJP: 35%.

A version of chapter 5 has been published: Palidis DJ, Gribble PL (2020). EEG correlates of physical effort and reward processing during reinforcement learning. Journal of neurophysiology, 124(2), 610–622. **Contributions:** DJP conceived and designed research; DJP performed experiments; DJP analyzed data; DJP and PLG interpreted results of experiments; DJP prepared figures; DJP drafted manuscript; DJP and PLG edited and revised manuscript; DJP and PLG approved final version of manuscript. Approximate total contribution by DJP: 85%.

Acknowledgments

I would like to thank my supervisor, Paul Gribble, for his guidance, support, and exceptional generosity. Paul always encouraged me and trusted me with my projects. He provided many opportunities for me to learn and grow, and the freedom to pursue my interests and ideas. I would also like to thank my previous supervisors: David Ostry for setting me on this path, and Miriam Spering for the training and experience that prepared me for graduate school.

I would like to thank Heather McGregor, Josh Cashaback, and the other members of the Gribble lab for their encouragement, advice, and technical training. I would like to thank Penny MacDonald and Andrew Vo for their assistance with the levodopa protocols. The entire sensory-motor group at the Brain and Mind Institute was incredibly valuable to my training, particularly Jörn Diedrichsen, Andrew Pruszynski, and the members of their respective labs. I am grateful to all of my colleagues and classmates for their friendships, and for the good times we shared.

I would like to thank my parents for propelling me through life with loving kindness and enduring support. They stimulated my interest in science from a young age, they paid for my education, and they continue to encourage and support me. I would also like to thank Anca for her companionship and care throughout this process.

Table of Contents

Abstract ii

Summary for Lay Audience iv

Co-Authorship Statement vi

Acknowledgments vii

Table of Contents viii

List of Tables xii

List of Figures xiii

List of Abbreviations xv

1 General Introduction 1

1.1 Overview 2

1.2 Literature review 5

1.2.1 Sensory error-based motor adaptation 6

1.2.2 Reward-based motor adaptation 13

1.2.3 Savings and anterograde interference 21

1.2.4 Effort-based motor adaptation 24

1.2.5 Event related potentials elicited by feedback 28

1.3 References 33

2 Neural signatures of reward and sensory error feedback processing in motor learning 71

2.1 Introduction 72

2.2 Materials and Methods 75

2.2.1 Experimental Design and Statistical Analysis 75

2.2.2 Participants 76

2.2.3 Apparatus/Behavioral Task 76
2.2.4 EEG Data Acquisition 83
2.2.5 Behavioral Data Analysis 84
2.2.6 EEG Data Denoising 85
2.2.7 Event Related Potential Averaging 85
2.2.8 Feedback-Related Negativity/ Reward Positivity Analysis 86
2.2.9 P300 Analysis 87
2.3 Results 88
2.3.1 Behavioral Results 88
2.3.2 Feedback-Related Negativity/ Reward Positivity Results 91
2.3.3 P300 Results 94
2.4 Discussion 97
2.5 References 100
3 Effects of levodopa on neural and behavioral responses to feedback during motor adaptation 106
3.1 Introduction 107
3.2 Materials and methods 110
3.2.1 Participants 110
3.2.2 Experimental design 111
3.2.3 Apparatus/Behavioral Task 112
3.2.4 EEG data acquisition 116
3.2.5 Behavioral data analysis 117
3.2.6 EEG preprocessing 118
3.2.7 EEG data analysis 118
3.2.8 Statistics 120

3.3 Results 121
3.3.1 Control measures 121
3.3.2 Behavioral results 124
3.3.3 Event-related potential results 128
3.4 Discussion 131
3.5 References 136
4 Effects of levodopa on force-field adaptation, savings and interference 152
4.1 Introduction 153
4.2 Materials and Methods 156
4.2.1 Participants 156
4.2.2 Procedure 156
4.2.3 Data analysis 159
4.2.4 Statistics 161
4.3 Results 162
4.3.1 Control measures 162
4.3.2 Force field adaptation results 162
4.4 Discussion 169
4.5 References 171
5 EEG correlates of physical effort and reward processing during reinforcement learning 178
5.1 Introduction 179
5.2 Materials and methods 182
5.2.1 Participants 182
5.2.2 Experimental setup 182
5.2.3 EMG and EEG recording 183

5.2.4 Visual feedback of EMG 183
5.2.5 Experimental task 184
5.2.6 Behavioral analysis 188
5.2.7 EEG preprocessing and denoising 189
5.2.8 Event related potential analysis 189
5.3 Results 193
5.3.1 Behavioral results 193
5.3.2 ERP results 193
5.3.3 IC-ERP results 198
5.4 Discussion 202
5.5 References 207
6 General Discussion 216
6.1 Summary of results 217
6.2 Limitations 220
6.3 The role of midfrontal reinforcement signals in motor adaptation 223
6.4 The P300 ERP component is modulated by error, reinforcement, and effort 228
6.5 The role of dopamine in motor adaptation 231
6.6 References 235
Appendix: Ethics Approval 251
Curriculum Vitae 2533

List of Tables

Table 3-1: Control measurements 123

Table 4-1: Control measurements 157

Table 4-2: Statistical results 167

List of Figures

Figure 2-1: Experimental setup 77

Figure 2-2: Reach angles of a representative participant 81

Figure 2-3: Behavior during the reward-based learning task 89

Figure 2-4: Behavior during the visuomotor rotation task 90

Figure 2-5: Fronto-central event related potential responses to reinforcement feedback 92

Figure 2-6: Fronto-central event related potential responses to sensory error feedback 93

Figure 2-7: Centro-parietal event related potential responses to reinforcement feedback 95

Figure 2-8: Centro-parietal event related potential responses to sensory error feedback. 96

Figure 3-1: Experimental setup 113

Figure 3-2: Behavior during the reward-based learning task 125

Figure 3-3: Reward induced modulation of trial-by-trial variability of reach angle 126

Figure 3-4: Behavior during the visuomotor rotation task 127

Figure 3-5: Event-related potentials elicited by reinforcement feedback 129

Figure 3-6: Event-related potentials elicited by sensory error feedback 130

Figure 4-1: Perpendicular deviation of reach trajectory during non-catch trials 163

Figure 4-2: Perpendicular deviation of reach trajectory during catch trials 164

Figure 4-3: Adaptation effects 165

Figure 5-1: Stimuli 186

Figure 5-2: Adaptive responses to effort and reward 194

Figure 5-3: Event-related potentials elicited by reinforcement feedback 195

Figure 5-4: Event-related potentials elicited by effort condition feedback 197

Figure 5-5: Independent components of event-related potentials (IC-ERPs) elicited by reinforcement feedback 199

Figure 5-6: Independent components of event-related potentials (IC-ERPs) elicited by effort condition feedback 200

List of Abbreviations

ACC	Anterior cingulate cortex
BF	Bayes Factor
CCW	Counterclockwise
CCWFF	Counterclockwise force field
CW	Clockwise
CWFF	Clockwise force field
dACC	Dorsal anterior cingulate cortex
DF	Degrees of freedom
EEG	Electroencephalography
EMG	Electromyography
EOG	Electrooculogram
ERP	Event-related potential
FDR	False discovery rate
FF	Force field
FF1a	Force field 1a (clockwise force field)
FF1b	Force field 1b (clockwise force field)
FF2	Force field 2 (counterclockwise force field)
FRN	Feedback-related negativity

IC-ERP	Independent component event related potential
ICA	Independent components analysis
LTD	Long term depression
LTP	Long term potentiation
MSN	Medium spiny neuron
MVC	Maximum voluntary contraction
NF	Null field
NFa	Null field a
NFb	Null field b
P	P-value
PD	Perpendicular deviation
ROI	Region of interest
RP	Reward positivity
SD	Standard deviation
SEM	Standard error of the mean
T	T-statistic
VMR	Visuomotor rotation
δ	Cohen's δ (effect size)

Chapter 1

General Introduction

1.1 Overview

Human motor behavior is highly adaptable. People learn to select actions and produce movements in ways that reduce error, reduce effort, and increase reward attainment. Behavioral adaptation occurs largely through feedback mechanisms by which the observed consequences of actions drive changes in subsequent motor output. Multiple processes contribute to motor adaptation with distinct neural substrates, timescales, and behavioral signatures.

Cortico-cerebellar circuits are heavily implicated in sensory-error based adaptation, which functions to rapidly reduce errors when the direct sensory consequences of motor output are perturbed (Shadmehr et al., 2010). Recent work has characterized an additional distinct process of reward-based motor adaptation, by which the parameters of movements are flexibly adapted in order to increase task success or the obtainment of extrinsic rewards (Izawa & Shadmehr, 2011). Plasticity in fronto-striatal circuits mediated by midbrain dopamine signaling is broadly thought to underlie reward-based learning for the selection of discrete actions (Schultz, 2016). However, it is not yet known whether the same processes also shape the kinematics and dynamics of movement according to reward outcomes.

Movements are selected and adapted not only to increase reward attainment but also to reduce physical effort as a cost (Selinger et al., 2015; Summerside et al., 2018). Reward circuits involving frontal cortical areas, the striatum, and the dopaminergic midbrain have also been implicated in motivation and decision making related to physical effort (Hauser et al., 2017; Klein-Flügge et al., 2016; Kurniawan et al., 2011; Suzuki et al., 2020). However, little is known regarding the neural learning mechanisms by which behavior is adapted and refined to reduce effort. This thesis presents a series of experiments testing the hypothesis that fronto-striatal and dopaminergic circuits underlie specific forms of motor adaptation that are characterized by the reinforcement of actions according to reward value and effort cost.

In the experiment presented in chapter 2, we designed behavioral paradigms to dissociate the neural correlates of feedback processing in sensory-error based motor adaptation and reward-based motor adaptation of reaching movements. Sensory-error based adaptation was isolated by perturbing visual feedback of hand position without disrupting the attainment of task goals or reward. Reward-based learning was isolated by manipulating binary reinforcement feedback indicating whether or not reward was attained while withholding visual feedback of the arm. We used electroencephalography (EEG) to measure event-related potentials (ERPs) elicited by sensory error and reward feedback. We found that a medial-frontal ERP component called the feedback-related negativity, or alternatively the reward positivity (FRN/RP) was elicited specifically by reinforcement feedback, while sensory error modulated a distinct posterior ERP component called the P300. The FRN/RP is a well characterized neural correlate of reward processing and a prominent theory states that it reflects a reinforcement learning process in the medial-frontal cortex (Walsh & Anderson, 2012). This result suggests that reinforcement signaling in the medial frontal cortex may play a role in motor adaptation that is specific to reward-based learning processes.

A role of phasic dopamine signaling in learning from reward outcomes has been characterized extensively by decades of research. A prominent theory of the FRN/RP suggests that it reflects a reinforcement learning process driven by phasic dopaminergic reward signaling (Holroyd & Coles, 2002). In the experiment presented in chapter 3, we tested the hypothesis that dopamine signaling mediates the neural processing of reward outcomes measured by the FRN/RP, and that this process underlies reward-based motor adaptation. We manipulated dopamine transmission by administering levodopa, a dopamine precursor, to healthy participants in a randomized, placebo controlled design. We recorded EEG while participants performed sensory error- and reward-based motor adaptation tasks similar to those described in chapter 1. We replicated our previous finding that the FRN/RP occurs specifically during reward-based motor adaptation. However, we did not detect any effects of levodopa on the FRN/RP or motor adaptation. Numerous studies have shown that levodopa impairs reward-based learning in tasks that

involve the selection of discrete actions, and there is evidence to suggest that levodopa impairs learning by disrupting neural activity in the ventral striatum (Cools et al., 2007; Hiebert et al., 2019; Vo et al., 2016). Our results suggest that reward-based motor adaptation and the FRN/RP likely do not depend on the same mechanisms shown to be impaired by levodopa in cognitive reward-based learning tasks. These results also add to a small number of recent studies that did not detect associations between dopamine function and reward-based motor adaptation (Holland et al., 2019; Quattrocchi et al., 2018).

Adaptation in sensory-error based learning tasks occurs more quickly when a particular perturbation is encountered a second time after washout of initial learning, a phenomenon known as savings (Coltman et al., 2019). Previous adaptation to a particular perturbation also causes anterograde interference during subsequent adaptation to an opposite perturbation (Miall et al., 2004). Some findings suggests that savings and anterograde interference may result from reinforcement of motor commands upon successful error reduction due to sensory error-based learning (Huang et al., 2011). In the experiment described in chapter 4, we tested the hypothesis that dopamine signaling mediates savings and anterograde interference. Participants adapted their movements to compensate for velocity-dependent forces applied to their hands during reaching movements. We measured adaptation, savings and anterograde interference and manipulated dopamine release by administering levodopa or placebo in a randomized, double-blinded design. Contrary to our hypothesis, we observed no effects of levodopa on adaptation, savings, or anterograde interference. These results suggests that savings and interference do not depend on reinforcement learning processes previously shown to be impaired by levodopa.

Humans not only adapt decision making and movement parameters to increase reward attainment, but also to reduce motor effort and energetic requirements. Economic models suggest that movement parameters are not necessarily determined to maximize reward but rather subjective utility, a function of reward discounted by effort (Shadmehr et al., 2016a). Neural structures implicated in reward learning including the striatum,

medial frontal cortex, and dopaminergic midbrain have also been shown to play a role in motivating effortful behavior and integrating effort with reward during decision making. In the experiment described in chapter 5, we tested the hypothesis that medial-frontal learning signals are sensitive to subjective utility characterized by effort-discounted reward. Participants performed a task in which they were required to accurately produce target levels of muscle activation to receive reward. Physical effort requirements varied probabilistically according to choices such that participants could adapt their behavior to reduce effort requirements. We found that the FRN/RP did not respond to feedback that indicated the effort requirements resulting from participants' choices. Reward feedback elicited an FRN/RP component that was sensitive only to reward, followed by a distinct medial-frontal signal that was modulated by an interaction between effort and reward. This result provides evidence that while FRN/RP reflects reward processing irrespective of effort, later responses to reward outcomes integrate information about preceding effort.

In summary, this thesis presents scientific findings characterizing neural and behavioral responses to task-relevant feedback in the context of adaptive motor behavior. We found that a medial-frontal ERP component called the FRN/RP measures neural responses that are specific to reinforcement outcomes while distinct ERP components are modulated by motor effort and sensory error. Contrary to our predictions, we found no evidence for a dopaminergic basis of the FRN/RP, reward-based motor adaptation, savings, or anterograde interference.

1.2 Literature review

The following review surveys scientific literature necessary to motivate the work presented in this thesis. **Section 1.2.1** reviews sensory error-based motor adaptation, and **section 1.2.2** reviews reward-based motor adaptation. These topics are relevant to chapters 2 and 3 of this thesis, which present work that aims to identify neural substrates of reward-based learning and to dissociate them from mechanisms of sensory error-based learning. **Section 1.2.3** reviews the phenomena of savings and anterograde interference in

motor adaptation. This section is relevant to chapter 4, which details an experiment testing whether savings and interference are affected by levodopa. **Section 1.2.4** reviews effort-based learning. This topic motivates chapter 5, which presents an experiment testing whether reinforcement signals in the medial frontal cortex are also sensitive to effort as a cost. Finally, **section 1.2.5** reviews electrophysiological correlates of feedback processing related to behavioral adaptation, as measured by EEG. Chapters 2 and 3 use EEG to measure neural responses to sensory error and reward feedback. Chapter 5 reports EEG responses to effort feedback as well as effects of effort on EEG responses to reward feedback.

1.2.1 Sensory error-based motor adaptation

Sensory feedback plays an essential role in the control of movement. A loss of proprioceptive input from the limbs due to peripheral deafferentation has devastating effects on motor function (Marsden et al., 1984; Rothwell et al., 1982). Despite the critical importance of sensory feedback for movement, the noisy and delayed nature of afferent signals poses significant problems that the sensorimotor system must overcome. There is substantial evidence to suggest that the nervous system achieves accurate control by predicting the sensory consequences of actions and adapting these predictions in response to errors. This section first discusses purported functions of sensory prediction in motor control. Second, it reviews how predictions regarding the sensory consequences of motor commands are thought to improve perception. The third part of this sections reviews how errors in sensory prediction are thought to drive sensory-error based motor adaptation, a topic important to the work presented in chapters 2, 3, and 4. The final part of this section reviews the neural basis of sensory prediction and sensory-error based learning.

Sensory prediction aids control:

Saccadic eye movements are generally too brief for sensory feedback to be used effectively in control. The delay at which visual information can be transmitted to the extraocular muscles is around 70ms, longer than the duration of many saccades (B. Fischer & Boch,

1983). However, if the brain can quickly and accurately predict the sensory consequences of motor commands sent to the muscles, these predictions can be used in place of afferent feedback. Guthrie et al., (1983) electrically stimulated the superior colliculus of monkeys to induce eye movements immediately before the onset of a voluntary saccade to a visual target. Although the stimulation-induced movements perturbed the position of the eyes, the subsequent voluntary saccades still accurately reached their targets despite removal of any visual or proprioceptive input. Thus, motor output signals from the colliculus can be used to accurately predict changes in eye position in the absence of sensory feedback. This mechanism may also be used to predict errors caused by inaccurate motor commands and issue corrections during the course of a movement. For example, repetition of a visual target or transcranial magnetic stimulation can affect motor commands and result in a reduction of saccade velocity (Xu-Wilson et al., 2009, Xu-Wilson et al., 2011). However, compensatory motor commands arriving later in the movement allow the eye to accurately reach the target. These findings suggest that errors due to variability in motor commands can be anticipated and corrected through internal feedback.

Unlike saccades, the durations of limb movements are often long enough for sensory feedback to influence control. However, the feedback loop between sensory input and motor output can contain total delays upwards of 100ms. Delays in feedback control systems result in instabilities with high feedback gains, and larger delays necessitate lower gains to maintain stable control (Miall & Wolpert, 1996). Feedback delays can be overcome by using an internal "forward model" to predict the sensory consequences of actions. These predictions can be used for control in an internal feedback loop without delay (Bhushan & Shadmehr, 1999; Mehta & Schaal, 2002; Miall et al., 1993; Todorov, 2004). Delayed external feedback can be used to update the forward model and to respond to unpredictable errors. Humans often respond to self-generated forces without delay (Diedrichsen et al., 2003; Flanagan & Wing, 1993; Gribble & Ostry, 1999; Kawato et al., 2003; Kurtzer et al., 2008) and estimate the state of the body in real time despite sensory delays (Miall et al., 2007; Vercher et al., 1996; Wagner & Smith, 2008). These

findings suggest that humans may predict the sensory consequences of motor commands to achieve stable and accurate control.

Sensory prediction improves perception:

In addition to improving control, sensory prediction can also aid perception. Sensory input often provides noisy and limited information. If the brain can predict the sensory consequences of motor commands, these predictions can be combined with noisy sensory information to estimate the state of the body with improved reliability. It has been shown that motor output improves perception and state-estimation across a variety of tasks and sensory modalities (Fuentes & Bastian, 2010; Izawa & Shadmehr, 2008; Körding & Wolpert, 2004; Sommer & Wurtz, 2006; Spering et al., 2011; Vaziri et al., 2006; Wolpert et al., 1995).

Predicting the sensory consequences of motor commands may also be important to distinguish self-generated sensory input from externally generated input (J. X. Brooks & Cullen, 2019). Predictable self-generated sensory signals are often cancelled or suppressed in the brain. For example, we maintain a stable and continuous perception of the visual world despite frequent rapid movements of the eyes (Wurtz, 2008, 2018). Suppression of self-generated sensation in the somatosensory system is evident by the fact that we are unable to tickle ourselves and that self-generated forces are perceived as being considerably smaller than externally-generated forces (Blakemore et al., 1998; Shergill et al., 2003). This somatosensory attenuation occurs across levels of sensory processing from the spinal cord to the somatosensory cortex (Chapman, 1994; Fetz et al., 2002; Seki et al., 2003; Seki & Fetz, 2012).

Filtering of predicted self-generated inputs may also be useful to enhance inputs caused by external perturbations (J. X. Brooks & Cullen, 2019). However, cancellation of self-generated input is not always observed and can be complex. Active movement can enhance proprioceptive inputs at the level of the spinal cord (Confais et al., 2017). Vestibular reafference may only be cancelled if proprioceptive sensory signals match the

predicted input, indicating a more complex mechanism than simple cancellation of self-generated input (J. X. Brooks & Cullen, 2019).

Sensory error drives adaptive changes in control and perception:

Predictions regarding the sensory consequences of actions contribute to accurate motor control and perception. To remain accurate, these predictions must be adaptable to changing dynamics that dictate the sensory consequences of motor commands. The mechanical properties of the body change throughout development, and the response properties of muscles may change with fatigue on a much shorter timescale. We must adapt to new dynamics every time we use a new tool. Perception and motor output tend to drift in the absence of sensory feedback, indicating that frequent recalibration may be required even in the absence of perturbations (L. E. Brown et al., 2003; Patterson et al., 2017; Wann & Ibrahim, 1992).

Two prominent experimental models of motor learning have been used in recent work in this area: force field adaptation and visuomotor rotation (VMR) tasks. In studies of force field adaptation, a robot applies velocity-dependent forces to the hand during reaches to targets (Shadmehr & Mussa-Ivaldi, 1994). In VMR tasks, a cursor on a digital display represents the position of the hand, and the mapping between the actual reach angle and the position of the cursor is rotated to induce errors (Krakauer et al., 2000). Optical prism goggles can also be used to impose visuomotor rotation by shifting the entire visual field. In both kinds of tasks participants quickly adapt their movements to compensate for the experimentally induced perturbations (i.e., external forces or visual feedback rotation, respectively).

Simple state-space models provide a good descriptive account of behavioral learning in these motor adaptation tasks. In typical state-space models, adaptation proceeds by compensating for a fixed proportion of errors experienced on each trial, resulting in exponential learning curves (Diedrichsen et al., 2005; O. Donchin et al., 2003; Thoroughman & Shadmehr, 2000). Bayesian models of adaptation suggest that the learning rate is not necessarily fixed but is affected by the uncertainty, consistency, and

relevance of errors (Berniker & Kording, 2011; Gonzalez Castro et al., 2014; Tan et al., 2014; Wei & Koerding, 2010). It has also been shown that multiple processes may contribute to adaptation in parallel at different timescales. This notion has been formalized with a two-state model of adaptation, in which learning is the sum of slow and a fast adaptation processes (M. A. Smith et al., 2006). The slow process is characterized by a lower learning rate and a lower rate of forgetting, while the fast process is characterized by a high learning rate but also a high rate of forgetting (J.-Y. Lee & Schweighofer, 2009; M. A. Smith et al., 2006; H. Tanaka et al., 2012).

Two processes that have been functionally differentiated in common motor adaptation tasks are explicit and implicit learning (McDougle et al., 2016; Schween et al., 2020; Taylor et al., 2014). Explicit learning constitutes the use of a strategy that is under conscious awareness and control. In VMR tasks, explicit learning occurs when the participant notices the perturbation and strategically aims their reaches to counteract the rotation. Implicit learning occurs without conscious awareness or control. If perturbations are introduced gradually, implicit learning can produce large adaptive changes in motor output without awareness (Rand & Heuer, 2019).

A prominent theory of motor adaptation states that learning is driven by sensory prediction errors that occur when sensory feedback differs from the predictions of an internal forward model. VMR and force field perturbations are thought to cause sensory prediction errors that are used to update the forward model. Accurate control is restored as the sensorimotor system learns to predict the effects of visual or mechanical perturbations. Mazzoni & Krakauer (2006) provide compelling evidence that sensory prediction errors drive adaptation. Participants were instructed to counteract a 45 deg counterclockwise VMR by strategically aiming 45 deg clockwise away from the target. The aiming strategy initially eliminated errors caused by the perturbation relative to the target, but a discrepancy between the motor command and sensory feedback still occurred, resulting in sensory-prediction error. Implicit adaptation caused the reach angle to gradually shift further in the clockwise direction despite increasing errors relative to the target. This finding suggests that the nervous system implicitly predicts the direct

sensory consequences of actions, and that violations of these predictions causes adaptation irrespective of task goals or conscious strategy.

Consistent with adaptation of a forward model, motor adaptation is accompanied by perceptual changes in participants' estimates of the sensory consequences of actions (Izawa et al., 2012; Izawa & Shadmehr, 2011; Rand & Heuer, 2019). However, perceptual changes in response to adaptation have also been observed using passive learning tasks and passive sensory measures that do not involve motor output (Mostafa et al., 2019; Ohashi et al., 2019; Ostry et al., 2010; Ostry & Gribble, 2016). This indicates a role for sensory plasticity that does not involve forward prediction based on motor commands. For example, in VMR there is a process of realignment between visual and proprioceptive inputs that does not depend on motor output (Block & Bastian, 2012; Henriques et al., 2014).

Neural basis of sensory prediction and learning:

It is thought that the cerebellum is an integral component of a forward model that predicts the sensory consequences of motor commands for control and perception (Bhanpuri et al., 2014; Kawato, 1999; Miall et al., 1993, 2007; Shadmehr & Krakauer, 2008; Streng et al., 2018; Weeks et al., 2017; Wolpert et al., 1998; Xu-Wilson et al., 2009). Cerebellar damage and disruption has been shown to cause behavioral deficits in state-estimation and predictive motor control (Bastian et al., 1996; Bhanpuri et al., 2011, 2013, 2014; Haggard et al., 1995; Kakei et al., 2019; Miall et al., 2007; Müller & Dichgans, 1994; Therrien & Bastian, 2015; van Donkelaar & Lee, 1994; Vercher & Gauthier, 1988; Weeks et al., 2017; Xu-Wilson et al., 2009). Cerebellar damage has also been shown to impair learning in VMR and force field adaptation tasks (Izawa et al., 2012; Maschke et al., 2004; Morton & Bastian, 2006; Schlerf et al., 2012; M. A. Smith & Shadmehr, 2005; Synofzik et al., 2008; Tseng et al., 2007; Weiner et al., 1983). In healthy participants cerebellar activation is increased during VMR and force field adaptation (Diedrichsen et al., 2005).These findings are consistent with the notion that the cerebellum contributes to the

control of movement by predicting the sensory consequences of action and by updating these predictions in response to environmental perturbations.

An influential model of the cerebellar circuit provides a potential mechanism for an adaptable forward model. Motor areas of the cerebellar cortex receive input from the motor cortex via the pons and project back to the motor cortex via the thalamus (Sokolov et al., 2017). Largely feedforward circuitry within the cerebellar cortex may transform efferent copies of motor commands to sensory predictions. These sensory predictions may be projected back to motor cortex to produce internal feedback control loops (M. Ito, 2006; Requarth & Sawtell, 2014). Sensory predictions may also be transmitted to the inferior olive where they converge with afferent sensory feedback to compute sensory prediction error signals (M. Ito, 2006). Sensory prediction error signals in the inferior olive may be used to update the forward model by inducing long-term depression in the cerebellar cortex through climbing fiber projections. During reaching movements, simple spike firing rate in purkinje neurons within the cerebellar cortex contain signals that predict the future sensory state and encode delayed sensory feedback (Ebner & Pasalar, 2008; Hewitt et al., 2015; Pasalar et al., 2006; Popa et al., 2012, 2013; Popa & Ebner, 2019; Streng et al., 2018; H. Tanaka et al., 2019). Sensory prediction error signals are observed in the rostral fastigial nucleus, a major output of the primate cerebellar cortex, in response to externally generated head movement while predictable self-generated sensory inputs are cancelled (J. X. Brooks et al., 2015; J. X. Brooks & Cullen, 2013). These findings support the notion that the cerebellum acts as an adaptable predictor of the sensory consequences of motor commands.

Cortical sensorimotor areas are also implicated in motor adaptation. The cerebellum is densely interconnected to motor, frontal and parietal cortices through a series of parallel loops (Sokolov et al., 2017). Lesions of the parietal cortex impair VMR adaptation (Mutha et al., 2011). Neuroimaging during VMR tasks reveals changes in activation of primary motor, premotor, posterior parietal, and supplementary motor cortex (Clower et al., 1996; Diedrichsen et al., 2005; Inoue et al., 2000; Krakauer et al., 2004). Stimulation studies suggest that the primary motor cortex may be particularly involved in retention of

motor memory related to VMR (Galea et al., 2011; Hadipour-Niktarash et al., 2007). These results are corroborated by learning-induced structural changes in primary motor cortex associated with long-term memory for VMR (Landi et al., 2011). Recordings from single neurons in primary motor cortex, premotor cortex, and supplementary motor cortex reveal changes in neural encoding of movements due to VMR adaptation (S. M. Chase et al., 2012; Paz et al., 2003, 2005; Paz & Vaadia, 2004; Vyas et al., 2018; S. P. Wise et al., 1998). Force field adaptation is also associated with changes in activation and neural population coding in the primary motor, supplementary motor, and premotor cortex (Diedrichsen et al., 2005; C.-S. R. Li et al., 2001; Padoa-Schioppa et al., 2004; Perich et al., 2018; Xiao et al., 2006).

A number of studies have also highlighted a role of primary sensory cortex in motor adaptation that may correspond to perceptual changes associated with learning (Ostry & Gribble, 2016). Force field adaptation is associated with changes in neural responses to somatosensory stimuli (Nasir et al., 2013). Disruption of primary somatosensory cortex interferes with acquisition and consolidation of force field adaptation (Kumar et al., 2019; Mathis et al., 2017).

1.2.2 Reward-based motor adaptation

Reward-based learning mechanisms have been studied extensively in the context of adaptive behavior involving decisions between discrete stimuli and actions. Although studies of motor adaptation have largely focused on sensory-error based learning, influences of reward-based learning processes are increasingly reported in recent research. The first part of this section reviews behavioral findings characterizing the role of reward outcomes in motor adaptation. Second, literature regarding the neural mechanisms of reward-based motor adaptation is reviewed. Third, the function of exploration in motor learning and reward-based learning generally is discussed. These topics are relevant to the work presented in chapters 2 and 3, which examines the neurophysiological basis of reward-based motor adaptation. The final part of this section reviews computational and neurophysiological theories of reinforcement learning. This

section motivates the overarching hypothesis of this thesis that fronto-striatal and dopaminergic reinforcement learning mechanisms contribute to adaptive motor control.

Reward and success can influence adaptive motor behavior:

In typical models of sensory-error based adaptation, the learning process functions to improve prediction and control of the perceptual consequences of motor output. Improving control in this way is generally expected to improve task performance, however factors such as task success and reward are typically not considered in models of sensory-error based learning. However, there is empirical evidence to suggest that instrumental and reward-related outcomes can modulate sensory-error based learning in humans. The retention of learning in visuomotor rotation tasks is improved with the addition of monetary reward or stimuli that indicate success (Galea et al., 2015; Shmuelof et al., 2012). The addition of reward or punishment contingent on error can increase the rate of adaptation (Galea et al., 2015; Nikooyan & Ahmed, 2014). Errors related to the goals of a given task can affect sensory-error based learning similarly to explicit reward or punishment feedback. It has been shown that equivalent sensory prediction errors cause greater adaptation when they result in a failure to reach a visual target, even without extrinsic reward contingent on reaching the target (Kim et al., 2019; Leow et al., 2020).

The findings outlined above indicate that reinforcement of successful or rewarding actions can influence sensory-error based learning. It has also been demonstrated that reward-based learning and sensory-error based learning can occur as largely separable processes. Adaptation to sensory error has been shown to occur automatically even when it interferes with task performance, supporting a distinction between reinforcement of successful actions and sensory error-based learning (Mazzoni & Krakauer, 2006). Reward-based learning can be isolated in paradigms analogous to typical visuomotor rotation tasks by withholding cursor feedback of hand position and instead providing only binary reinforcement feedback indicating the success or failure of each movement (Izawa & Shadmehr, 2011). If reinforcement is contingent on producing reach angles that are rotated relative to the visual target, reward-based learning can result in adaptation of

reach angle comparable to sensory-error based learning. However, reward-based learning does not produce the same pattern of spatial generalization nor does it alter the predicted sensory consequences of actions (Izawa & Shadmehr, 2011). In typical motor adaptation tasks, sensory-error based learning processes tend to play a dominant role in adaptation, while reinforcement learning can compensate for degraded or absent sensory feedback (Cashaback et al., 2017; Izawa & Shadmehr, 2011). Reinforcement learning can also compensate for deficits in sensory error-based learning caused by injury to the cerebellum (Therrien et al., 2016).

While sensory-error based learning in gradual visuomotor rotation tasks can produce large changes in reach angle without awareness, reward-based learning for reach angle rotations beyond 5-10 degrees occurs largely through strategic and conscious responses to the experimental manipulation (Codol et al., 2018; Holland et al., 2018, 2019). Explicit strategy may be necessary for learning in reward-based rotation paradigms due to the discrepancy between rewarded movements and the visually instructed targets. Reward-based motor adaptation may occur without awareness during a paradigm in which participants always reach towards a veridical target, but success depends on adapting movements to produce particular joint angle configurations (Mehler et al., 2017). Such adaptation is possible because reaching movements are often highly redundant. For example, a given target hand position can be achieved through an infinite number of different combinations of joint angles in the arm. Although sensory error-based learning can correct biases that affect average motor error, a reinforcement learning process may be necessary to discover solutions in a redundant movement space that minimize the variance of motor error, a hallmark of motor skill acquisition (Mehler et al., 2017).

Neural mechanisms of reward-based motor learning:

Sidarta et al., (2016) used fMRI to measure changes in resting-state functional connectivity associated with reward-based motor learning. The amount of reward obtained during the task correlated with increases in connectivity between the putamen and ventromedial prefrontal cortex, regions implicated in reward processing generally.

The degree of learning was correlated with increases in connectivity throughout somatosensory and motor cortical areas. The involvement of somatosensory regions is corroborated by behavioral findings that reward-based motor learning produces increases in task-relevant somatosensory acuity (Bernardi et al., 2015). Furthermore, passive somatosensory training with reinforcement can produce learning equivalent to active learning (Bernardi et al., 2015). Both somatosensory and visuo-spatial working memory capacity have been shown to correlate with performance in reward-based motor learning tasks (Holland et al., 2019; Sidarta et al., 2016). These findings are consistent with contributions of a strategy in which participants aim towards the location of previously rewarded movements maintained in working memory, likely involving the prefrontal cortex.

Uehara et al. (2018) found that reward-based motor learning caused changes in physiological responses to non-invasive brain stimulation that indicate long-term potentiation in the primary motor cortex. Dopaminergic projections to primary motor cortex are essential for some forms of motor skill learning in rodents (Beeler et al., 2010; Hosp et al., 2011; Luft & Schwarz, 2009; Molina-Luna et al., 2009). Recently reports have revealed reward signals in premotor and primary cortices of non-human primates, including reward prediction error (An et al., 2019; Moore et al., 2020; Ramakrishnan et al., 2017; Ramkumar et al., 2016). These findings suggest a possibility that plasticity in motor cortex, possibly mediated by dopamine, could contribute to reward-based motor learning.

The role of exploration in reward-based learning:

In sensory-error based learning, errors experienced in one movement are reliably reduced in the subsequent movement, or even within the same movement (Crevecoeur et al., 2020). This is consistent with the idea that an internal sensorimotor mapping allows for corrections to be informed by the error itself. However, reinforcement learning essentially relies on a process of "trial and error". A successful action can only be reinforced after it has already been produced and rewarded. If an action is unsuccessful or leads to punishment, then it will be deterred in the future. However, errors do not necessarily

inform the correct response if there are more than one possible alternative. Thus, variability in behavior is essential for reinforcement learning to discover and select rewarding actions. In a reward-based motor learning task, participants with greater baseline variability in task-relevant dimensions of movement learned more than those with lower baseline variability (H. G. Wu et al., 2014).

In some reinforcement learning algorithms, variability of actions is actively generated and regulated to improve learning through efficient exploration (Sutton & Barto, 2011). Even if a particular action produces reward, there is still a tradeoff between continuing to exploit the known reward or exploring other actions that may potentially be more rewarding. During reward-based motor adaptation, trial-by-trial variability in movement kinematics is larger following non-reward outcomes than reward outcomes (Cashaback et al., 2019; Kooij & Smeets, 2019; Mastrigt et al., 2020; Pekny et al., 2015). An increase in motor variability following non-reward may reflect exploration in search of more valuable actions. High throughput experiments in rats have shown that movement variability is regulated according to the recent history of reward outcomes extending at least 10 trials in the past (Dhawale et al., 2019). On a much longer timescale, variability was found to be modulated by uncertainty induced by non-stationary reward contingencies.

The role of regulated motor variability in learning has been studied extensively in songbirds' vocal learning. A basal ganglia-forebrain circuit has been shown to actively generate motor variability required for learning (Kao et al., 2005; Ölveczky et al., 2011). This circuit is essential for song learning during development but not for producing songs that have already been learned (Bottjer et al., 1984). Natural variability in motor output also allows for adaptation of fully learned songs in adult songbirds through a process consistent with reinforcement learning (Andalman & Fee, 2009; Tumer & Brainard, 2007). Dopamine signaling in this circuit has been demonstrated to mediate context-dependent regulation of song variability and also to act as a learning signal that reinforces motor commands that are successful in producing vocal targets (Hoffmann et al., 2016; Kao et al., 2008; Woolley, 2019; Woolley et al., 2014). Pekny et al. (2015) demonstrated

that people with Parkinson's Disease fail to upregulate variability of arm movements following unrewarded outcomes. This suggests that dopaminergic signals in the human basal ganglia may also be important for adaptive regulation of motor variability during reward-based reach adaptation.

Theories of biological reinforcement learning:

The neural substrates of reward-based motor adaptation remain largely unexplored. However, extensive work has characterized the physiological and computational processes supporting reward-based learning in more cognitive paradigms involving discrete choices and stimuli. A natural starting point is to test for the involvement of well characterized reinforcement learning processes in reward-based motor adaptation.

In formal models of reinforcement learning, agents make choices according to a value function that represents an estimate of the expected values of reward outcomes to be gained from being in a particular state or executing a particular action (Sutton & Barto, 2011). The differences between observed reward outcomes and those predicted by the value function are known as reward prediction errors. Typically, the value estimate corresponding to an action or state is updated by a proportion of the reward prediction error that is determined by a learning rate parameter. If the actual outcome is better than predicted, a positive reward prediction error occurs and the value estimate is increased accordingly, resulting in an increased probability of selecting that action in the future. If the actual outcome is worse than expected, a negative reward prediction error serves to diminish the value estimate and deter that action. Phasic changes in the firing rate of many midbrain dopamine neurons match reward prediction error signals predicted by computational models of reinforcement learning (Bayer & Glimcher, 2005; García-García et al., 2017; Jocham & Ullsperger, 2009; Schultz et al., 1997; Waelti et al., 2001; Watabe-Uchida et al., 2017). These dopaminergic signals are thought to mediate synaptic plasticity in the striatum and frontal cortex that may underlie reward-based learning (Mohebi et al., 2019; Otani et al., 2003; Reynolds & Wickens, 2002; Shindou et al., 2019; Steinberg et al., 2013; J. X. Wang et al., 2018; R. A. Wise, 2004).

Highly influential models of biological reinforcement learning state that dopaminergic reward prediction error signals shape future behavior by biasing action selection processes in the basal ganglia. Prefrontal and motor cortices provide major inputs to the basal ganglia through synapses onto the striatum. The internal segment of the globus pallidus (GPi), the primary output of the basal ganglia, projects back to the cortex via the thalamus to form a cortico-basal ganglia-thalamic loop circuit (G. E. Alexander et al., 1986; Parent & Hazrati, 1995). Basal ganglia output through the GPi provides widespread tonic inhibition that is thought to prevent motor output in a default state. Inputs to the striatum can facilitate movements through a direct inhibitory pathway between the striatum and the GPi that serves to release the tonic inhibitory output from the basal ganglia. An indirect pathway from the striatum to the GPi passes through the external segment of the globus pallidus and the subthalamic nucleus. The indirect pathway prevents movement by bolstering the inhibitory output of the GPi. Longstanding theories of action selection in this circuit propose that its organization maintains action specific representation of movement, with parallel channels throughout the loop supporting individual actions. Early models proposed that the direct pathway serves to release actions specified by the motor cortex while the indirect pathway broadly inhibits competing actions (G. E. Alexander & Crutcher, 1990; Mink, 1996). However, this account is largely incompatible with recent reports that both the direct and indirect pathways show action specific activation during action selection (Cui et al., 2013; Klaus et al., 2017; J. G. Parker et al., 2018). Newer computation models of the cortico-basal ganglia-thalamic circuit propose that the direct and indirect pathways compete within individual action channels to facilitate and suppress specific movements, respectively (Bariselli et al., 2019; Dunovan et al., 2019; Dunovan & Verstynen, 2016).

Synapses of cortical inputs onto striatal medium spiny neurons exhibit bidirectional Hebbian spike timing dependent plasticity, being strengthened or weakened through LTP or LTD, respectively. This plasticity is strongly dependent on dopamine, suggesting a mechanism for dopaminergic reward prediction error signals to mediate reinforcement learning (Calabresi et al., 2007; Lerner & Kreitzer, 2011; Pawlak & Kerr, 2008; Perrin &

Venance, 2019; Reynolds et al., 2001; Reynolds & Wickens, 2002; Shen et al., 2008). Although the specific conditions required for LTP and LTD in cortico-striatal synapses are still not well established, current evidence is largely consistent with models of reinforcement learning in the basal ganglia circuit for action selection. Increased dopamine release signaling reward outcomes may facilitate LTP in synapses targeting direct pathway MSNs, while decreased dopamine release caused by unrewarding or punishing outcomes instead may elicit LTD. This would strengthen or weaken cortical inputs to the striatum that promote a particular action in accordance with the outcome. Interestingly, dopamine seems to have opposite effects on plasticity in indirect pathway MSNs due to different prevalence of specific dopamine receptor types (Lerner & Kreitzer, 2011; Shen et al., 2008). These findings are consistent with the oppositional role of the direct and indirect pathways in models of action selection (Rubin et al., 2020).

Reward prediction error signals are widely observed in not only in the striatum but also various subregions of the prefrontal cortex (Delgado et al., 2004; Diederen et al., 2017; Knutson & Cooper, 2005; O'Doherty et al., 2004; Oya et al., 2005; Pessiglione et al., 2006; Ramnani et al., 2004; Rogers et al., 2004; Rutledge et al., 2010; Schultz, 2016; S. C. Tanaka et al., 2004). Dopamine may drive reward prediction error signals in the frontal cortex through dense projections from the dopaminergic midbrain throughout the frontal lobes (Tzschentke, 2001). However, the prefrontal cortex seems to contain all of the information necessary to compute reward prediction error and implement a reinforcement learning process without dopaminergic inputs. Neural activity in the prefrontal cortex encodes a transformation from predicted value to choice, and an updating of value estimates according to outcomes (Barraclough et al., 2004; Hunt et al., 2015, 2018; Kennerley et al., 2008; Padoa-Schioppa & Assad, 2006; H. Seo & Lee, 2008; Tsutsui et al., 2016). These findings are recapitulated by a computational theory of the prefrontal cortex as a meta-learning system that implements a distinct reinforcement learning process entirely through recurrent dynamics (J. X. Wang et al., 2018). In this model, learning can occur through changes in the dynamical state of neural activity without synaptic plasticity mediated by dopamine. The prefrontal cortex may play a

particularly important role in model-based reinforcement learning, which allows for flexible planning according to a mental representation of the task, and can be limited by working memory processes (Babayan et al., 2018; Collins & Frank, 2012; Daw et al., 2011; Deserno et al., 2015; Dolan & Dayan, 2013; Doll et al., 2016; Gläscher et al., 2010; Russek et al., 2017; Sambrook et al., 2018; Shahar et al., 2019; Sharp et al., 2016; Wunderlich et al., 2012). This is in contrast to model-free learning, characterized by reinforcement of simple stimulus-response associations that facilitate habitual, reflexive responding. It remains to be seen how these different forms of learning relate to reward-based motor adaptation.

1.2.3 Savings and anterograde interference

In visuomotor rotation (VMR) and force field (FF) learning tasks, adaptation to a perturbation quickly washes out when the perturbation is removed. However, adaptation occurs more quickly during re-exposure to a previously encountered perturbation after loss of initial learning (Kojima et al., 2004). This phenomenon, known as savings, indicates some persistent memory for motor adaptation. A memory of adaptation to a perturbation can also cause anterograde interference during subsequent adaptation to an opposite perturbation, resulting in decreased learning rate (Miall et al., 2004). The first part of this section characterizes accounts of savings and interference in terms of sensory error-based learning mechanisms. The second part of this section discusses findings supporting the hypothesis that a reinforcement learning process can contribute to savings and interference. This section is relevant to the work presented in chapter 4, which tests whether dopamine release mediates savings and interference.

Motor memories for sensory-error based learning:

Some models account for savings and interference in motor adaptation through dynamics occurring due to multiple adaptive processes that learn at different timescales. The influential 2-state model of motor adaptation can account for savings through interactions between a slow learning mechanism that adapts and decays at a low rate and a fast learning mechanism that adapts and decays at a high rate (M. A. Smith et al., 2006).

When a perturbation is suddenly removed following adaptation, errors initially occur in the direction opposite of the removed perturbation. These aftereffects quickly wash out which seems to indicate that adaptation is lost. However, the slow process may retain some adaptation while the fast process quickly compensates to cancel the aftereffects. Residual adaptation in the slow process can then be uncovered upon re-exposure to the initial perturbation, resulting is savings. The 2-state model can also explain anterograde interference, as adapting to a perturbation in one direction biases the slow state against subsequent adaptation to an opposite perturbation. However, the 2-state model fails to account for the fact that savings can occur after extended periods of washout that should be sufficient to de-adapt the slow process (Zarahn et al., 2008). Results from dual adaptation paradigms have been used to support another model with a single fast learning process and multiple slow learning processes (J.-Y. Lee & Schweighofer, 2009). In this model, each slow learning process corresponds to a distinct motor memory, with contextual cues mediating switching between memory states. Savings can occur in this model through retrieval of latent motor memories, even following extended washout.

In the models described previously, savings occurs through retention of adaptation itself while the rate of learning in response to errors remains constant. Other have suggested that savings occurs through meta-learning that upregulates the adaptive responses to errors caused by familiar perturbations (Coltman et al., 2019; Zarahn et al., 2008). Herzfeld et al. (2014) provided compelling evidence that the brain stores memories of sensory prediction errors and that savings occurs through increased rate of learning in response to errors that are already present in memory.

Savings through reinforcement learning:

Savings has also been well documented in the literature related to operant conditioning through reinforcement (Kehoe & Macrae, 1997). Huang et al. (2011) suggested that although adaptation to visuomotor rotations relies primarily on sensory-error based learning, savings occurs due to additional influences of a reinforcement learning mechanism. They hypothesized that when error-based learning counteracts the errors

induced by a perturbation, the resulting motor commands are reinforced due to successful performance in the task. When the same perturbation is encountered in the future, those previously reinforced actions could be expressed, resulting in savings. They showed that savings occurred between adaptation to opposite rotations when the targets were arranged such that successful adaptation to both perturbations resulted in the same arm movements. This suggests that savings can occur through simply recalling a movement that was previously reinforced. Retention of sensory error-based learning would not be expected to cause savings between rotations of opposite direction, and instead would likely cause interference. In typical experiments, opposite perturbations do not result in the same adapted movements, and instead require conflicting motor solutions. In this case, reinforcement of successful actions predicts the effects of anterograde interference that are normally observed. Savings between opposite rotations with the same movement solution is only observed when the first perturbation is introduced abruptly, not gradually (Orban de Xivry & Lefèvre, 2015). Because abrupt perturbations initially produce large errors, adaptation may cause reward prediction errors by restoring success in the task, resulting in reinforcement. Gradual perturbations do not disrupt performance and thus adaptation is not associated with a significant increase in task success.

The idea that savings occurs through reinforcement of successful actions also depends on the fact that exposure to perturbations triggers the expression of previously reinforced actions. Indeed, it has been shown that experiencing random perturbations or simply the withholding of visual stimuli that indicate task success can cause re-expression of previous force-field adaptation following wash out of learning (Pekny et al., 2011). Furthermore, decades of research in the field of operant conditioning have demonstrated similar "resurgence" of previously reinforced actions in response to withholding of reward (Epstein & Skinner, 1980). Reinforcement learning is thought to depend heavily on dopamine signaling in the basal ganglia. Loss of dopamine neurons in the basal ganglia due to Parkinson's Disease causes severe impairment in savings and anterograde

interference despite initial adaptation being intact (Bédard & Sanes, 2011; Leow et al., 2012, 2013; Marinelli et al., 2009).

The conditions that result in savings are still rather unclear. In some experiments, a memory for errors can explain the occurrence of savings without the repetition of reinforced actions (Herzfeld et al., 2014; Leow et al., 2016). However, in other cases repetition of successful actions can explain savings while a memory of errors is neither necessary nor sufficient to produce savings (Orban de Xivry & Lefèvre, 2015). It is possible that both mechanisms can contribute to savings under particular conditions.

1.2.4 Effort-based motor adaptation

The literature reviewed in this section is relevant to the work presented in chapter 5, which tests whether neural signals implicated in reward-based learning respond to physical effort as a cost that devalues reward. Effects of physical effort are first reviewed in the context of value-based decision making, and second in the context of motor control. The third part of this section discusses the neurophysiology of physical effort. The fourth part of this section discusses findings that effort does not always function as a cost, rather it can sometimes increase the value of reward. Finally, effort-based learning is reviewed.

Effort influences decision making:

Humans and other animals tend to make decisions that result in more rewarding and less physically effortful outcomes. When deciding between offers consisting of varying reward values and effort requirements, people weigh the benefits and effort costs associated with potential choices (Hartmann et al., 2013; Kennerley et al., 2008; Körding et al., 2004; Prévost et al., 2010). If the likelihood of obtaining reward is determined by effort expenditure, people tend to expend more effort when the reward stakes are higher (Hauser et al., 2017; Pessiglione et al., 2007). Choices involving effort and reward are often well described by simple mathematical models in which subjective utility is a function of reward discounted by effort, and the probability of choosing an offer is a function of its utility relative to other choices.

Effort influences motor control:

Typical studies of effort-based decision making consist of choices between actions involving isometric force production, such as squeezing a hand-grip with varying amounts of force. Factors related to effort are also considered during the planning and selection of potential movements. During tasks in which people decide between different reaching movements, people strongly prefer actions that involve lower effective mass of the arm (Cos et al., 2011, 2014; Shadmehr et al., 2016a). These findings indicate that choices are sensitive to biomechanical properties of the body such that people prefer to select movements that are stable and energetically efficient. Biomechanical determinants of effort not only influence discrete choices between actions but also how movements are executed and controlled. People tend to walk at speeds and step frequencies that are energetically efficient (Bertram & Ruina, 2001; Donelan et al., 2001; Holt et al., 1991; Molen et al., 1972; Ralston, 1958; Selinger et al., 2015; Umberger & Martin, 2007; Zarrugh et al., 1974). Modelling work has attempted to identify whether the nervous system plans and executes reaching movements to minimize various effort-related cost functions. Some proposed cost functions are related muscle effort, such as the mechanical work performed by muscle torques or metabolic energy expenditure by muscles (R. M. Alexander, 1997; Berret et al., 2011; Taniai & Nishii, 2015). Other cost functions have been proposed that are related to control effort, such as the integral of squared motor neuron activation or the amount of information that must be processed by the nervous system to implement a particular control policy (Dounskaia & Shimansky, 2016; Guigon et al., 2007; Rigoux & Guigon, 2012; Todorov, 2004; Wochner et al., 2020). Although an effort cost related to control signals may correlate with energetic effort for similar types of movement, control effort can better account for some movement decisions (Morel et al., 2017).

Selecting the speed of goal-directed movement may in itself constitute an economic decision. Shadmehr et al. (2016) propose that movements incur an effort cost directly related to metabolic energy expenditure. Energy costs increase as a function of speed for both walking and reaching, suggesting that people should prefer to move more slowly

(Ralston, 1958; Shadmehr et al., 2016b). However, moving more slowly delays reward attainment for goal directed action, and it is well established in decision making research that the subjective utility of reward is discounted hyperbolically by temporal delay (Jimura et al., 2009; Kobayashi & Schultz, 2008; Prévost et al., 2010). According to the formulation of utility proposed by Shadmehr et al. (2016), a particular movement speed optimizes utility depending on the reward value of the outcome and the mechanics of the movement. The optimal speed increases for larger rewards, which accounts for findings that reward increases movement vigor (Opris et al., 2011; Summerside et al., 2018). This model can also account for decreased speed when movements must transport additional mass or when a delay between movements is imposed (Gordon et al., 1994; Haith et al., 2012). However, reward also reliably increases the vigor of movements with negligible energetic cost such as saccades. This effect that may be better accounted for by control effort than energetic effort (Kawagoe et al., 1998; Reppert et al., 2015; Sedaghat-Nejad et al., 2019; Yoon et al., 2020).

Neurophysiology of effort:

Fronto-striatal circuits implicated in reward-based learning and decision making are also involved in processing information about physical effort. Dopamine signaling plays an important role in motivating decisions to expend effort in order to obtain rewards (Denk et al., 2005; Hosking et al., 2015; Phillips et al., 2007; Salamone et al., 1994). It has been proposed that tonic dopamine release encodes the recent history of reward and invigorates action when the reward rate is high (Kurniawan et al., 2011; Niv et al., 2007). Invigoration of movement by tonic dopamine release may occur through excitatory effects on the basal ganglia direct pathway mediated by D1 receptors and inhibitory effects on the indirect pathway mediated by D2 receptors. It can be adaptive to expend more effort to move quickly in response to an increase in the rate of reward attainment because the opportunity cost of moving slowly is related to the abundance of reward in an environment.

Neural activity in the striatum has been shown to encode information about effort and to integrate effort and reward during decision making (Croxson et al., 2009; Kurniawan et al., 2010; Suzuki et al., 2020). In the prefrontal cortex, the anterior cingulate region has been consistently shown to represent effort, reward, and utility (Croxson et al., 2009; Hauser et al., 2017; Kennerley et al., 2008; Klein-Flügge et al., 2016; Porter et al., 2019; Prévost et al., 2010; Skvortsova et al., 2014; Walton et al., 2003). Localized disruptions of dopamine signaling in both the striatum and the anterior cingulate impair the motivation to produce effortful behavior (Filla et al., 2018; Salamone et al., 2007; Schweimer et al., 2005; Schweimer & Hauber, 2006).

Effort can increase the value of reward:

Economic theories assert that effort is a cost that devalues reward. Paradoxically, it has been found in humans and animals that effort can enhance the reinforcing quality of rewards (Clement et al., 2000; Inzlicht et al., 2018; Lydall et al., 2010; Zentall, 2010). It may be that prospective effort devalues reward, while retrospective effort amplifies reinforcement. For example, when given a choice between responses requiring high and low effort, participants choose to produce less effort in the immediate future. However, when given a choice between conditioned reinforcers that previously followed either low or high effort, humans and other animals tend to prefer the reinforcer that required greater effort in the past (Alessandri et al., 2008; Clement et al., 2000; Hernandez Lallement et al., 2014). Reward prediction error signals recorded from dopamine neurons have been shown to be amplified for rewards that require more effort to obtain (S. Tanaka et al., 2019). These findings may be explained by contrast effects, whereby rewards are valued relative to the immediate context. In the context of a difficult effort, a reward may be considered more valuable.

Learning to reduce effort:

Although many studies have characterized preferences and choices related to effort, relatively few have examined behavioral adaptation to changing and uncertain effort contingencies. A few experiments have been reported in which participants must learn

arbitrary and changing stimulus-response associations involving both reward and effort (Hauser et al., 2017; Skvortsova et al., 2014). These studies found that behavioral learning for effort and reward contingencies are both consistent with prediction-error mediated learning mechanisms, but that the striatum and ventromedial prefrontal cortex are involved in reward-based learning while the anterior cingulate cortex is involved in effort-based learning. Some findings also suggest that motor control is adaptive to reduce effort expenditure in response to perturbations. When a robotic exoskeleton is used to change the energetically optimal gait pattern, participants can adapt their step frequency to become more efficient (Selinger et al., 2015, 2019). Typically, movements of the elbow joint create torques at the shoulder due to intersegmental dynamics, and shoulder muscle activity predictively compensates for these interaction torques. When the shoulder joint is locked in place, this shoulder muscle activity gradually attenuates (Maeda et al., 2018, 2020). Because the shoulder joint is locked, shoulder muscle activity does not contribute to error, and instead this adaptation may reflect an optimization of effort cost.

1.2.5 Event related potentials elicited by feedback

In the work presented in chapters 2, 3, and 5, Electroencephalography (EEG) was used to measure neural event-related potentials (ERPs) elicited by feedback conveying information about sensory error, reward, and effort. This section reviews the topic of ERP correlates of feedback processing related to behavioral adaptation. The first part of this section characterizes an ERP component called the feedback related negativity/ reward positivity (FRN/RP), a medial-frontal signal implicated in error monitoring and reward-based learning. Second, the neurophysiological basis of the FRN/RP is discussed in detail. The final part of this section reviews the P300, a centro-parietal ERP component elicited by nearly all forms of task-relevant feedback that is also implicated in behavioral adaptation.

Feedback related negativity / reward positivity:

Upon commission of errors such as unwarranted responses in speeded reaction time tasks, a negative voltage deflection occurs at the medial-frontal scalp in event-related

potentials (ERPs) measured by EEG (Falkenstein et al., 1991; Gehring et al., 1993). This deflection, called the error related negativity (ERN), peaks 50-100ms after erroneous responses and is thought to reflect a process in the anterior cingulate involved in detecting errors and adapting behaviors (Holroyd & Coles, 2002; Ullsperger et al., 2014).

When success or reward is uncertain, extrinsic feedback that signals an unfavorable outcome also elicits a negative ERP deflection compared to feedback signaling a rewarding outcome (Holroyd & Coles, 2002; Miltner et al., 1997). Deemed the feedback-related negativity (FRN), this ERP component occurs 200-350ms after feedback stimuli with a medial-frontal scalp distribution nearly identical to the ERN. The FRN was originally interpreted as an ERP negativity reflecting error processing in response to unfavorable external feedback. However, converging lines of evidence have demonstrated that a large portion of the variance attributed to the FRN is actually due to a positive ERP deflection in response to rewards (Becker et al., 2014; Carlson et al., 2011; Foti et al., 2011; Krigolson, 2018; Proudfit, 2015; Sambrook & Goslin, 2016; Walsh & Anderson, 2012). The negativity observed following unfavorable outcomes may be a feature of the ERP waveform elicited by any task-relevant feedback that is cancelled by an overlapping positivity in response to reward (Baker & Holroyd, 2011; Holroyd et al., 2006, 2008). In light of these findings, the FRN is often referred to as the reward positivity (RP) in recent literature (Proudfit, 2015).

According to highly influential theories of the feedback related negativity/ reward positivity (FRN/RP), it reflects reward prediction error signals in the medial frontal cortex that affect behavioral adaptation according to the principles of reinforcement learning (Holroyd & Coles, 2002; Holroyd & Umemoto, 2016; Walsh & Anderson, 2012). Multiple studies, including a recent meta-analysis of 55 datasets, have demonstrated that the FRN/RP satisfies formal requirements of a signal encoding reward prediction error (Luque et al., 2012; Sambrook & Goslin, 2015, 2016; Talmi et al., 2012). Like reward-prediction error signals observed in midbrain dopamine neurons, the FRN/RP initially responds to reinforcement outcomes themselves but shifts to stimuli that predict those outcomes through associative learning (Eppinger et al., 2008; Holroyd et al., 2011; Liao et

al., 2011; Nieuwenhuis et al., 2002; Schultz, 2007; Walsh & Anderson, 2011). While some studies have observed trial-by-trial relationships between the FRN and learning behavior (Arbel et al., 2013; Cohen & Ranganath, 2007; Frank et al., 2005; Hewig et al., 2011; Holroyd & Krigolson, 2007; Lohse et al., 2020; Philiastides et al., 2010; Santesso et al., 2008, 2008; van der Helden et al., 2010; Yasuda et al., 2004), others studies have found that the FRN/RP does not predict behavioral adaptation (H. W. Chase et al., 2011; Martín et al., 2013; Yeung & Sanfey, 2004). The relationship between the FRN/RP and behavior may depend on the degree to which different neural control systems affect behavior in a particular task. For example, the FRN/RP may express "model-free" reward prediction errors that incrementally update habitual response tendencies even while deliberative planning or rule-based decision making processes control behavior (H. W. Chase et al., 2011; Sambrook et al., 2018).

Neurophysiological significance of the FRN/RP:

Convergent evidence from studies using source localization, fMRI, invasive recordings, and lesion studies suggest that the ERN is primarily generated in the anterior cingulate cortex (ACC; Emeric et al., 2008; Godlove et al., 2011; Keil et al., 2010; Stemmer et al., 2004; Swick & Turken, 2002; M. Ullsperger & von Cramon, 2001). The FRN/RP is generally thought to be functionally equivalent to the ERN, and has also been consistently localized to the ACC through various methods (Becker et al., 2014; Gehring & Willoughby, 2002; Hauser, Iannaccone, Ball, et al., 2014; Hauser, Iannaccone, Stämpfli, et al., 2014; Santesso et al., 2008; E. H. Smith et al., 2015; Warren et al., 2015). Recordings from single neurons in the ACC of non-human primates have revealed rich information about actions, reward prediction error, and behavioral changes in response to feedback (Amiez et al., 2005; Gemba et al., 1986; S. Ito et al., 2003; Kennerley et al., 2011; Matsumoto et al., 2007; Niki & Watanabe, 1979; Quilodran et al., 2008; Sallet et al., 2007; Schall et al., 2002; Shima & Tanji, 1998; Ullsperger et al., 2014; Williams et al., 2004). A highly influential theory states that the FRN/RP reflects reward prediction error signals in the ACC driven by dopaminergic inputs to affect adaptive behavior (Holroyd & Coles, 2002; Holroyd & Umemoto, 2016; Walsh & Anderson, 2012). Midbrain dopamine

neurons could influence the ACC directly through dense innervation of the frontal cortex. Alternatively, dopaminergic reward prediction error signals could drive the FRN/RP through projections to the striatum which is interconnected with the ACC through cortico-basal ganglia-thalamic loop circuits. Single-trial measures of the FRN/RP correlate with positive BOLD responses to reward in the ventral striatum and ACC, suggesting a function role of concerted frontostriatal circuits in generating the FRN/RP (Becker et al., 2014). Although phasic dopamine signals have not been conclusively shown to cause the generation of the FRN/RP, numerous findings suggest a role of dopamine. The magnitude of the FRN/RP is diminished in Parkinson's disease (D. R. Brown et al., 2020; Martínez-Horta et al., 2014), correlated with gray matter volume of dopaminergic midbrain structures (Carlson et al., 2015), and related to dopaminergic genetic polymorphisms (Enge et al., 2017; Marco-Pallarés et al., 2009). Some effects of pharmacological dopamine manipulations on the FRN/RP have been reported, but results have been mixed (Forster et al., 2017; Mueller et al., 2014; Santesso et al., 2009; Schutte et al., 2020).

P300:

Extrinsic feedback related to reward, error, or task performance also elicits a centro-parietal ERP component with a positive voltage peak at latencies of 300-600ms called the P300 (Glazer et al., 2018; San Martín, 2012). For both wins and losses, the amplitude of the P300 is positively related to outcome magnitude and inversely related to outcome probability (Bellebaum et al., 2010; Gu, Lei, et al., 2011; Hajcak et al., 2007; Pfabigan et al., 2011; Sato et al., 2005; Watts et al., 2017; Xu et al., 2011; Yeung & Sanfey, 2004). Effects of outcome valence on P300 magnitude are mixed and may depend on the experimental context (Bellebaum et al., 2010; Campbell et al., 1979; Frank et al., 2005; Gu, Lei, et al., 2011; Gu, Wu, et al., 2011; Kreussel et al., 2012; Polezzi et al., 2010; Sato et al., 2005; Yeung et al., 2005; Yeung & Sanfey, 2004; Zhou et al., 2010). P300 components elicited by feedback are also often modulated by high-level factors such as motivational salience, temporal waiting cost and cognitive effort (Glazer et al., 2018; Ma et al., 2014; J. Wang et al., 2014).

A longstanding and highly influential hypothesis suggests that the P300 is a measure of neural activity that functions to revise internal models of behavioral tasks in response to relevant stimuli (E. Donchin & Coles, 1988). This is consistent with findings that magnitude, surprise, and salience increase the P300 response to task feedback. Furthermore, P300 amplitude has been linked to behavioral adjustment in response to feedback (H. W. Chase et al., 2011; A. G. Fischer & Ullsperger, 2013; Sun & Wang, 2020; Yeung & Sanfey, 2004).

Widespread cortical association networks have been implicated in producing the P300, including parietal, temporal, and prefrontal areas (Polich, 2007; Soltani & Knight, 2000). The P300 is thought to reflect phasic norepinephrine release to diffuse cortical targets in response to salient stimuli to facilitate sensory encoding and adaptive behavioral adjustments (Dayan & Yu, 2006; Nieuwenhuis et al., 2005; San Martín, 2012; Vazey et al., 2018).

1.3 References

Alessandri, J., Darcheville, J.-C., Delevoye-Turrell, Y., & Zentall, T. R. (2008). Preference for rewards that follow greater effort and greater delay. *Learning & Behavior*, *36*(4), 352–358. https://doi.org/10.3758/LB.36.4.352

Alexander, G. E., & Crutcher, M. D. (1990). Functional architecture of basal ganglia circuits: Neural substrates of parallel processing. *Trends in Neurosciences*, *13*(7), 266–271. https://doi.org/10.1016/0166-2236(90)90107-l

Alexander, G. E., DeLong, M. R., & Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Review of Neuroscience*, *9*, 357–381. https://doi.org/10.1146/annurev.ne.09.030186.002041

Alexander, R. M. (1997). A minimum energy cost hypothesis for human arm trajectories. *Biological Cybernetics*, *76*(2), 97–105. https://doi.org/10.1007/s004220050324

Amiez, C., Joseph, J.-P., & Procyk, E. (2005). Anterior cingulate error-related activity is modulated by predicted reward. *The European Journal of Neuroscience*, *21*(12), 3447–3452. https://doi.org/10.1111/j.1460-9568.2005.04170.x

An, J., Yadav, T., Hessburg, J. P., & Francis, J. T. (2019). Reward Expectation Modulates Local Field Potentials, Spiking Activity and Spike-Field Coherence in the Primary Motor Cortex. *ENeuro*, *6*(3). https://doi.org/10.1523/ENEURO.0178-19.2019

Andalman, A. S., & Fee, M. S. (2009). A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(30), 12518–12523. https://doi.org/10.1073/pnas.0903214106

Arbel, Y., Goforth, K., & Donchin, E. (2013). The Good, the Bad, or the Useful? The Examination of the Relationship between the Feedback-related Negativity (FRN) and Long-term Learning Outcomes. *Journal of Cognitive Neuroscience*, *25*(8), 1249–1260. https://doi.org/10.1162/jocn_a_00385

Babayan, B. M., Uchida, N., & Gershman, S. J. (2018). Belief state representation in the dopamine system. *Nature Communications*, *9*(1), 1891. https://doi.org/10.1038/s41467-018-04397-0

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

Bariselli, S., Fobbs, W. C., Creed, M. C., & Kravitz, A. V. (2019). A competitive model for striatal action selection. *Brain Research*, *1713*, 70–79. https://doi.org/10.1016/j.brainres.2018.10.009

Barraclough, D. J., Conroy, M. L., & Lee, D. (2004). Prefrontal cortex and decision making in a mixed-strategy game. *Nature Neuroscience*, *7*(4), 404–410. https://doi.org/10.1038/nn1209

Bastian, A. J., Martin, T. A., Keating, J. G., & Thach, W. T. (1996). Cerebellar ataxia: Abnormal control of interaction torques across multiple joints. *Journal of Neurophysiology*, *76*(1), 492–509. https://doi.org/10.1152/jn.1996.76.1.492

Bayer, H. M., & Glimcher, P. W. (2005). Midbrain Dopamine Neurons Encode a Quantitative Reward Prediction Error Signal. *Neuron*, *47*(1), 129–141. https://doi.org/10.1016/j.neuron.2005.05.020

Becker, M. P. I., Nitsch, A. M., Miltner, W. H. R., & Straube, T. (2014). A single-trial estimation of the feedback-related negativity and its relation to BOLD responses in a time-estimation task. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *34*(8), 3005–3012. https://doi.org/10.1523/JNEUROSCI.3684-13.2014

Bédard, P., & Sanes, J. N. (2011). Basal ganglia-dependent processes in recalling learned visual-motor adaptations. *Experimental Brain Research*, *209*(3), 385–393. https://doi.org/10.1007/s00221-011-2561-y

Beeler, J. A., Cao, Z. F. H., Kheirbek, M. A., Ding, Y., Koranda, J., Murakami, M., Kang, U. J., & Zhuang, X. (2010). Dopamine-dependent motor learning: Insight into levodopa's long-duration response. *Annals of Neurology*, *67*(5), 639–647. https://doi.org/10.1002/ana.21947

Bellebaum, C., Polezzi, D., & Daum, I. (2010). It is less than you expected: The feedback-related negativity reflects violations of reward magnitude expectations. *Neuropsychologia*, *48*(11), 3343–3350. https://doi.org/10.1016/j.neuropsychologia.2010.07.023

Bernardi, N. F., Darainy, M., & Ostry, D. J. (2015). Somatosensory Contribution to the Initial Stages of Human Motor Learning. *Journal of Neuroscience*, *35*(42), 14316–14326. https://doi.org/10.1523/JNEUROSCI.1344-15.2015

Berniker, M., & Kording, K. P. (2011). Estimating the Relevance of World Disturbances to Explain Savings, Interference and Long-Term Motor Adaptation Effects. *PLOS Computational Biology*, *7*(10), e1002210. https://doi.org/10.1371/journal.pcbi.1002210

Berret, B., Chiovetto, E., Nori, F., & Pozzo, T. (2011). Evidence for Composite Cost Functions in Arm Movement Planning: An Inverse Optimal Control Approach. *PLOS Computational Biology*, *7*(10), e1002183. https://doi.org/10.1371/journal.pcbi.1002183

Bertram, J. E., & Ruina, A. (2001). Multiple walking speed-frequency relations are predicted by constrained optimization. *Journal of Theoretical Biology*, *209*(4), 445–453. https://doi.org/10.1006/jtbi.2001.2279

Bhanpuri, N. H., Okamura, A. M., & Bastian, A. J. (2011). Active force perception depends on cerebellar function. *Journal of Neurophysiology*, *107*(6), 1612–1620. https://doi.org/10.1152/jn.00983.2011

Bhanpuri, N. H., Okamura, A. M., & Bastian, A. J. (2013). Predictive Modeling by the Cerebellum Improves Proprioception. *Journal of Neuroscience*, *33*(36), 14301–14306. https://doi.org/10.1523/JNEUROSCI.0784-13.2013

Bhanpuri, N. H., Okamura, A. M., & Bastian, A. J. (2014). Predicting and correcting ataxia using a model of cerebellar function. *Brain*, *137*(7), 1931–1944. https://doi.org/10.1093/brain/awu115

Bhushan, N., & Shadmehr, R. (1999). Computational nature of human adaptive control during learning of reaching movements in force fields. *Biological Cybernetics*, *81*(1), 39–60. https://doi.org/10.1007/s004220050543

Blakemore, S.-J., Wolpert, D. M., & Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nature Neuroscience*, *1*(7), 635–640. https://doi.org/10.1038/2870

Block, H. J., & Bastian, A. J. (2012). Cerebellar involvement in motor but not sensory adaptation. *Neuropsychologia*, *50*(8), 1766–1775. https://doi.org/10.1016/j.neuropsychologia.2012.03.034

Bottjer, S. W., Miesner, E. A., & Arnold, A. P. (1984). Forebrain lesions disrupt development but not maintenance of song in passerine birds. *Science (New York, N.Y.)*, *224*(4651), 901–903. https://doi.org/10.1126/science.6719123

Brooks, J. X., Carriot, J., & Cullen, K. E. (2015). Learning to expect the unexpected: Rapid updating in primate cerebellum during voluntary self-motion. *Nature Neuroscience*, *18*(9), 1310–1317. https://doi.org/10.1038/nn.4077

Brooks, J. X., & Cullen, K. E. (2013). The Primate Cerebellum Selectively Encodes Unexpected Self-Motion. *Current Biology*, *23*(11), 947–955. https://doi.org/10.1016/j.cub.2013.04.029

Brooks, J. X., & Cullen, K. E. (2019). Predictive Sensing: The Role of Motor Signals in Sensory Processing. *Biological Psychiatry. Cognitive Neuroscience and Neuroimaging*, *4*(9), 842–850. https://doi.org/10.1016/j.bpsc.2019.06.003

Brown, D. R., Richardson, S. P., & Cavanagh, J. F. (2020). An EEG marker of reward processing is diminished in Parkinson's disease. *Brain Research*, *1727*, 146541. https://doi.org/10.1016/j.brainres.2019.146541

Brown, L. E., Rosenbaum, D. A., & Sainburg, R. L. (2003). Limb Position Drift: Implications for Control of Posture and Movement. *Journal of Neurophysiology*, *90*(5), 3105–3118. https://doi.org/10.1152/jn.00013.2003

Calabresi, P., Picconi, B., Tozzi, A., & Di Filippo, M. (2007). Dopamine-mediated regulation of corticostriatal synaptic plasticity. *Trends in Neurosciences*, *30*(5), 211–219. https://doi.org/10.1016/j.tins.2007.03.001

Campbell, K. B., Courchesne, E., Picton, T. W., & Squires, K. C. (1979). Evoked potential correlates of human information processing. *Biological Psychology*, *8*(1), 45–68. https://doi.org/10.1016/0301-0511(79)90004-8

Carlson, J. M., Foti, D., Harmon-Jones, E., & Proudfit, G. H. (2015). Midbrain volume predicts fMRI and ERP measures of reward reactivity. *Brain Structure and Function*, *220*(3), 1861–1866. https://doi.org/10.1007/s00429-014-0725-9

Carlson, J. M., Foti, D., Mujica-Parodi, L. R., Harmon-Jones, E., & Hajcak, G. (2011). Ventral striatal and medial prefrontal BOLD activation is correlated with reward-related electrocortical activity: A combined ERP and fMRI study. *NeuroImage*, *57*(4), 1608–1616. https://doi.org/10.1016/j.neuroimage.2011.05.037

Cashaback, J. G. A., Lao, C. K., Palidis, D. J., Coltman, S. K., McGregor, H. R., & Gribble, P. L. (2019). The gradient of the reinforcement landscape influences sensorimotor learning. *PLOS Computational Biology*, *15*(3), e1006839. https://doi.org/10.1371/journal.pcbi.1006839

Cashaback, J. G. A., McGregor, H. R., Mohatarem, A., & Gribble, P. L. (2017). Dissociating error-based and reinforcement-based loss functions during sensorimotor learning. *PLoS Computational Biology*, *13*(7), e1005623. https://doi.org/10.1371/journal.pcbi.1005623

Chapman, C. E. (1994). Active versus passive touch: Factors influencing the transmission of somatosensory signals to primary somatosensory cortex. *Canadian Journal of Physiology and Pharmacology*, *72*(5), 558–570. https://doi.org/10.1139/y94-080

Chase, H. W., Swainson, R., Durham, L., Benham, L., & Cools, R. (2011). Feedback-related Negativity Codes Prediction Error but Not Behavioral Adjustment during Probabilistic Reversal Learning. *Journal of Cognitive Neuroscience*, *23*(4), 936–946. https://doi.org/10.1162/jocn.2010.21456

Chase, S. M., Kass, R. E., & Schwartz, A. B. (2012). Behavioral and neural correlates of visuomotor adaptation observed through a brain-computer interface in primary motor cortex. *Journal of Neurophysiology*, *108*(2), 624–644. https://doi.org/10.1152/jn.00371.2011

Clement, T. S., Feltus, J. R., Kaiser, D. H., & Zentall, T. R. (2000). "work ethic" in pigeons: Reward value is directly related to the effort or time required to obtain the reward. *Psychonomic Bulletin & Review*, *7*(1), 100–106. https://doi.org/10.3758/BF03210727

Clower, D. M., Hoffman, J. M., Votaw, J. R., Faber, T. L., Woods, R. P., & Alexander, G. E. (1996). Role of posterior parietal cortex in the recalibration of visually guided reaching. *Nature*, *383*(6601), 618–621. https://doi.org/10.1038/383618a0

Codol, O., Holland, P. J., & Galea, J. M. (2018). The relationship between reinforcement and explicit control during visuomotor adaptation. *Scientific Reports*, *8*(1), 9121. https://doi.org/10.1038/s41598-018-27378-1

Cohen, M. X., & Ranganath, C. (2007). Reinforcement Learning Signals Predict Future Decisions. *Journal of Neuroscience*, *27*(2), 371–378. https://doi.org/10.1523/JNEUROSCI.4421-06.2007

Collins, A. G. E., & Frank, M. J. (2012). How much of reinforcement learning is working memory, not reinforcement learning? A behavioral, computational, and neurogenetic analysis. *The European Journal of Neuroscience*, *35*(7), 1024–1035. https://doi.org/10.1111/j.1460-9568.2011.07980.x

Coltman, S. K., Cashaback, J. G. A., & Gribble, P. L. (2019). Both fast and slow learning processes contribute to savings following sensorimotor adaptation. *Journal of Neurophysiology*, *121*(4), 1575–1583. https://doi.org/10.1152/jn.00794.2018

Confais, J., Kim, G., Tomatsu, S., Takei, T., & Seki, K. (2017). Nerve-Specific Input Modulation to Spinal Neurons during a Motor Task in the Monkey. *Journal of Neuroscience*, *37*(10), 2612–2626. https://doi.org/10.1523/JNEUROSCI.2561-16.2017

Cools, R., Lewis, S. J. G., Clark, L., Barker, R. A., & Robbins, T. W. (2007). L-DOPA disrupts activity in the nucleus accumbens during reversal learning in Parkinson's disease. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, *32*(1), 180–189. https://doi.org/10.1038/sj.npp.1301153

Cos, I., Bélanger, N., & Cisek, P. (2011). The influence of predicted arm biomechanics on decision making. *Journal of Neurophysiology*, *105*(6), 3022–3033. https://doi.org/10.1152/jn.00975.2010

Cos, I., Duque, J., & Cisek, P. (2014). Rapid prediction of biomechanical costs during action decisions. *Journal of Neurophysiology*, *112*(6), 1256–1266. https://doi.org/10.1152/jn.00147.2014

Crevecoeur, F., Mathew, J., Bastin, M., & Lefèvre, P. (2020). Feedback Adaptation to Unpredictable Force Fields in 250 ms. *ENeuro*, *7*(2). https://doi.org/10.1523/ENEURO.0400-19.2020

Croxson, P. L., Walton, M. E., O'Reilly, J. X., Behrens, T. E. J., & Rushworth, M. F. S. (2009). Effort-Based Cost–Benefit Valuation and the Human Brain. *Journal of Neuroscience*, *29*(14), 4531–4541. https://doi.org/10.1523/JNEUROSCI.4515-08.2009

Cui, G., Jun, S. B., Jin, X., Pham, M. D., Vogel, S. S., Lovinger, D. M., & Costa, R. M. (2013). Concurrent activation of striatal direct and indirect pathways during action initiation. *Nature*, *494*(7436), 238–242. https://doi.org/10.1038/nature11846

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, *69*(6), 1204–1215. https://doi.org/10.1016/j.neuron.2011.02.027

Dayan, P., & Yu, A. J. (2006). Phasic norepinephrine: A neural interrupt signal for unexpected events. *Network: Computation in Neural Systems*, *17*(4), 335–350. https://doi.org/10.1080/09548980601004024

Delgado, M. R., Stenger, V. A., & Fiez, J. A. (2004). Motivation-dependent Responses in the Human Caudate Nucleus. *Cerebral Cortex*, *14*(9), 1022–1030. https://doi.org/10.1093/cercor/bhh062

Denk, F., Walton, M. E., Jennings, K. A., Sharp, T., Rushworth, M. F. S., & Bannerman, D. M. (2005). Differential involvement of serotonin and dopamine systems in cost-benefit decisions about delay or effort. *Psychopharmacology*, *179*(3), 587–596. https://doi.org/10.1007/s00213-004-2059-4

Deserno, L., Huys, Q. J. M., Boehme, R., Buchert, R., Heinze, H.-J., Grace, A. A., Dolan, R. J., Heinz, A., & Schlagenhauf, F. (2015). Ventral striatal dopamine reflects behavioral and neural signatures of model-based control during sequential decision making. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(5), 1595–1600. https://doi.org/10.1073/pnas.1417219112

Dhawale, A. K., Miyamoto, Y. R., Smith, M. A., & Ölveczky, B. P. (2019). Adaptive Regulation of Motor Variability. *Current Biology*, *29*(21), 3551-3562.e7. https://doi.org/10.1016/j.cub.2019.08.052

Diederen, K. M. J., Ziauddeen, H., Vestergaard, M. D., Spencer, T., Schultz, W., & Fletcher, P. C. (2017). Dopamine Modulates Adaptive Prediction Error Coding in the Human Midbrain and Striatum. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *37*(7), 1708–1720. https://doi.org/10.1523/JNEUROSCI.1979-16.2016

Diedrichsen, J., Hashambhoy, Y., Rane, T., & Shadmehr, R. (2005). Neural Correlates of Reach Errors. *Journal of Neuroscience*, *25*(43), 9919–9931. https://doi.org/10.1523/JNEUROSCI.1874-05.2005

Diedrichsen, J., Verstynen, T., Hon, A., Lehman, S. L., & Ivry, R. B. (2003). Anticipatory adjustments in the unloading task: Is an efference copy necessary for learning? *Experimental Brain Research*, *148*(2), 272–276. https://doi.org/10.1007/s00221-002-1318-z

Dolan, R. J., & Dayan, P. (2013). Goals and Habits in the Brain. *Neuron*, *80*(2), 312–325. https://doi.org/10.1016/j.neuron.2013.09.007

Doll, B. B., Bath, K. G., Daw, N. D., & Frank, M. J. (2016). Variability in Dopamine Genes Dissociates Model-Based and Model-Free Reinforcement Learning. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *36*(4), 1211–1222. https://doi.org/10.1523/JNEUROSCI.1901-15.2016

Donchin, E., & Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences*, *11*(3), 357–374. https://doi.org/10.1017/S0140525X00058027

Donchin, O., Francis, J. T., & Shadmehr, R. (2003). Quantifying Generalization from Trial-by-Trial Behavior of Adaptive Systems that Learn with Basis Functions: Theory and Experiments in Human Motor Control. *Journal of Neuroscience*, *23*(27), 9032–9045. https://doi.org/10.1523/JNEUROSCI.23-27-09032.2003

Donelan, J. M., Kram, R., & Kuo, A. D. (2001). Mechanical and metabolic determinants of the preferred step width in human walking. *Proceedings. Biological Sciences*, *268*(1480), 1985–1992. https://doi.org/10.1098/rspb.2001.1761

Dounskaia, N., & Shimansky, Y. (2016). Strategy of arm movement control is determined by minimization of neural effort for joint coordination. *Experimental Brain Research*, *234*(6), 1335–1350. https://doi.org/10.1007/s00221-016-4610-z

Dunovan, K., & Verstynen, T. (2016). Believer-Skeptic Meets Actor-Critic: Rethinking the Role of Basal Ganglia Pathways during Decision-Making and Reinforcement Learning. *Frontiers in Neuroscience*, *10*. https://doi.org/10.3389/fnins.2016.00106

Dunovan, K., Vich, C., Clapp, M., Verstynen, T., & Rubin, J. (2019). Reward-driven changes in striatal pathway competition shape evidence evaluation in decision-making. *PLOS Computational Biology*, *15*(5), e1006998. https://doi.org/10.1371/journal.pcbi.1006998

Ebner, T. J., & Pasalar, S. (2008). Cerebellum Predicts the Future Motor State. *The Cerebellum*, *7*(4), 583–588. https://doi.org/10.1007/s12311-008-0059-3

Emeric, E. E., Brown, J. W., Leslie, M., Pouget, P., Stuphorn, V., & Schall, J. D. (2008). Performance monitoring local field potentials in the medial frontal cortex of primates: Anterior cingulate cortex. *Journal of Neurophysiology*, *99*(2), 759–772. https://doi.org/10.1152/jn.00896.2006

Enge, S., Mothes, H., Fleischhauer, M., Reif, A., & Strobel, A. (2017). Genetic variation of dopamine and serotonin function modulates the feedback-related negativity during altruistic punishment. *Scientific Reports*, *7*(1), 2996. https://doi.org/10.1038/s41598-017-02594-3

Eppinger, B., Kray, J., Mock, B., & Mecklinger, A. (2008). Better or worse than expected? Aging, learning, and the ERN. *Neuropsychologia*, *46*(2), 521–539. https://doi.org/10.1016/j.neuropsychologia.2007.09.001

Epstein, R., & Skinner, B. F. (1980). Resurgence of responding after the cessation of response-independent reinforcement. *Proceedings of the National Academy of Sciences of the United States of America*, *77*(10), 6251–6253. https://doi.org/10.1073/pnas.77.10.6251

Falkenstein, M., Hohnsbein, J., Hoormann, J., & Blanke, L. (1991). Effects of crossmodal divided attention on late ERP components. II. Error processing in choice reaction tasks. *Electroencephalography and Clinical Neurophysiology*, *78*(6), 447–455. https://doi.org/10.1016/0013-4694(91)90062-9

Fetz, E. E., Perlmutter, S. I., Prut, Y., & Seki, K. (2002). Functional properties of primate spinal interneurones during voluntary hand movements. *Advances in Experimental Medicine and Biology*, *508*, 265–271. https://doi.org/10.1007/978-1-4615-0713-0_32

Filla, I., Bailey, M. R., Schipani, E., Winiger, V., Mezias, C., Balsam, P. D., & Simpson, E. H. (2018). Striatal dopamine D2 receptors regulate effort but not value-based decision making and alter the dopaminergic encoding of cost. *Neuropsychopharmacology*, *43*(11), 2180–2189. https://doi.org/10.1038/s41386-018-0159-9

Fischer, A. G., & Ullsperger, M. (2013). Real and Fictive Outcomes Are Processed Differently but Converge on a Common Adaptive Mechanism. *Neuron*, *79*(6), 1243–1255. https://doi.org/10.1016/j.neuron.2013.07.006

Fischer, B., & Boch, R. (1983). Saccadic eye movements after extremely short reaction times in the monkey. *Brain Research*, *260*(1), 21–26. https://doi.org/10.1016/0006-8993(83)90760-6

Flanagan, J. R., & Wing, A. M. (1993). Modulation of grip force with load force during point-to-point arm movements. *Experimental Brain Research*, *95*(1), 131–143. https://doi.org/10.1007/BF00229662

Forster, S. E., Zirnheld, P., Shekhar, A., Steinhauer, S. R., O'Donnell, B. F., & Hetrick, W. P. (2017). Event-related potentials reflect impaired temporal interval learning following haloperidol administration. *Psychopharmacology*, *234*(17), 2545–2562. https://doi.org/10.1007/s00213-017-4645-2

Foti, D., Weinberg, A., Dien, J., & Hajcak, G. (2011). Event-related potential activity in the basal ganglia differentiates rewards from nonrewards: Temporospatial principal components analysis and source localization of the feedback negativity. *Human Brain Mapping*, *32*(12), 2207–2216. https://doi.org/10.1002/hbm.21182

Frank, M. J., Woroch, B. S., & Curran, T. (2005). Error-Related Negativity Predicts Reinforcement Learning and Conflict Biases. *Neuron*, *47*(4), 495–501. https://doi.org/10.1016/j.neuron.2005.06.020

Fuentes, C. T., & Bastian, A. J. (2010). Where is your arm? Variations in proprioception across space and tasks. *Journal of Neurophysiology*, *103*(1), 164–171. https://doi.org/10.1152/jn.00494.2009

Galea, J. M., Mallia, E., Rothwell, J., & Diedrichsen, J. (2015). The dissociable effects of punishment and reward on motor learning. *Nature Neuroscience*, *18*(4), 597–602. https://doi.org/10.1038/nn.3956

Galea, J. M., Vazquez, A., Pasricha, N., Orban de Xivry, J.-J., & Celnik, P. (2011). Dissociating the Roles of the Cerebellum and Motor Cortex during Adaptive Learning: The Motor Cortex Retains What the Cerebellum Learns. *Cerebral Cortex*, *21*(8), 1761–1770. https://doi.org/10.1093/cercor/bhq246

García-García, I., Zeighami, Y., & Dagher, A. (2017). Reward Prediction Errors in Drug Addiction and Parkinson's Disease: From Neurophysiology to Neuroimaging. *Current Neurology and Neuroscience Reports*, *17*(6), 46. https://doi.org/10.1007/s11910-017-0755-9

Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., & Donchin, E. (1993). A Neural System for Error Detection and Compensation. *Psychological Science*, *4*(6), 385–390.

Gehring, W. J., & Willoughby, A. R. (2002). The Medial Frontal Cortex and the Rapid Processing of Monetary Gains and Losses. *Science*, *295*(5563), 2279–2282. https://doi.org/10.1126/science.1066893

Gemba, H., Sasaki, K., & Brooks, V. B. (1986). "Error" potentials in limbic cortex (anterior cingulate area 24) of monkeys during motor learning. *Neuroscience Letters*, *70*(2), 223–227. https://doi.org/10.1016/0304-3940(86)90467-2

Gläscher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). States versus rewards: Dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*, *66*(4), 585–595. https://doi.org/10.1016/j.neuron.2010.04.016

Glazer, J. E., Kelley, N. J., Pornpattananangkul, N., Mittal, V. A., & Nusslock, R. (2018). Beyond the FRN: Broadening the time-course of EEG and ERP components implicated in reward processing. *International Journal of Psychophysiology*, *132*, 184–202. https://doi.org/10.1016/j.ijpsycho.2018.02.002

Godlove, D. C., Emeric, E. E., Segovis, C. M., Young, M. S., Schall, J. D., & Woodman, G. F. (2011). Event-Related Potentials Elicited by Errors during the Stop-Signal Task. I. Macaque Monkeys. *The Journal of Neuroscience*, *31*(44), 15640–15649. https://doi.org/10.1523/JNEUROSCI.3349-11.2011

Gonzalez Castro, L. N., Hadjiosif, A. M., Hemphill, M. A., & Smith, M. A. (2014). Environmental Consistency Determines the Rate of Motor Adaptation. *Current Biology*, *24*(10), 1050–1061. https://doi.org/10.1016/j.cub.2014.03.049

Gordon, J., Ghilardi, M. F., Cooper, S. E., & Ghez, C. (1994). Accuracy of planar reaching movements. II. Systematic extent errors resulting from inertial anisotropy. *Experimental Brain Research*, *99*(1), 112–130. https://doi.org/10.1007/BF00241416

Gribble, P. L., & Ostry, D. J. (1999). Compensation for Interaction Torques During Single- and Multijoint Limb Movement. *Journal of Neurophysiology*, *82*(5), 2310–2326. https://doi.org/10.1152/jn.1999.82.5.2310

Gu, R., Lei, Z., Broster, L., Wu, T., Jiang, Y., & Luo, Y. (2011). Beyond valence and magnitude: A flexible evaluative coding system in the brain. *Neuropsychologia*, *49*(14), 3891–3897. https://doi.org/10.1016/j.neuropsychologia.2011.10.006

Gu, R., Wu, T., Jiang, Y., & Luo, Y.-J. (2011). Woulda, coulda, shoulda: The evaluation and the impact of the alternative outcome. *Psychophysiology*, *48*(10), 1354–1360. https://doi.org/10.1111/j.1469-8986.2011.01215.x

Guigon, E., Baraduc, P., & Desmurget, M. (2007). Computational Motor Control: Redundancy and Invariance. *Journal of Neurophysiology*, *97*(1), 331–347. https://doi.org/10.1152/jn.00290.2006

Guthrie, B. L., Porter, J. D., & Sparks, D. L. (1983). Corollary discharge provides accurate eye position information to the oculomotor system. *Science*, *221*(4616), 1193–1195. https://doi.org/10.1126/science.6612334

Hadipour-Niktarash, A., Lee, C. K., Desmond, J. E., & Shadmehr, R. (2007). Impairment of Retention But Not Acquisition of a Visuomotor Skill Through Time-Dependent Disruption of Primary Motor Cortex. *Journal of Neuroscience*, *27*(49), 13413–13419. https://doi.org/10.1523/JNEUROSCI.2570-07.2007

Haggard, P., Miall, R. C., Wade, D., Fowler, S., Richardson, A., Anslow, P., & Stein, J. (1995). Damage to cerebellocortical pathways after closed head injury: A behavioural and magnetic resonance imaging study. *Journal of Neurology, Neurosurgery & Psychiatry*, *58*(4), 433–438. https://doi.org/10.1136/jnnp.58.4.433

Haith, A. M., Reppert, T. R., & Shadmehr, R. (2012). Evidence for hyperbolic temporal discounting of reward in control of movements. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *32*(34), 11727–11736. https://doi.org/10.1523/JNEUROSCI.0424-12.2012

Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2007). It's worse than you thought: The feedback negativity and violations of reward prediction in gambling tasks. *Psychophysiology*, *44*(6), 905–912. https://doi.org/10.1111/j.1469-8986.2007.00567.x

Hartmann, M. N., Hager, O. M., Tobler, P. N., & Kaiser, S. (2013). Parabolic discounting of monetary rewards by physical effort. *Behavioural Processes*, *100*, 192–196. https://doi.org/10.1016/j.beproc.2013.09.014

Hauser, T. U., Eldar, E., & Dolan, R. J. (2017). Separate mesocortical and mesolimbic pathways encode effort and reward learning signals. *Proceedings of the National Academy of Sciences*, *114*(35), E7395–E7404. https://doi.org/10.1073/pnas.1705643114

Hauser, T. U., Iannaccone, R., Ball, J., Mathys, C., Brandeis, D., Walitza, S., & Brem, S. (2014). Role of the medial prefrontal cortex in impaired decision making in juvenile attention-deficit/hyperactivity disorder. *JAMA Psychiatry*, *71*(10), 1165–1173. https://doi.org/10.1001/jamapsychiatry.2014.1093

Hauser, T. U., Iannaccone, R., Stämpfli, P., Drechsler, R., Brandeis, D., Walitza, S., & Brem, S. (2014). The feedback-related negativity (FRN) revisited: New insights into the localization, meaning and network organization. *NeuroImage*, *84*, 159–168. https://doi.org/10.1016/j.neuroimage.2013.08.028

Henriques, D. Y. P., Filippopulos, F., Straube, A., & Eggert, T. (2014). The cerebellum is not necessary for visually driven recalibration of hand proprioception. *Neuropsychologia*, *64*, 195–204. https://doi.org/10.1016/j.neuropsychologia.2014.09.029

Hernandez Lallement, J., Kuss, K., Trautner, P., Weber, B., Falk, A., & Fliessbach, K. (2014). Effort increases sensitivity to reward and loss magnitude in the human brain. *Social Cognitive and Affective Neuroscience*, *9*(3), 342–349. https://doi.org/10.1093/scan/nss147

Herzfeld, D. J., Vaswani, P. A., Marko, M. K., & Shadmehr, R. (2014). A memory of errors in sensorimotor learning. *Science*, *345*(6202), 1349–1353. https://doi.org/10.1126/science.1253138

Hewig, J., Kretschmer, N., Trippe, R. H., Hecht, H., Coles, M. G. H., Holroyd, C. B., & Miltner, W. H. R. (2011). Why humans deviate from rational choice. *Psychophysiology*, *48*(4), 507–514. https://doi.org/10.1111/j.1469-8986.2010.01081.x

Hewitt, A. L., Popa, L. S., & Ebner, T. J. (2015). Changes in Purkinje Cell Simple Spike Encoding of Reach Kinematics during Adaption to a Mechanical Perturbation. *Journal of Neuroscience*, *35*(3), 1106–1124. https://doi.org/10.1523/JNEUROSCI.2579-14.2015

Hiebert, N. M., Owen, A. M., Ganjavi, H., Mendonça, D., Jenkins, M. E., Seergobin, K. N., & MacDonald, P. A. (2019). Dorsal striatum does not mediate feedback-based, stimulus-response learning: An event-related fMRI study in patients with Parkinson's disease tested on and off dopaminergic therapy. *NeuroImage*, *185*, 455–470. https://doi.org/10.1016/j.neuroimage.2018.10.045

Hoffmann, L. A., Saravanan, V., Wood, A. N., He, L., & Sober, S. J. (2016). Dopaminergic Contributions to Vocal Learning. *Journal of Neuroscience*, *36*(7), 2176–2189. https://doi.org/10.1523/JNEUROSCI.3883-15.2016

Holland, P., Codol, O., & Galea, J. M. (2018). Contribution of explicit processes to reinforcement-based motor learning. *Journal of Neurophysiology*, *119*(6), 2241–2255. https://doi.org/10.1152/jn.00901.2017

Holland, P., Codol, O., Oxley, E., Taylor, M., Hamshere, E., Joseph, S., Huffer, L., & Galea, J. M. (2019). Domain-Specific Working Memory, But Not Dopamine-Related Genetic Variability, Shapes Reward-Based Motor Learning. *The Journal of Neuroscience*, *39*(47), 9383–9396. https://doi.org/10.1523/JNEUROSCI.0583-19.2019

Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*(4), 679–709. https://doi.org/10.1037/0033-295X.109.4.679

Holroyd, C. B., Hajcak, G., & Larsen, J. T. (2006). The good, the bad and the neutral: Electrophysiological responses to feedback stimuli. *Brain Research*, *1105*(1), 93–101. https://doi.org/10.1016/j.brainres.2005.12.015

Holroyd, C. B., & Krigolson, O. E. (2007). Reward prediction error signals associated with a modified time estimation task. *Psychophysiology*, *44*(6), 913–917. https://doi.org/10.1111/j.1469-8986.2007.00561.x

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

Holroyd, C. B., Pakzad-Vaezi, K. L., & Krigolson, O. E. (2008). The feedback correct-related positivity: Sensitivity of the event-related brain potential to unexpected positive feedback. *Psychophysiology*, *45*(5), 688–697. https://doi.org/10.1111/j.1469-8986.2008.00668.x

Holroyd, C. B., & Umemoto, A. (2016). The research domain criteria framework: The case for anterior cingulate cortex. *Neuroscience & Biobehavioral Reviews*, *71*, 418–443. https://doi.org/10.1016/j.neubiorev.2016.09.021

Holt, K. G., Hamill, J., & Andres, R. O. (1991). Predicting the minimal energy costs of human walking. *Medicine and Science in Sports and Exercise*, *23*(4), 491–498.

Hosking, J. G., Floresco, S. B., & Winstanley, C. A. (2015). Dopamine Antagonism Decreases Willingness to Expend Physical, But Not Cognitive, Effort: A Comparison of Two Rodent Cost/Benefit Decision-Making Tasks. *Neuropsychopharmacology*, *40*(4), 1005–1015. https://doi.org/10.1038/npp.2014.285

Hosp, J. A., Pekanovic, A., Rioult-Pedotti, M. S., & Luft, A. R. (2011). Dopaminergic Projections from Midbrain to Primary Motor Cortex Mediate Motor Skill Learning. *Journal of Neuroscience*, *31*(7), 2481–2487. https://doi.org/10.1523/JNEUROSCI.5411-10.2011

Huang, V. S., Haith, A., Mazzoni, P., & Krakauer, J. W. (2011). Rethinking Motor Learning and Savings in Adaptation Paradigms: Model-Free Memory for Successful Actions Combines with Internal Models. *Neuron*, *70*(4), 787–801. https://doi.org/10.1016/j.neuron.2011.04.012

Hunt, L. T., Behrens, T. E., Hosokawa, T., Wallis, J. D., & Kennerley, S. W. (2015). Capturing the temporal evolution of choice across prefrontal cortex. *ELife*, *4*, e11945. https://doi.org/10.7554/eLife.11945

Hunt, L. T., Malalasekera, W., de Berker, A., Miranda, B., Farmer, S., Behrens, T., & Kennerley, S. (2018). Triple Dissociation of Attention and Decision Computations across Prefrontal Cortex. *Nature Neuroscience*, *21*(10), 1471–1481. https://doi.org/10.1038/s41593-018-0239-5

Inoue, K., Kawashima, R., Satoh, K., Kinomura, S., Sugiura, M., Goto, R., Ito, M., & Fukuda, H. (2000). A PET study of visuomotor learning under optical rotation. *NeuroImage*, *11*(5 Pt 1), 505–516. https://doi.org/10.1006/nimg.2000.0554

Inzlicht, M., Shenhav, A., & Olivola, C. Y. (2018). The Effort Paradox: Effort Is Both Costly and Valued. *Trends in Cognitive Sciences*, *22*(4), 337–349. https://doi.org/10.1016/j.tics.2018.01.007

Ito, M. (2006). Cerebellar circuitry as a neuronal machine. *Progress in Neurobiology*, *78*(3), 272–303. https://doi.org/10.1016/j.pneurobio.2006.02.006

Ito, S., Stuphorn, V., Brown, J. W., & Schall, J. D. (2003). Performance monitoring by the anterior cingulate cortex during saccade countermanding. *Science (New York, N.Y.)*, *302*(5642), 120–122. https://doi.org/10.1126/science.1087847

Izawa, J., Criscimagna-Hemminger, S. E., & Shadmehr, R. (2012). Cerebellar Contributions to Reach Adaptation and Learning Sensory Consequences of Action. *Journal of Neuroscience*, *32*(12), 4230–4239. https://doi.org/10.1523/JNEUROSCI.6353-11.2012

Izawa, J., & Shadmehr, R. (2008). On-Line Processing of Uncertain Information in Visuomotor Control. *Journal of Neuroscience*, *28*(44), 11360–11368. https://doi.org/10.1523/JNEUROSCI.3063-08.2008

Izawa, J., & Shadmehr, R. (2011). Learning from Sensory and Reward Prediction Errors during Motor Adaptation. *PLOS Computational Biology*, *7*(3), e1002012. https://doi.org/10.1371/journal.pcbi.1002012

Jimura, K., Myerson, J., Hilgard, J., Braver, T. S., & Green, L. (2009). Are people really more patient than other animals? Evidence from human discounting of real liquid rewards. *Psychonomic Bulletin & Review*, *16*(6). https://doi.org/10.3758/PBR.16.6.1071

Jocham, G., & Ullsperger, M. (2009). Neuropharmacology of performance monitoring. *Neuroscience and Biobehavioral Reviews*, *33*(1), 48–60. https://doi.org/10.1016/j.neubiorev.2008.08.011

Kakei, S., Lee, J., Mitoma, H., Tanaka, H., Manto, M., & Hampe, C. S. (2019). Contribution of the Cerebellum to Predictive Motor Control and Its Evaluation in Ataxic Patients. *Frontiers in Human Neuroscience*, *13*. https://doi.org/10.3389/fnhum.2019.00216

Kao, M. H., Doupe, A. J., & Brainard, M. S. (2005). Contributions of an avian basal ganglia-forebrain circuit to real-time modulation of song. *Nature*, *433*(7026), 638–643. https://doi.org/10.1038/nature03127

Kao, M. H., Wright, B. D., & Doupe, A. J. (2008). Neurons in a forebrain nucleus required for vocal plasticity rapidly switch between precise firing and variable bursting depending on social context. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *28*(49), 13232–13247. https://doi.org/10.1523/JNEUROSCI.2250-08.2008

Kawagoe, R., Takikawa, Y., & Hikosaka, O. (1998). Expectation of reward modulates cognitive signals in the basal ganglia. *Nature Neuroscience*, *1*(5), 411–416. https://doi.org/10.1038/1625

Kawato, M. (1999). Internal models for motor control and trajectory planning. *Current Opinion in Neurobiology*, *9*(6), 718–727. https://doi.org/10.1016/S0959-4388(99)00028-8

Kawato, M., Kuroda, T., Imamizu, H., Nakano, E., Miyauchi, S., & Yoshioka, T. (2003). Internal forward models in the cerebellum: FMRI study on grip force and load force coupling. In *Progress in Brain Research* (Vol. 142, pp. 171–188). Elsevier. https://doi.org/10.1016/S0079-6123(03)42013-X

Kehoe, E. J., & Macrae, M. (1997). Savings in animal learning: Implications for relapse and maintenance after therapy. *Behavior Therapy*, *28*(1), 141–155. https://doi.org/10.1016/S0005-7894(97)80039-1

Keil, J., Weisz, N., Paul-Jordanov, I., & Wienbruch, C. (2010). Localization of the magnetic equivalent of the ERN and induced oscillatory brain activity. *NeuroImage*, *51*(1), 404–411. https://doi.org/10.1016/j.neuroimage.2010.02.003

Kennerley, S. W., Behrens, T. E. J., & Wallis, J. D. (2011). Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. *Nature Neuroscience*, *14*(12), 1581–1589. https://doi.org/10.1038/nn.2961

Kennerley, S. W., Dahmubed, A. F., Lara, A. H., & Wallis, J. D. (2008). Neurons in the Frontal Lobe Encode the Value of Multiple Decision Variables. *Journal of Cognitive Neuroscience*, *21*(6), 1162–1178. https://doi.org/10.1162/jocn.2009.21100

Kim, H. E., Parvin, D. E., & Ivry, R. B. (2019). The influence of task outcome on implicit motor learning. *ELife*, *8*. https://doi.org/10.7554/eLife.39882

Klaus, A., Martins, G. J., Paixao, V. B., Zhou, P., Paninski, L., & Costa, R. M. (2017). The Spatiotemporal Organization of the Striatum Encodes Action Space. *Neuron*, *95*(5), 1171-1180.e7. https://doi.org/10.1016/j.neuron.2017.08.015

Klein-Flügge, M. C., Kennerley, S. W., Friston, K., & Bestmann, S. (2016). Neural Signatures of Value Comparison in Human Cingulate Cortex during Decisions Requiring an Effort-Reward Trade-off. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *36*(39), 10002–10015. https://doi.org/10.1523/JNEUROSCI.0292-16.2016

Knutson, B., & Cooper, J. C. (2005). Functional magnetic resonance imaging of reward prediction. *Current Opinion in Neurology*, *18*(4), 411–417. https://doi.org/10.1097/01.wco.0000173463.24758.f6

Kobayashi, S., & Schultz, W. (2008). Influence of Reward Delays on Responses of Dopamine Neurons. *The Journal of Neuroscience*, *28*(31), 7837. https://doi.org/10.1523/JNEUROSCI.1600-08.2008

Kojima, Y., Iwamoto, Y., & Yoshida, K. (2004). Memory of Learning Facilitates Saccadic Adaptation in the Monkey. *Journal of Neuroscience*, *24*(34), 7531–7539. https://doi.org/10.1523/JNEUROSCI.1741-04.2004

Kooij, K. van der, & Smeets, J. B. J. (2019). Reward-based motor adaptation can generalize across actions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(1), 71–81. https://doi.org/10.1037/xlm0000573

Körding, K. P., Fukunaga, I., Howard, I. S., Ingram, J. N., & Wolpert, D. M. (2004). A Neuroeconomics Approach to Inferring Utility Functions in Sensorimotor Control. *PLOS Biology*, *2*(10), e330. https://doi.org/10.1371/journal.pbio.0020330

Körding, K. P., & Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature*, *427*(6971), 244–247. https://doi.org/10.1038/nature02169

Krakauer, J. W., Ghilardi, M.-F., Mentis, M., Barnes, A., Veytsman, M., Eidelberg, D., & Ghez, C. (2004). Differential Cortical and Subcortical Activations in Learning Rotations and Gains for Reaching: A PET Study. *Journal of Neurophysiology*, *91*(2), 924–933. https://doi.org/10.1152/jn.00675.2003

Krakauer, J. W., Pine, Z. M., Ghilardi, M.-F., & Ghez, C. (2000). Learning of Visuomotor Transformations for Vectorial Planning of Reaching Trajectories. *Journal of Neuroscience*, *20*(23), 8916–8924. https://doi.org/10.1523/JNEUROSCI.20-23-08916.2000

Kreussel, L., Hewig, J., Kretschmer, N., Hecht, H., Coles, M. G. H., & Miltner, W. H. R. (2012). The influence of the magnitude, probability, and valence of potential wins and losses on the amplitude of the feedback negativity. *Psychophysiology*, *49*(2), 207–219. https://doi.org/10.1111/j.1469-8986.2011.01291.x

Krigolson, O. E. (2018). Event-related brain potentials and the study of reward processing: Methodological considerations. *International Journal of Psychophysiology*, *132*, 175–183. https://doi.org/10.1016/j.ijpsycho.2017.11.007

Kumar, N., Manning, T. F., & Ostry, D. J. (2019). Somatosensory cortex participates in the consolidation of human motor memory. *PLOS Biology*, *17*(10), e3000469. https://doi.org/10.1371/journal.pbio.3000469

Kurniawan, I. T., Guitart-Masip, M., & Dolan, R. J. (2011). Dopamine and Effort-Based Decision Making. *Frontiers in Neuroscience*, *5*. https://doi.org/10.3389/fnins.2011.00081

Kurniawan, I. T., Seymour, B., Talmi, D., Yoshida, W., Chater, N., & Dolan, R. J. (2010). Choosing to Make an Effort: The Role of Striatum in Signaling Physical Effort of a Chosen Action. *Journal of Neurophysiology*, *104*(1), 313–321. https://doi.org/10.1152/jn.00027.2010

Kurtzer, I. L., Pruszynski, J. A., & Scott, S. H. (2008). Long-latency reflexes of the human arm reflect an internal model of limb dynamics. *Current Biology: CB*, *18*(6), 449–453. https://doi.org/10.1016/j.cub.2008.02.053

Landi, S. M., Baguear, F., & Della-Maggiore, V. (2011). One Week of Motor Adaptation Induces Structural Changes in Primary Motor Cortex That Predict Long-Term Memory One Year Later. *Journal of Neuroscience*, *31*(33), 11808–11813. https://doi.org/10.1523/JNEUROSCI.2253-11.2011

Lee, J.-Y., & Schweighofer, N. (2009). Dual Adaptation Supports a Parallel Architecture of Motor Memory. *Journal of Neuroscience*, *29*(33), 10396–10404. https://doi.org/10.1523/JNEUROSCI.1294-09.2009

Leow, L.-A., de Rugy, A., Loftus, A. M., & Hammond, G. (2013). Different mechanisms contributing to savings and anterograde interference are impaired in Parkinson's disease. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00055

Leow, L.-A., de Rugy, A., Marinovic, W., Riek, S., & Carroll, T. J. (2016). Savings for visuomotor adaptation require prior history of error, not prior repetition of successful actions. *Journal of Neurophysiology*, *116*(4), 1603–1614. https://doi.org/10.1152/jn.01055.2015

Leow, L.-A., Loftus, A. M., & Hammond, G. R. (2012). Impaired savings despite intact initial learning of motor adaptation in Parkinson's disease. *Experimental Brain Research*, *218*(2), 295–304. https://doi.org/10.1007/s00221-012-3060-5

Leow, L.-A., Marinovic, W., Rugy, A. de, & Carroll, T. J. (2020). Task Errors Drive Memories That Improve Sensorimotor Adaptation. *Journal of Neuroscience*, *40*(15), 3075–3088. https://doi.org/10.1523/JNEUROSCI.1506-19.2020

Lerner, T. N., & Kreitzer, A. C. (2011). Neuromodulatory control of striatal plasticity and behavior. *Current Opinion in Neurobiology*, *21*(2), 322–327. https://doi.org/10.1016/j.conb.2011.01.005

Li, C.-S. R., Padoa-Schioppa, C., & Bizzi, E. (2001). Neuronal Correlates of Motor Performance and Motor Learning in the Primary Motor Cortex of Monkeys Adapting to an External Force Field. *Neuron*, *30*(2), 593–607. https://doi.org/10.1016/S0896-6273(01)00301-4

Liao, Y., Gramann, K., Feng, W., Deák, G. O., & Li, H. (2011). This ought to be good: Brain activity accompanying positive and negative expectations and outcomes. *Psychophysiology*, *48*(10), 1412–1419. https://doi.org/10.1111/j.1469-8986.2011.01205.x

Lohse, K. R., Miller, M. W., Daou, M., Valerius, W., & Jones, M. (2020). Dissociating the contributions of reward-prediction errors to trial-level adaptation and long-term learning. *Biological Psychology*, *149*, 107775. https://doi.org/10.1016/j.biopsycho.2019.107775

Luft, A. R., & Schwarz, S. (2009). Dopaminergic signals in primary motor cortex. *International Journal of Developmental Neuroscience*, *27*(5), 415–421. https://doi.org/10.1016/j.ijdevneu.2009.05.004

Luque, D., López, F. J., Marco-Pallares, J., Càmara, E., & Rodríguez-Fornells, A. (2012). Feedback-related Brain Potential Activity Complies with Basic Assumptions of Associative Learning Theory. *Journal of Cognitive Neuroscience*, *24*(4), 794–808. https://doi.org/10.1162/jocn_a_00145

Lydall, E. S., Gilmour, G., & Dwyer, D. M. (2010). Rats place greater value on rewards produced by high effort: An animal analogue of the "effort justification" effect. *Journal of Experimental Social Psychology*, *46*(6), 1134–1137. https://doi.org/10.1016/j.jesp.2010.05.011

Ma, Q., Meng, L., Wang, L., & Shen, Q. (2014). I endeavor to make it: Effort increases valuation of subsequent monetary reward. *Behavioural Brain Research*, *261*, 1–7. https://doi.org/10.1016/j.bbr.2013.11.045

Maeda, R. S., Cluff, T., Gribble, P. L., & Pruszynski, J. A. (2018). Feedforward and Feedback Control Share an Internal Model of the Arm's Dynamics. *The Journal of Neuroscience*, *38*(49), 10505–10514. https://doi.org/10.1523/JNEUROSCI.1709-18.2018

Maeda, R. S., Gribble, P. L., & Pruszynski, J. A. (2020). Learning New Feedforward Motor Commands Based on Feedback Responses. *Current Biology: CB*, *30*(10), 1941-1948.e3. https://doi.org/10.1016/j.cub.2020.03.005

Marco-Pallarés, J., Cucurell, D., Cunillera, T., Krämer, U. M., Càmara, E., Nager, W., Bauer, P., Schüle, R., Schöls, L., Münte, T. F., & Rodriguez-Fornells, A. (2009). Genetic variability in the dopamine system (dopamine receptor D4, catechol-O-methyltransferase) modulates neurophysiological responses to gains and losses. *Biological Psychiatry*, *66*(2), 154–161. https://doi.org/10.1016/j.biopsych.2009.01.006

Marinelli, L., Crupi, D., Di Rocco, A., Bove, M., Eidelberg, D., Abbruzzese, G., & Felice Ghilardi, M. (2009). Learning and consolidation of visuo-motor adaptation in Parkinson's disease. *Parkinsonism & Related Disorders*, *15*(1), 6–11. https://doi.org/10.1016/j.parkreldis.2008.02.012

Marsden, C. D., Rothwell, J. C., & Day, B. L. (1984). The use of peripheral feedback in the control of movement. *Trends in Neurosciences*, *7*(7), 253–257. https://doi.org/10.1016/S0166-2236(84)80218-0

Martín, R. S., Appelbaum, L. G., Pearson, J. M., Huettel, S. A., & Woldorff, M. G. (2013). Rapid Brain Responses Independently Predict Gain Maximization and Loss Minimization during Economic Decision Making. *Journal of Neuroscience*, *33*(16), 7011–7019. https://doi.org/10.1523/JNEUROSCI.4242-12.2013

Martínez-Horta, S., Riba, J., Bobadilla, R. F. de, Pagonabarraga, J., Pascual-Sedano, B., Antonijoan, R. M., Romero, S., Mañanas, M. À., García-Sanchez, C., & Kulisevsky, J. (2014). Apathy in Parkinson's Disease: Neurophysiological Evidence of Impaired Incentive Processing. *Journal of Neuroscience*, *34*(17), 5918–5926. https://doi.org/10.1523/JNEUROSCI.0251-14.2014

Maschke, M., Gomez, C. M., Ebner, T. J., & Konczak, J. (2004). Hereditary cerebellar ataxia progressively impairs force adaptation during goal-directed arm movements. *Journal of Neurophysiology*, *91*(1), 230–238. https://doi.org/10.1152/jn.00557.2003

Mastrigt, N. M. van, Smeets, J. B. J., & Kooij, K. van der. (2020). Quantifying exploration in reward-based motor learning. *PLOS ONE*, *15*(4), e0226789. https://doi.org/10.1371/journal.pone.0226789

Mathis, M. W., Mathis, A., & Uchida, N. (2017). Somatosensory Cortex Plays an Essential Role in Forelimb Motor Adaptation in Mice. *Neuron*, *93*(6), 1493-1503.e6. https://doi.org/10.1016/j.neuron.2017.02.049

Matsumoto, M., Matsumoto, K., Abe, H., & Tanaka, K. (2007). Medial prefrontal cell activity signaling prediction errors of action values. *Nature Neuroscience*, *10*(5), 647–656. https://doi.org/10.1038/nn1890

Mazzoni, P., & Krakauer, J. W. (2006). An Implicit Plan Overrides an Explicit Strategy during Visuomotor Adaptation. *Journal of Neuroscience*, *26*(14), 3642–3645. https://doi.org/10.1523/JNEUROSCI.5317-05.2006

McDougle, S. D., Ivry, R. B., & Taylor, J. A. (2016). Taking Aim at the Cognitive Side of Learning in Sensorimotor Adaptation Tasks. *Trends in Cognitive Sciences*, *20*(7), 535–544. https://doi.org/10.1016/j.tics.2016.05.002

Mehler, D. M. A., Reichenbach, A., Klein, J., & Diedrichsen, J. (2017). Minimizing endpoint variability through reinforcement learning during reaching movements involving shoulder, elbow and wrist. *PloS One*, *12*(7), e0180803. https://doi.org/10.1371/journal.pone.0180803

Mehta, B., & Schaal, S. (2002). Forward Models in Visuomotor Control. *Journal of Neurophysiology*, *88*(2), 942–953. https://doi.org/10.1152/jn.2002.88.2.942

Miall, R. C., Christensen, L. O. D., Cain, O., & Stanley, J. (2007). Disruption of State Estimation in the Human Lateral Cerebellum. *PLOS Biology*, *5*(11), e316. https://doi.org/10.1371/journal.pbio.0050316

Miall, R. C., Jenkinson, N., & Kulkarni, K. (2004). Adaptation to rotated visual feedback: A re-examination of motor interference. *Experimental Brain Research*, *154*(2), 201–210. https://doi.org/10.1007/s00221-003-1630-2

Miall, R. C., Weir, D. J., Wolpert, D. M., & Stein, J. F. (1993). Is the Cerebellum a Smith Predictor? *Journal of Motor Behavior*, *25*(3), 203–216. https://doi.org/10.1080/00222895.1993.9942050

Miall, R. C., & Wolpert, D. M. (1996). Forward Models for Physiological Motor Control. *Neural Networks*, *9*(8), 1265–1279. https://doi.org/10.1016/S0893-6080(96)00035-4

Miltner, W. H., Braun, C. H., & Coles, M. G. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: Evidence for a "generic" neural system for error detection. *Journal of Cognitive Neuroscience*, *9*(6), 788–798. https://doi.org/10.1162/jocn.1997.9.6.788

Mink, J. W. (1996). The basal ganglia: Focused selection and inhibition of competing motor programs. *Progress in Neurobiology*, *50*(4), 381–425. https://doi.org/10.1016/s0301-0082(96)00042-1

Mohebi, A., Pettibone, J. R., Hamid, A. A., Wong, J.-M. T., Vinson, L. T., Patriarchi, T., Tian, L., Kennedy, R. T., & Berke, J. D. (2019). Dissociable dopamine dynamics for learning and motivation. *Nature*, *570*(7759), 65–70. https://doi.org/10.1038/s41586-019-1235-y

Molen, N. H., Rozendal, R. H., & Boon, W. (1972). Graphic representation of the relationship between oxygen-consumption and characteristics of normal gait of the human male. *Proceedings of the Koninklijke Nederlandse Akademie Van Wetenschappen. Series C. Biological and Medical Sciences*, *75*(4), 305–314.

Molina-Luna, K., Pekanovic, A., Röhrich, S., Hertler, B., Schubring-Giese, M., Rioult-Pedotti, M.-S., & Luft, A. R. (2009). Dopamine in Motor Cortex Is Necessary for Skill Learning and Synaptic Plasticity. *PLOS ONE*, *4*(9), e7082. https://doi.org/10.1371/journal.pone.0007082

Moore, B., Khang, S., & Francis, J. T. (2020). Noise-Correlation Is Modulated by Reward Expectation in the Primary Motor Cortex Bilaterally During Manual and Observational Tasks in Primates. *Frontiers in Behavioral Neuroscience*, *14*. https://doi.org/10.3389/fnbeh.2020.541920

Morel, P., Ulbrich, P., & Gail, A. (2017). What makes a reach movement effortful? Physical effort discounting supports common minimization principles in decision making and motor control. *PLOS Biology*, *15*(6), e2001323. https://doi.org/10.1371/journal.pbio.2001323

Morton, S. M., & Bastian, A. J. (2006). Cerebellar Contributions to Locomotor Adaptations during Splitbelt Treadmill Walking. *Journal of Neuroscience*, *26*(36), 9107–9116. https://doi.org/10.1523/JNEUROSCI.2622-06.2006

Mostafa, A. A., Hart, B. M. 't, & Henriques, D. Y. P. (2019). Motor learning without moving: Proprioceptive and predictive hand localization after passive visuoproprioceptive discrepancy training. *PLOS ONE*, *14*(8), e0221861. https://doi.org/10.1371/journal.pone.0221861

Mueller, E. M., Burgdorf, C., Chavanon, M.-L., Schweiger, D., Hennig, J., Wacker, J., & Stemmler, G. (2014). The COMT Val158Met polymorphism regulates the effect of a dopamine antagonist on the feedback-related negativity. *Psychophysiology*, *51*(8), 805–809. https://doi.org/10.1111/psyp.12226

Müller, F., & Dichgans, J. (1994). Dyscoordination of pinch and lift forces during grasp in patients with cerebellar lesions. *Experimental Brain Research*, *101*(3), 485–492. https://doi.org/10.1007/BF00227341

Mutha, P. K., Sainburg, R. L., & Haaland, K. Y. (2011). Left Parietal Regions Are Critical for Adaptive Visuomotor Control. *Journal of Neuroscience*, *31*(19), 6972–6981. https://doi.org/10.1523/JNEUROSCI.6432-10.2011

Nasir, S. M., Darainy, M., & Ostry, D. J. (2013). Sensorimotor adaptation changes the neural coding of somatosensory stimuli. *Journal of Neurophysiology*, *109*(8), 2077–2085. https://doi.org/10.1152/jn.00719.2012

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

Nieuwenhuis, S., Ridderinkhof, K. R., Talsma, D., Coles, M. G. H., Holroyd, C. B., Kok, A., & van der Molen, M. W. (2002). A computational account of altered error processing in older age: Dopamine and the error-related negativity. *Cognitive, Affective, & Behavioral Neuroscience*, *2*(1), 19–36. https://doi.org/10.3758/CABN.2.1.19

Niki, H., & Watanabe, M. (1979). Prefrontal and cingulate unit activity during timing behavior in the monkey. *Brain Research*, *171*(2), 213–224. https://doi.org/10.1016/0006-8993(79)90328-7

Nikooyan, A. A., & Ahmed, A. A. (2014). Reward feedback accelerates motor learning. *Journal of Neurophysiology*, *113*(2), 633–646. https://doi.org/10.1152/jn.00032.2014

Niv, Y., Daw, N. D., Joel, D., & Dayan, P. (2007). Tonic dopamine: Opportunity costs and the control of response vigor. *Psychopharmacology*, *191*(3), 507–520. https://doi.org/10.1007/s00213-006-0502-4

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable Roles of Ventral and Dorsal Striatum in Instrumental Conditioning. *Science*, *304*(5669), 452–454. https://doi.org/10.1126/science.1094285

Ohashi, H., Valle-Mena, R., Gribble, P. L., & Ostry, D. J. (2019). Movements following force-field adaptation are aligned with altered sense of limb position. *Experimental Brain Research*, *237*(5), 1303–1313. https://doi.org/10.1007/s00221-019-05509-y

Ölveczky, B. P., Otchy, T. M., Goldberg, J. H., Aronov, D., & Fee, M. S. (2011). Changes in the neural control of a complex motor sequence during learning. *Journal of Neurophysiology*, *106*(1), 386–397. https://doi.org/10.1152/jn.00018.2011

Opris, I., Lebedev, M., & Nelson, R. J. (2011). Motor Planning under Unpredictable Reward: Modulations of Movement Vigor and Primate Striatum Activity. *Frontiers in Neuroscience*, *5*, 61. https://doi.org/10.3389/fnins.2011.00061

Orban de Xivry, J.-J., & Lefèvre, P. (2015). Formation of model-free motor memories during motor adaptation depends on perturbation schedule. *Journal of Neurophysiology*, *113*(7), 2733–2741. https://doi.org/10.1152/jn.00673.2014

Ostry, D. J., Darainy, M., Mattar, A. A. G., Wong, J., & Gribble, P. L. (2010). Somatosensory Plasticity and Motor Learning. *Journal of Neuroscience*, *30*(15), 5384–5393. https://doi.org/10.1523/JNEUROSCI.4571-09.2010

Ostry, D. J., & Gribble, P. L. (2016). Sensory plasticity in human motor learning. *Trends in Neurosciences*, *39*(2), 114–123. https://doi.org/10.1016/j.tins.2015.12.006

Otani, S., Daniel, H., Roisin, M.-P., & Crepel, F. (2003). Dopaminergic modulation of long-term synaptic plasticity in rat prefrontal neurons. *Cerebral Cortex (New York, N.Y.: 1991)*, *13*(11), 1251–1256. https://doi.org/10.1093/cercor/bhg092

Oya, H., Adolphs, R., Kawasaki, H., Bechara, A., Damasio, A., & Howard, M. A. (2005). Electrophysiological correlates of reward prediction error recorded in the human prefrontal cortex. *Proceedings of the National Academy of Sciences*, *102*(23), 8351–8356. https://doi.org/10.1073/pnas.0500899102

Padoa-Schioppa, C., & Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature*, *441*(7090), 223–226. https://doi.org/10.1038/nature04676

Padoa-Schioppa, C., Li, C.-S. R., & Bizzi, E. (2004). Neuronal Activity in the Supplementary Motor Area of Monkeys Adapting to a New Dynamic Environment. *Journal of Neurophysiology*, *91*(1), 449–473. https://doi.org/10.1152/jn.00876.2002

Parent, A., & Hazrati, L. N. (1995). Functional anatomy of the basal ganglia. I. The cortico-basal ganglia-thalamo-cortical loop. *Brain Research. Brain Research Reviews*, *20*(1), 91–127. https://doi.org/10.1016/0165-0173(94)00007-c

Parker, J. G., Marshall, J. D., Ahanonu, B., Wu, Y.-W., Kim, T. H., Grewe, B. F., Zhang, Y., Li, J. Z., Ding, J. B., Ehlers, M. D., & Schnitzer, M. J. (2018). Diametric neural ensemble dynamics in parkinsonian and dyskinetic states. *Nature*, *557*(7704), 177–182. https://doi.org/10.1038/s41586-018-0090-6

Pasalar, S., Roitman, A. V., Durfee, W. K., & Ebner, T. J. (2006). Force field effects on cerebellar Purkinje cell discharge with implications for internal models. *Nature Neuroscience*, *9*(11), 1404–1411. https://doi.org/10.1038/nn1783

Patterson, J. R., Brown, L. E., Wagstaff, D. A., & Sainburg, R. L. (2017). Limb position drift results from misalignment of proprioceptive and visual maps. *Neuroscience*, *346*, 382–394. https://doi.org/10.1016/j.neuroscience.2017.01.040

Pawlak, V., & Kerr, J. N. D. (2008). Dopamine Receptor Activation Is Required for Corticostriatal Spike-Timing-Dependent Plasticity. *Journal of Neuroscience*, *28*(10), 2435–2446. https://doi.org/10.1523/JNEUROSCI.4402-07.2008

Paz, R., Boraud, T., Natan, C., Bergman, H., & Vaadia, E. (2003). Preparatory activity in motor cortex reflects learning of local visuomotor skills. *Nature Neuroscience*, *6*(8), 882–890. https://doi.org/10.1038/nn1097

Paz, R., Natan, C., Boraud, T., Bergman, H., & Vaadia, E. (2005). Emerging Patterns of Neuronal Responses in Supplementary and Primary Motor Areas during Sensorimotor Adaptation. *Journal of Neuroscience*, *25*(47), 10941–10951. https://doi.org/10.1523/JNEUROSCI.0164-05.2005

Paz, R., & Vaadia, E. (2004). Learning-Induced Improvement in Encoding and Decoding of Specific Movement Directions by Neurons in the Primary Motor Cortex. *PLOS Biology*, *2*(2), e45. https://doi.org/10.1371/journal.pbio.0020045

Pekny, S. E., Criscimagna-Hemminger, S. E., & Shadmehr, R. (2011). Protection and Expression of Human Motor Memories. *Journal of Neuroscience*, *31*(39), 13829–13839. https://doi.org/10.1523/JNEUROSCI.1704-11.2011

Pekny, S. E., Izawa, J., & Shadmehr, R. (2015). Reward-dependent modulation of movement variability. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *35*(9), 4015–4024. https://doi.org/10.1523/JNEUROSCI.3244-14.2015

Perich, M. G., Gallego, J. A., & Miller, L. E. (2018). A Neural Population Mechanism for Rapid Learning. *Neuron*, *100*(4), 964-976.e7. https://doi.org/10.1016/j.neuron.2018.09.030

Perrin, E., & Venance, L. (2019). Bridging the gap between striatal plasticity and learning. *Current Opinion in Neurobiology*, *54*, 104–112. https://doi.org/10.1016/j.conb.2018.09.007

Pessiglione, M., Schmidt, L., Draganski, B., Kalisch, R., Lau, H., Dolan, R. J., & Frith, C. D. (2007). How the Brain Translates Money into Force: A Neuroimaging Study of Subliminal Motivation. *Science*, *316*(5826), 904–906. https://doi.org/10.1126/science.1140459

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*, *442*(7106), 1042–1045. https://doi.org/10.1038/nature05051

Pfabigan, D. M., Alexopoulos, J., Bauer, H., & Sailer, U. (2011). Manipulation of feedback expectancy and valence induces negative and positive reward prediction error signals manifest in event-related brain potentials. *Psychophysiology*, *48*(5), 656–664. https://doi.org/10.1111/j.1469-8986.2010.01136.x

Philiastides, M. G., Biele, G., Vavatzanidis, N., Kazzer, P., & Heekeren, H. R. (2010). Temporal dynamics of prediction error processing during reward-based decision making. *NeuroImage*, *53*(1), 221–232. https://doi.org/10.1016/j.neuroimage.2010.05.052

Phillips, P. E. M., Walton, M. E., & Jhou, T. C. (2007). Calculating utility: Preclinical evidence for cost–benefit analysis by mesolimbic dopamine. *Psychopharmacology*, *191*(3), 483–495. https://doi.org/10.1007/s00213-006-0626-6

Polezzi, D., Sartori, G., Rumiati, R., Vidotto, G., & Daum, I. (2010). Brain correlates of risky decision-making. *NeuroImage*, *49*(2), 1886–1894. https://doi.org/10.1016/j.neuroimage.2009.08.068

Polich, J. (2007). Updating P300: An Integrative Theory of P3a and P3b. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, *118*(10), 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019

Popa, L. S., & Ebner, T. J. (2019). Cerebellum, Predictions and Errors. *Frontiers in Cellular Neuroscience*, *12*. https://doi.org/10.3389/fncel.2018.00524

Popa, L. S., Hewitt, A. L., & Ebner, T. J. (2012). Predictive and Feedback Performance Errors Are Signaled in the Simple Spike Discharge of Individual Purkinje Cells. *Journal of Neuroscience*, *32*(44), 15345–15358. https://doi.org/10.1523/JNEUROSCI.2151-12.2012

Popa, L. S., Hewitt, A. L., & Ebner, T. J. (2013). Purkinje Cell Simple Spike Discharge Encodes Error Signals Consistent with a Forward Internal Model. *The Cerebellum*, *12*(3), 331–333. https://doi.org/10.1007/s12311-013-0452-4

Porter, B. S., Hillman, K. L., & Bilkey, D. K. (2019). Anterior cingulate cortex encoding of effortful behavior. *Journal of Neurophysiology*, *121*(2), 701–714. https://doi.org/10.1152/jn.00654.2018

Prévost, C., Pessiglione, M., Météreau, E., Cléry-Melin, M.-L., & Dreher, J.-C. (2010). Separate Valuation Subsystems for Delay and Effort Decision Costs. *Journal of Neuroscience*, *30*(42), 14080–14090. https://doi.org/10.1523/JNEUROSCI.2752-10.2010

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

Quattrocchi, G., Monaco, J., Ho, A., Irmen, F., Strube, W., Ruge, D., Bestmann, S., & Galea, J. M. (2018). Pharmacological Dopamine Manipulation Does Not Alter Reward-Based Improvements in Memory Retention during a Visuomotor Adaptation Task. *ENeuro*, *5*(3). https://doi.org/10.1523/ENEURO.0453-17.2018

Quilodran, R., Rothé, M., & Procyk, E. (2008). Behavioral shifts and action valuation in the anterior cingulate cortex. *Neuron*, *57*(2), 314–325. https://doi.org/10.1016/j.neuron.2007.11.031

Ralston, H. J. (1958). Energy-speed relation and optimal speed during level walking. *Internationale Zeitschrift Fur Angewandte Physiologie, Einschliesslich Arbeitsphysiologie*, *17*(4), 277–283. https://doi.org/10.1007/BF00698754

Ramakrishnan, A., Byun, Y. W., Rand, K., Pedersen, C. E., Lebedev, M. A., & Nicolelis, M. A. L. (2017). Cortical neurons multiplex reward-related signals along with sensory and motor information. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(24), E4841–E4850. https://doi.org/10.1073/pnas.1703668114

Ramkumar, P., Dekleva, B., Cooler, S., Miller, L., & Kording, K. (2016). Premotor and Motor Cortices Encode Reward. *PLoS ONE*, *11*(8). https://doi.org/10.1371/journal.pone.0160851

Ramnani, N., Elliott, R., Athwal, B. S., & Passingham, R. E. (2004). Prediction error for free monetary reward in the human prefrontal cortex. *NeuroImage*, *23*(3), 777–786. https://doi.org/10.1016/j.neuroimage.2004.07.028

Rand, M. K., & Heuer, H. (2019). Visual and proprioceptive recalibrations after exposure to a visuomotor rotation. *European Journal of Neuroscience*, *50*(8), 3296–3310. https://doi.org/10.1111/ejn.14433

Reppert, T. R., Lempert, K. M., Glimcher, P. W., & Shadmehr, R. (2015). Modulation of Saccade Vigor during Value-Based Decision Making. *Journal of Neuroscience*, *35*(46), 15369–15378. https://doi.org/10.1523/JNEUROSCI.2621-15.2015

Requarth, T., & Sawtell, N. B. (2014). Plastic Corollary Discharge Predicts Sensory Consequences of Movements in a Cerebellum-Like Circuit. *Neuron*, *82*(4), 896–907. https://doi.org/10.1016/j.neuron.2014.03.025

Reynolds, J. N., Hyland, B. I., & Wickens, J. R. (2001). A cellular mechanism of reward-related learning. *Nature*, *413*(6851), 67–70. https://doi.org/10.1038/35092560

Reynolds, J. N., & Wickens, J. R. (2002). Dopamine-dependent plasticity of corticostriatal synapses. *Neural Networks*, *15*(4), 507–521. https://doi.org/10.1016/S0893-6080(02)00045-X

Rigoux, L., & Guigon, E. (2012). A Model of Reward- and Effort-Based Optimal Decision Making and Motor Control. *PLOS Computational Biology*, *8*(10), e1002716. https://doi.org/10.1371/journal.pcbi.1002716

Rogers, R. D., Ramnani, N., Mackay, C., Wilson, J. L., Jezzard, P., Carter, C. S., & Smith, S. M. (2004). Distinct portions of anterior cingulate cortex and medial prefrontal cortex are activated by reward processing in separable phases of decision-making cognition. *Biological Psychiatry*, *55*(6), 594–602. https://doi.org/10.1016/j.biopsych.2003.11.012

Rothwell, J. C., Traub, M. M., Day, B. L., Obeso, J. A., Thomas, P. K., & Marsden, C. D. (1982). Manual motor performance in a deafferented man. *Brain: A Journal of Neurology*, *105 (Pt 3)*, 515–542. https://doi.org/10.1093/brain/105.3.515

Rubin, J. E., Vich, C., Clapp, M., Noneman, K., & Verstynen, T. (2020). The credit assignment problem in cortico-basal ganglia-thalamic networks: A review, a problem and a possible solution. *The European Journal of Neuroscience*. https://doi.org/10.1111/ejn.14745

Russek, E. M., Momennejad, I., Botvinick, M. M., Gershman, S. J., & Daw, N. D. (2017). Predictive representations can link model-based reinforcement learning to model-free mechanisms. *PLoS Computational Biology*, *13*(9), e1005768. https://doi.org/10.1371/journal.pcbi.1005768

Rutledge, R. B., Dean, M., Caplin, A., & Glimcher, P. W. (2010). Testing the reward prediction error hypothesis with an axiomatic model. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *30*(40), 13525–13536. https://doi.org/10.1523/JNEUROSCI.1747-10.2010

Salamone, J. D., Correa, M., Farrar, A., & Mingote, S. M. (2007). Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits. *Psychopharmacology*, *191*(3), 461–482. https://doi.org/10.1007/s00213-006-0668-9

Salamone, J. D., Cousins, M. S., & Bucher, S. (1994). Anhedonia or anergia? Effects of haloperidol and nucleus accumbens dopamine depletion on instrumental response selection in a T-maze cost/benefit procedure. *Behavioural Brain Research*, *65*(2), 221–229. https://doi.org/10.1016/0166-4328(94)90108-2

Sallet, J., Quilodran, R., Rothé, M., Vezoli, J., Joseph, J.-P., & Procyk, E. (2007). Expectations, gains, and losses in the anterior cingulate cortex. *Cognitive, Affective & Behavioral Neuroscience*, *7*(4), 327–336. https://doi.org/10.3758/cabn.7.4.327

Sambrook, T. D., & Goslin, J. (2015). A neural reward prediction error revealed by a meta-analysis of ERPs using great grand averages. *Psychological Bulletin*, *141*(1), 213–235. http://dx.doi.org.proxy1.lib.uwo.ca/10.1037/bul0000006

Sambrook, T. D., & Goslin, J. (2016). Principal components analysis of reward prediction errors in a reinforcement learning task. *NeuroImage*, *124*, 276–286. https://doi.org/10.1016/j.neuroimage.2015.07.032

Sambrook, T. D., Hardwick, B., Wills, A. J., & Goslin, J. (2018). Model-free and model-based reward prediction errors in EEG. *NeuroImage*, *178*, 162–171. https://doi.org/10.1016/j.neuroimage.2018.05.023

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

Santesso, D. L., Dillon, D. G., Birk, J. L., Holmes, A. J., Goetz, E., Bogdan, R., & Pizzagalli, D. A. (2008). Individual differences in reinforcement learning: Behavioral, electrophysiological, and neuroimaging correlates. *NeuroImage*, *42*(2), 807–816. https://doi.org/10.1016/j.neuroimage.2008.05.032

Santesso, D. L., Evins, A. E., Frank, M. J., Schetter, E. C., Bogdan, R., & Pizzagalli, D. A. (2009). Single dose of a dopamine agonist impairs reinforcement learning in humans: Evidence from event-related potentials and computational modeling of striatal-cortical function. *Human Brain Mapping*, *30*(7), 1963–1976.

Sato, A., Yasuda, A., Ohira, H., Miyawaki, K., Nishikawa, M., Kumano, H., & Kuboki, T. (2005). Effects of value and reward magnitude on feedback negativity and P300. *NeuroReport*, *16*(4), 407–411.

Schall, J. D., Stuphorn, V., & Brown, J. W. (2002). Monitoring and control of action by the frontal lobes. *Neuron*, *36*(2), 309–322. https://doi.org/10.1016/s0896-6273(02)00964-9

Schlerf, J. E., Xu, J., Klemfuss, N. M., Griffiths, T. L., & Ivry, R. B. (2012). Individuals with cerebellar degeneration show similar adaptation deficits with large and small visuomotor errors. *Journal of Neurophysiology*, *109*(4), 1164–1173. https://doi.org/10.1152/jn.00654.2011

Schultz, W. (2007). Multiple dopamine functions at different time courses. *Annual Review of Neuroscience*, *30*, 259–288. https://doi.org/10.1146/annurev.neuro.28.061604.135722

Schultz, W. (2016). Dopamine reward prediction error coding. *Dialogues in Clinical Neuroscience*, *18*(1), 23–32.

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

Schutte, I., Deschamps, P. K. H., van Harten, P. N., & Kenemans, J. L. (2020). Dopaminergic and noradrenergic manipulation of anticipatory reward and probability event-related potentials. *Psychopharmacology*, *237*(7), 2019–2030. https://doi.org/10.1007/s00213-020-05515-x

Schween, R., McDougle, S. D., Hegele, M., & Taylor, J. A. (2020). Assessing explicit strategies in force field adaptation. *Journal of Neurophysiology*, *123*(4), 1552–1565. https://doi.org/10.1152/jn.00427.2019

Schweimer, J., & Hauber, W. (2006). Dopamine D1 receptors in the anterior cingulate cortex regulate effort-based decision making. *Learning & Memory*, *13*(6), 777–782. https://doi.org/10.1101/lm.409306

Schweimer, J., Saft, S., & Hauber, W. (2005). Involvement of catecholamine neurotransmission in the rat anterior cingulate in effort-related decision making. *Behavioral Neuroscience*, *119*(6), 1687–1692. https://doi.org/10.1037/0735-7044.119.6.1687

Sedaghat-Nejad, E., Herzfeld, D. J., & Shadmehr, R. (2019). Reward Prediction Error Modulates Saccade Vigor. *The Journal of Neuroscience*, *39*(25), 5010–5017. https://doi.org/10.1523/JNEUROSCI.0432-19.2019

Seki, K., & Fetz, E. E. (2012). Gating of Sensory Input at Spinal and Cortical Levels during Preparation and Execution of Voluntary Movement. *Journal of Neuroscience*, *32*(3), 890–902. https://doi.org/10.1523/JNEUROSCI.4958-11.2012

Seki, K., Perlmutter, S. I., & Fetz, E. E. (2003). Sensory input to primate spinal cord is presynaptically inhibited during voluntary movement. *Nature Neuroscience*, *6*(12), 1309–1316. https://doi.org/10.1038/nn1154

Selinger, J. C., O'Connor, S. M., Wong, J. D., & Donelan, J. M. (2015). Humans Can Continuously Optimize Energetic Cost during Walking. *Current Biology*, *25*(18), 2452–2456. https://doi.org/10.1016/j.cub.2015.08.016

Selinger, J. C., Wong, J. D., Simha, S. N., & Donelan, J. M. (2019). How humans initiate energy optimization and converge on their optimal gaits. *The Journal of Experimental Biology*, *222*(Pt 19). https://doi.org/10.1242/jeb.198234

Seo, H., & Lee, D. (2008). Cortical mechanisms for reinforcement learning in competitive games. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1511), 3845–3857. https://doi.org/10.1098/rstb.2008.0158

Shadmehr, R., Huang, H. J., & Ahmed, A. A. (2016a). A Representation of Effort in Decision-Making and Motor Control. *Current Biology*, *26*(14), 1929–1934. https://doi.org/10.1016/j.cub.2016.05.065

Shadmehr, R., Huang, H. J., & Ahmed, A. A. (2016b). Effort, reward, and vigor in decision-making and motor control. *Current Biology : CB*, *26*(14), 1929–1934. https://doi.org/10.1016/j.cub.2016.05.065

Shadmehr, R., & Krakauer, J. W. (2008). A computational neuroanatomy for motor control. *Experimental Brain Research*, *185*(3), 359–381. https://doi.org/10.1007/s00221-008-1280-5

Shadmehr, R., & Mussa-Ivaldi, F. A. (1994). Adaptive representation of dynamics during learning of a motor task. *Journal of Neuroscience*, *14*(5), 3208–3224. https://doi.org/10.1523/JNEUROSCI.14-05-03208.1994

Shadmehr, R., Smith, M. A., & Krakauer, J. W. (2010). Error correction, sensory prediction, and adaptation in motor control. *Annual Review of Neuroscience*, *33*, 89–108. https://doi.org/10.1146/annurev-neuro-060909-153135

Shahar, N., Moran, R., Hauser, T. U., Kievit, R. A., McNamee, D., Moutoussis, M., NSPN Consortium, & Dolan, R. J. (2019). Credit assignment to state-independent task representations and its relationship with model-based decision making. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(32), 15871–15876. https://doi.org/10.1073/pnas.1821647116

Sharp, M. E., Foerde, K., Daw, N. D., & Shohamy, D. (2016). Dopamine selectively remediates "model-based" reward learning: A computational approach. *Brain: A Journal of Neurology*, *139*(Pt 2), 355–364. https://doi.org/10.1093/brain/awv347

Shen, W., Flajolet, M., Greengard, P., & Surmeier, D. J. (2008). Dichotomous Dopaminergic Control of Striatal Synaptic Plasticity. *Science*, *321*(5890), 848–851. https://doi.org/10.1126/science.1160575

Shergill, S. S., Bays, P. M., Frith, C. D., & Wolpert, D. M. (2003). Two Eyes for an Eye: The Neuroscience of Force Escalation. *Science*, *301*(5630), 187–187. https://doi.org/10.1126/science.1085327

Shima, K., & Tanji, J. (1998). Role for cingulate motor area cells in voluntary movement selection based on reward. *Science (New York, N.Y.)*, *282*(5392), 1335–1338. https://doi.org/10.1126/science.282.5392.1335

Shindou, T., Shindou, M., Watanabe, S., & Wickens, J. (2019). A silent eligibility trace enables dopamine-dependent synaptic plasticity for reinforcement learning in the mouse striatum. *European Journal of Neuroscience*, *49*(5), 726–736. https://doi.org/10.1111/ejn.13921

Shmuelof, L., Huang, V. S., Haith, A. M., Delnicki, R. J., Mazzoni, P., & Krakauer, J. W. (2012). Overcoming motor "forgetting" through reinforcement of learned actions. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *32*(42), 14617–14621. https://doi.org/10.1523/JNEUROSCI.2184-12.2012

Sidarta, A., Vahdat, S., Bernardi, N. F., & Ostry, D. J. (2016). Somatic and Reinforcement-Based Plasticity in the Initial Stages of Human Motor Learning. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *36*(46), 11682–11692. https://doi.org/10.1523/JNEUROSCI.1767-16.2016

Skvortsova, V., Palminteri, S., & Pessiglione, M. (2014). Learning to minimize efforts versus maximizing rewards: Computational principles and neural correlates. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *34*(47), 15621–15630. https://doi.org/10.1523/JNEUROSCI.1350-14.2014

Smith, E. H., Banks, G. P., Mikell, C. B., Cash, S. S., Patel, S. R., Eskandar, E. N., & Sheth, S. A. (2015). Frequency-Dependent Representation of Reinforcement-Related Information in the Human Medial and Lateral Prefrontal Cortex. *Journal of Neuroscience*, *35*(48), 15827–15836. https://doi.org/10.1523/JNEUROSCI.1864-15.2015

Smith, M. A., Ghazizadeh, A., & Shadmehr, R. (2006). Interacting Adaptive Processes with Different Timescales Underlie Short-Term Motor Learning. *PLOS Biology*, *4*(6), e179. https://doi.org/10.1371/journal.pbio.0040179

Smith, M. A., & Shadmehr, R. (2005). Intact Ability to Learn Internal Models of Arm Dynamics in Huntington's Disease But Not Cerebellar Degeneration. *Journal of Neurophysiology*, *93*(5), 2809–2821. https://doi.org/10.1152/jn.00943.2004

Sokolov, A. A., Miall, R. C., & Ivry, R. B. (2017). The Cerebellum: Adaptive Prediction for Movement and Cognition. *Trends in Cognitive Sciences*, *21*(5), 313–332. https://doi.org/10.1016/j.tics.2017.02.005

Soltani, M., & Knight, R. T. (2000). Neural origins of the P300. *Critical Reviews in Neurobiology*, *14*(3–4), 199–224.

Sommer, M. A., & Wurtz, R. H. (2006). Influence of the thalamus on spatial visual processing in frontal cortex. *Nature*, *444*(7117), 374–377. https://doi.org/10.1038/nature05279

Spering, M., Schütz, A. C., Braun, D. I., & Gegenfurtner, K. R. (2011). Keep your eyes on the ball: Smooth pursuit eye movements enhance prediction of visual motion. *Journal of Neurophysiology*, *105*(4), 1756–1767. https://doi.org/10.1152/jn.00344.2010

Steinberg, E. E., Keiflin, R., Boivin, J. R., Witten, I. B., Deisseroth, K., & Janak, P. H. (2013). A causal link between prediction errors, dopamine neurons and learning. *Nature Neuroscience*, *16*(7), 966–973. https://doi.org/10.1038/nn.3413

Stemmer, B., Segalowitz, S. J., Witzke, W., & Schönle, P. W. (2004). Error detection in patients with lesions to the medial prefrontal cortex: An ERP study. *Neuropsychologia*, *42*(1), 118–130. https://doi.org/10.1016/s0028-3932(03)00121-0

Streng, M. L., Popa, L. S., & Ebner, T. J. (2018). Modulation of sensory prediction error in Purkinje cells during visual feedback manipulations. *Nature Communications*, *9*(1), 1099. https://doi.org/10.1038/s41467-018-03541-0

Summerside, E. M., Shadmehr, R., & Ahmed, A. A. (2018). Vigor of reaching movements: Reward discounts the cost of effort. *Journal of Neurophysiology*, *119*(6), 2347–2357. https://doi.org/10.1152/jn.00872.2017

Sun, S., & Wang, S. (2020). The neural basis of feedback-guided behavioral adjustment. *Neuroscience Letters*, *736*, 135243. https://doi.org/10.1016/j.neulet.2020.135243

Sutton, R. S., & Barto, A. G. (2011). *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press.

Suzuki, S., Lawlor, V. M., Cooper, J. A., Arulpragasam, A. R., & Treadway, M. T. (2020). Distinct regions of the striatum underlying effort, movement initiation and effort discounting. *Nature Human Behaviour*, 1–11. https://doi.org/10.1038/s41562-020-00972-y

Swick, D., & Turken, A. U. (2002). Dissociation between conflict detection and error monitoring in the human anterior cingulate cortex. *Proceedings of the National Academy of Sciences*, *99*(25), 16354–16359. https://doi.org/10.1073/pnas.252521499

Synofzik, M., Lindner, A., & Thier, P. (2008). The cerebellum updates predictions about the visual consequences of one's behavior. *Current Biology: CB*, *18*(11), 814–818. https://doi.org/10.1016/j.cub.2008.04.071

Talmi, D., Fuentemilla, L., Litvak, V., Duzel, E., & Dolan, R. J. (2012). An MEG signature corresponding to an axiomatic model of reward prediction error. *NeuroImage*, *59*(1), 635–645. https://doi.org/10.1016/j.neuroimage.2011.06.051

Tan, H., Jenkinson, N., & Brown, P. (2014). Dynamic Neural Correlates of Motor Error Monitoring and Adaptation during Trial-to-Trial Learning. *Journal of Neuroscience*, *34*(16), 5678–5688. https://doi.org/10.1523/JNEUROSCI.4739-13.2014

Tanaka, H., Ishikawa, T., & Kakei, S. (2019). Neural Evidence of the Cerebellum as a State Predictor. *The Cerebellum*, *18*(3), 349–371. https://doi.org/10.1007/s12311-018-0996-4

Tanaka, H., Krakauer, J. W., & Sejnowski, T. J. (2012). Generalization and Multirate Models of Motor Adaptation. *Neural Computation*, *24*(4), 939–966. https://doi.org/10.1162/NECO_a_00262

Tanaka, S. C., Doya, K., Okada, G., Ueda, K., Okamoto, Y., & Yamawaki, S. (2004). Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nature Neuroscience*, *7*(8), 887–893. https://doi.org/10.1038/nn1279

Tanaka, S., O'Doherty, J. P., & Sakagami, M. (2019). The cost of obtaining rewards enhances the reward prediction error signal of midbrain dopamine neurons. *Nature Communications*, *10*(1), 3674. https://doi.org/10.1038/s41467-019-11334-2

Taniai, Y., & Nishii, J. (2015). Optimality of Upper-Arm Reaching Trajectories Based on the Expected Value of the Metabolic Energy Cost. *Neural Computation*, *27*(8), 1721–1737. https://doi.org/10.1162/NECO_a_00757

Taylor, J. A., Krakauer, J. W., & Ivry, R. B. (2014). Explicit and Implicit Contributions to Learning in a Sensorimotor Adaptation Task. *Journal of Neuroscience*, *34*(8), 3023–3032. https://doi.org/10.1523/JNEUROSCI.3619-13.2014

Therrien, A. S., & Bastian, A. J. (2015). Cerebellar damage impairs internal predictions for sensory and motor function. *Current Opinion in Neurobiology*, *33*, 127–133. https://doi.org/10.1016/j.conb.2015.03.013

Therrien, A. S., Wolpert, D. M., & Bastian, A. J. (2016). Effective reinforcement learning following cerebellar damage requires a balance between exploration and motor noise. *Brain*, *139*(1), 101–114. https://doi.org/10.1093/brain/awv329

Thoroughman, K. A., & Shadmehr, R. (2000). Learning of action through adaptive combination of motor primitives. *Nature*, *407*(6805), 742–747. https://doi.org/10.1038/35037588

Todorov, E. (2004). Optimality principles in sensorimotor control. *Nature Neuroscience*, *7*(9), 907–915. https://doi.org/10.1038/nn1309

Tseng, Y., Diedrichsen, J., Krakauer, J. W., Shadmehr, R., & Bastian, A. J. (2007). Sensory Prediction Errors Drive Cerebellum-Dependent Adaptation of Reaching. *Journal of Neurophysiology*, *98*(1), 54–62. https://doi.org/10.1152/jn.00266.2007

Tsutsui, K.-I., Grabenhorst, F., Kobayashi, S., & Schultz, W. (2016). A dynamic code for economic object valuation in prefrontal cortex neurons. *Nature Communications*, *7*(1), 12554. https://doi.org/10.1038/ncomms12554

Tumer, E. C., & Brainard, M. S. (2007). Performance variability enables adaptive plasticity of "crystallized" adult birdsong. *Nature*, *450*(7173), 1240–1244. https://doi.org/10.1038/nature06390

Tzschentke, T. M. (2001). Pharmacology and behavioral pharmacology of the mesocortical dopamine system. *Progress in Neurobiology*, *63*(3), 241–320. https://doi.org/10.1016/S0301-0082(00)00033-2

Uehara, S., Mawase, F., & Celnik, P. (2018). Learning Similar Actions by Reinforcement or Sensory-Prediction Errors Rely on Distinct Physiological Mechanisms. *Cerebral Cortex*, *28*(10), 3478–3490. https://doi.org/10.1093/cercor/bhx214

Ullsperger, M., Danielmeier, C., & Jocham, G. (2014). Neurophysiology of performance monitoring and adaptive behavior. *Physiological Reviews*, *94*(1), 35–79. https://doi.org/10.1152/physrev.00041.2012

Ullsperger, M., & von Cramon, D. Y. (2001). Subprocesses of performance monitoring: A dissociation of error processing and response competition revealed by event-related fMRI and ERPs. *NeuroImage*, *14*(6), 1387–1401. https://doi.org/10.1006/nimg.2001.0935

Umberger, B. R., & Martin, P. E. (2007). Mechanical power and efficiency of level walking with different stride rates. *Journal of Experimental Biology*, *210*(18), 3255–3265. https://doi.org/10.1242/jeb.000950

van der Helden, J., Boksem, M. A. S., & Blom, J. H. G. (2010). The importance of failure: Feedback-related negativity predicts motor learning efficiency. *Cerebral Cortex (New York, N.Y.: 1991)*, *20*(7), 1596–1603. https://doi.org/10.1093/cercor/bhp224

van Donkelaar, P., & Lee, R. G. (1994). Interactions between the eye and hand motor systems: Disruptions due to cerebellar dysfunction. *Journal of Neurophysiology*, *72*(4), 1674–1685. https://doi.org/10.1152/jn.1994.72.4.1674

Vazey, E. M., Moorman, D. E., & Aston-Jones, G. (2018). Phasic locus coeruleus activity regulates cortical encoding of salience information. *Proceedings of the National Academy of Sciences*, *115*(40), E9439–E9448. https://doi.org/10.1073/pnas.1803716115

Vaziri, S., Diedrichsen, J., & Shadmehr, R. (2006). Why Does the Brain Predict Sensory Consequences of Oculomotor Commands? Optimal Integration of the Predicted and the Actual Sensory Feedback. *Journal of Neuroscience*, *26*(16), 4188–4197. https://doi.org/10.1523/JNEUROSCI.4747-05.2006

Vercher, J. L., & Gauthier, G. M. (1988). Cerebellar involvement in the coordination control of the oculo-manual tracking system: Effects of cerebellar dentate nucleus lesion. *Experimental Brain Research*, *73*(1), 155–166. https://doi.org/10.1007/BF00279669

Vercher, J. L., Gauthier, G. M., Guedon, O., Blouin, J., Cole, J., & Lamarre, Y. (1996). Self-moved target eye tracking in control and deafferented subjects: Roles of arm motor command and proprioception in arm-eye coordination. *Journal of Neurophysiology*, *76*(2), 1133–1144. https://doi.org/10.1152/jn.1996.76.2.1133

Vo, A., Seergobin, K. N., Morrow, S. A., & MacDonald, P. A. (2016). Levodopa impairs probabilistic reversal learning in healthy young adults. *Psychopharmacology*, *233*(14), 2753–2763. https://doi.org/10.1007/s00213-016-4322-x

Vyas, S., Even-Chen, N., Stavisky, S. D., Ryu, S. I., Nuyujukian, P., & Shenoy, K. V. (2018). Neural Population Dynamics Underlying Motor Learning Transfer. *Neuron*, *97*(5), 1177-1186.e3. https://doi.org/10.1016/j.neuron.2018.01.040

Waelti, P., Dickinson, A., & Schultz, W. (2001). Dopamine responses comply with basic assumptions of formal learning theory. *Nature*, *412*(6842), 43–48. https://doi.org/10.1038/35083500

Wagner, M. J., & Smith, M. A. (2008). Shared Internal Models for Feedforward and Feedback Control. *Journal of Neuroscience*, *28*(42), 10663–10673. https://doi.org/10.1523/JNEUROSCI.5479-07.2008

Walsh, M. M., & Anderson, J. R. (2011). Learning from delayed feedback: Neural responses in temporal credit assignment. *Cognitive, Affective & Behavioral Neuroscience*, *11*(2), 131–143. https://doi.org/10.3758/s13415-011-0027-0

Walsh, M. M., & Anderson, J. R. (2012). Learning from experience: Event-related potential correlates of reward processing, neural adaptation, and behavioral choice. *Neuroscience and Biobehavioral Reviews*, *36*(8), 1870–1884. https://doi.org/10.1016/j.neubiorev.2012.05.008

Walton, M. E., Bannerman, D. M., Alterescu, K., & Rushworth, M. F. S. (2003). Functional Specialization within Medial Frontal Cortex of the Anterior Cingulate for Evaluating Effort-Related Decisions. *Journal of Neuroscience*, *23*(16), 6475–6479. https://doi.org/10.1523/JNEUROSCI.23-16-06475.2003

Wang, J., Chen, J., Lei, Y., & Li, P. (2014). P300, not feedback error-related negativity, manifests the waiting cost of receiving reward information. *NeuroReport*, *25*(13), 1044–1048. https://doi.org/10.1097/WNR.0000000000000226

Wang, J. X., Kurth-Nelson, Z., Kumaran, D., Tirumala, D., Soyer, H., Leibo, J. Z., Hassabis, D., & Botvinick, M. M. (2018). Prefrontal cortex as a meta-reinforcement learning system. *Nature Neuroscience*, *21*(6), 860–868. https://doi.org/10.1038/s41593-018-0147-8

Wann, J. P., & Ibrahim, S. F. (1992). Does limb proprioception drift? *Experimental Brain Research*, *91*(1), 162–166. https://doi.org/10.1007/BF00230024

Warren, C. M., Hyman, J. M., Seamans, J. K., & Holroyd, C. B. (2015). Feedback-related negativity observed in rodent anterior cingulate cortex. *Journal of Physiology-Paris*, *109*(1), 87–94. https://doi.org/10.1016/j.jphysparis.2014.08.008

Watabe-Uchida, M., Eshel, N., & Uchida, N. (2017). Neural Circuitry of Reward Prediction Error. *Annual Review of Neuroscience*, *40*, 373–394. https://doi.org/10.1146/annurev-neuro-072116-031109

Watts, A. T. M., Bachman, M. D., & Bernat, E. M. (2017). Expectancy effects in feedback processing are explained primarily by time-frequency delta not theta. *Biological Psychology*, *129*, 242–252. https://doi.org/10.1016/j.biopsycho.2017.08.054

Weeks, H. M., Therrien, A. S., & Bastian, A. J. (2017). The cerebellum contributes to proprioception during motion. *Journal of Neurophysiology*, *118*(2), 693–702. https://doi.org/10.1152/jn.00417.2016

Wei, K., & Koerding, K. (2010). Uncertainty of feedback and state estimation determines the speed of motor adaptation. *Frontiers in Computational Neuroscience*, *4*. https://doi.org/10.3389/fncom.2010.00011

Weiner, M. J., Hallett, M., & Funkenstein, H. H. (1983). Adaptation to lateral displacement of vision in patients with lesions of the central nervous system. *Neurology*, *33*(6), 766–772. https://doi.org/10.1212/wnl.33.6.766

Williams, Z. M., Bush, G., Rauch, S. L., Cosgrove, G. R., & Eskandar, E. N. (2004). Human anterior cingulate neurons and the integration of monetary reward with motor responses. *Nature Neuroscience*, *7*(12), 1370–1375. https://doi.org/10.1038/nn1354

Wise, R. A. (2004). Dopamine, learning and motivation. *Nature Reviews Neuroscience*, *5*(6), 483–494. https://doi.org/10.1038/nrn1406

Wise, S. P., Moody, S. L., Blomstrom, K. J., & Mitz, A. R. (1998). Changes in motor cortical activity during visuomotor adaptation. *Experimental Brain Research*, *121*(3), 285–299. https://doi.org/10.1007/s002210050462

Wochner, I., Driess, D., Zimmermann, H., Haeufle, D. F. B., Toussaint, M., & Schmitt, S. (2020). Optimality Principles in Human Point-to-Manifold Reaching Accounting for Muscle Dynamics. *Frontiers in Computational Neuroscience*, *14*. https://doi.org/10.3389/fncom.2020.00038

Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science*, *269*(5232), 1880–1882. https://doi.org/10.1126/science.7569931

Wolpert, D. M., Miall, R. C., & Kawato, M. (1998). Internal models in the cerebellum. *Trends in Cognitive Sciences*, *2*(9), 338–347. https://doi.org/10.1016/S1364-6613(98)01221-2

Woolley, S. C. (2019). Dopaminergic regulation of vocal-motor plasticity and performance. *Current Opinion in Neurobiology*, *54*, 127–133. https://doi.org/10.1016/j.conb.2018.10.008

Woolley, S. C., Rajan, R., Joshua, M., & Doupe, A. J. (2014). Emergence of context-dependent variability across a basal ganglia network. *Neuron*, *82*(1), 208–223. https://doi.org/10.1016/j.neuron.2014.01.039

Wu, H. G., Miyamoto, Y. R., Castro, L. N. G., Ölveczky, B. P., & Smith, M. A. (2014). Temporal structure of motor variability is dynamically regulated and predicts motor learning ability. *Nature Neuroscience*, *17*(2), 312–321. https://doi.org/10.1038/nn.3616

Wunderlich, K., Smittenaar, P., & Dolan, R. J. (2012). Dopamine enhances model-based over model-free choice behavior. *Neuron*, *75*(3), 418–424. https://doi.org/10.1016/j.neuron.2012.03.042

Wurtz, R. H. (2008). NEURONAL MECHANISMS OF VISUAL STABILITY. *Vision Research*, *48*(20), 2070–2089. https://doi.org/10.1016/j.visres.2008.03.021

Wurtz, R. H. (2018). Corollary Discharge Contributions to Perceptual Continuity Across Saccades. *Annual Review of Vision Science*, *4*(1), 215–237. https://doi.org/10.1146/annurev-vision-102016-061207

Xiao, J., Padoa-Schioppa, C., & Bizzi, E. (2006). Neuronal correlates of movement dynamics in the dorsal and ventral premotor area in the monkey. *Experimental Brain Research*, *168*(1), 106–119. https://doi.org/10.1007/s00221-005-0074-2

Xu, Q., Shen, Q., Chen, P., Ma, Q., Sun, D., & Pan, Y. (2011). How an uncertain cue modulates subsequent monetary outcome evaluation: An ERP study. *Neuroscience Letters*, *505*(2), 200–204. https://doi.org/10.1016/j.neulet.2011.10.024

Xu-Wilson, M., Chen-Harris, H., Zee, D. S., & Shadmehr, R. (2009). Cerebellar contributions to adaptive control of saccades in humans. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *29*(41), 12930–12939. https://doi.org/10.1523/JNEUROSCI.3115-09.2009

Xu-Wilson, M., Tian, J., Shadmehr, R., & Zee, D. S. (2011). TMS Perturbs Saccade Trajectories and Unmasks an Internal Feedback Controller for Saccades. *Journal of Neuroscience*, *31*(32), 11537–11546. https://doi.org/10.1523/JNEUROSCI.1584-11.2011

Yasuda, A., Sato, A., Miyawaki, K., Kumano, H., & Kuboki, T. (2004). Error-related negativity reflects detection of negative reward prediction error. *NeuroReport*, *15*(16), 2561–2565.

Yeung, N., Holroyd, C. B., & Cohen, J. D. (2005). ERP Correlates of Feedback and Reward Processing in the Presence and Absence of Response Choice. *Cerebral Cortex*, *15*(5), 535–544. https://doi.org/10.1093/cercor/bhh153

Yeung, N., & Sanfey, A. G. (2004). Independent Coding of Reward Magnitude and Valence in the Human Brain. *Journal of Neuroscience*, *24*(28), 6258–6264. https://doi.org/10.1523/JNEUROSCI.4537-03.2004

Yoon, T., Jaleel, A., Ahmed, A. A., & Shadmehr, R. (2020). Saccade vigor and the subjective economic value of visual stimuli. *Journal of Neurophysiology*, *123*(6), 2161–2172. https://doi.org/10.1152/jn.00700.2019

Zarahn, E., Weston, G. D., Liang, J., Mazzoni, P., & Krakauer, J. W. (2008). Explaining Savings for Visuomotor Adaptation: Linear Time-Invariant State-Space Models Are Not Sufficient. *Journal of Neurophysiology*, *100*(5), 2537–2548. https://doi.org/10.1152/jn.90529.2008

Zarrugh, M. Y., Todd, F. N., & Ralston, H. J. (1974). Optimization of energy expenditure during level walking. *European Journal of Applied Physiology and Occupational Physiology*, *33*(4), 293–306. https://doi.org/10.1007/BF00430237

Zentall, T. R. (2010). Justification of Effort by Humans and Pigeons: Cognitive Dissonance or Contrast? *Current Directions in Psychological Science*, *19*(5), 296–300. https://doi.org/10.1177/0963721410383381

Zhou, Z., Yu, R., & Zhou, X. (2010). To do or not to do? Action enlarges the FRN and P300 effects in outcome evaluation. *Neuropsychologia*, *48*(12), 3606–3613. https://doi.org/10.1016/j.neuropsychologia.2010.08.010

Chapter 2

Neural signatures of reward and sensory error feedback processing in motor learning

2.1 Introduction

It is thought that at least two distinct learning processes can simultaneously contribute to sensorimotor adaptation, sensory error-based learning and reward-based learning (Galea et al., 2015; Huang et al., 2011; Izawa & Shadmehr, 2011; Nikooyan & Ahmed, 2014; Shmuelof et al., 2012). Electroencephalography (EEG) has been used to identify neural signatures of error and feedback processing in various motor learning and movement execution tasks (Krigolson et al., 2008; MacLean et al., 2015; Savoie et al., 2018; Torrecillos et al., 2014). However, it remains unclear how these neural responses relate to distinct reward- and sensory error-based motor learning mechanisms, as these processes are potentially confounded in typical experimental paradigms. Here we identified neural responses to feedback in separate motor adaptation paradigms designed to isolate reward-based learning and sensory error-based learning.

Visuomotor rotation (VMR) paradigms are a common experimental model for sensory-error based learning (Inoue et al., 2000; Krakauer, 2009; Krakauer et al., 2005; Rand & Heuer, 2019; H. Tanaka et al., 2009). In VMR tasks, participants typically reach to one or more targets arranged around a central start position. Visual feedback of hand position is rotated about the start position relative to the actual angle of reach. These perturbations are thought to produce sensory prediction errors caused by sensory feedback that indicates a state of the motor system that differs from the intended or predicted consequence of a motor command. Sensory prediction error is thought to elicit plasticity in cortico-cerebellar circuits that mediates motor adaptation to compensate for visuomotor perturbations (Block & Bastian, 2012; Diedrichsen et al., 2005; Izawa & Shadmehr, 2011; Mazzoni & Krakauer, 2006; M. A. Smith & Shadmehr, 2005; Synofzik et al., 2008; H. Tanaka et al., 2009).

Recent research suggests that a reward-based learning process can also contribute to motor adaptation in parallel to sensory error-based learning (Cashaback et al., 2017; Codol et al., 2018; Holland et al., 2018; Izawa & Shadmehr, 2011; Kooij & Smeets, 2019; Therrien et al., 2016). Reward-based learning has been isolated experimentally by

providing participants with only binary reinforcement feedback that indicates success or failure, without visual feedback of hand position. Reward-based motor learning has been modeled using computational reinforcement learning algorithms, which map actions to abstract representations of reward or success rather than to the sensory consequences of action (Dhawale et al., 2019; Izawa & Shadmehr, 2011).

The overarching hypothesis of this thesis is that fronto-striatal circuits contribute to reward-based motor adaptation by implementing a computational reinforcement learning process. We first sought to identify neural correlates of reward-prediction error, a canonical neural reinforcement learning signal, during reward-based motor adaptation. We were particularly interested in a medial-frontal event-related potential (ERP) component known as the feedback-related negativity/reward positivity (FRN/RP). The FRN/RP is characterized by a relatively positive voltage deflection in the midfrontal EEG in response to rewarding outcomes compared to a negativity observed in response to unfavorable outcomes (Falkenstein et al., 1991; Gehring et al., 1993; Proudfit, 2015). The reinforcement learning theory of the FRN/RP states that the component is a measure of reward prediction error signals generated by the medial frontal cortex that serve to reinforce successful actions and/or deter unrewarding actions (Holroyd & Coles, 2002; Sambrook & Goslin, 2015; Walsh & Anderson, 2012). Alternative accounts of the FRN/RP state that it is not specifically tied to the rewarding quality of outcomes but that it is a general prediction error or salience signal elicited by any unexpected outcome (W. H. Alexander & Brown, 2011; Hauser, Iannaccone, Stämpfli, et al., 2014; Hird et al., 2018; Soder & Potts, 2018; Talmi et al., 2013).

Our goal was to test the reinforcement learning theory of the FRN/RP in the context of motor adaptation by determining whether the FRN/RP is elicited specifically by reinforcement outcomes but not sensory error. Counterfactually, if the FRN/RP reflects a general error monitoring or prediction error signal, then sensory error should also elicit an FRN/RP. Previous studies have reported FRN/RP-like neural responses to sensory error feedback (Krigolson et al., 2008; MacLean et al., 2015; Savoie et al., 2018; Torrecillos et al., 2014). However, perturbation of sensory outcomes typically coincides with a failure

to meet the goals of the task, making it difficult to differentiate reinforcement-related error processing from sensory error processing. Furthermore, previous research has introduced perturbations during ongoing movements. In this case it is difficult to distinguish adaptation-related error processing from the recruitment of neural resources for ongoing control, which can elicit N200 ERP components with similar scalp distributions and timing as the FRN/RP (Folstein & Petten, 2008). We designed motor adaptation tasks to isolate reward-based and sensory error-based learning processes. We predicted that the FRN/RP would be elicited specifically by reinforcement outcomes but not sensory error. We provided feedback only at movement end point to avoid confounds due to movements themselves or error processing related to ongoing control. We further tested the hypothesis that the FRN/RP reflects a reward prediction error signal by manipulating the likelihood of reward feedback across conditions. Specifically, we predicted that the differential ERP response to reward vs. non reward feedback would be larger for infrequent outcomes compared to frequent outcomes.

The FRN/RP potential is superimposed on the P300, a well-characterized positive ERP component that peaks later than the FRN/RP and with a more posterior scalp distribution. It has been proposed that the P300 reflects the updating of internal models of stimulus context upon processing of unexpected stimuli to facilitate adaptive responding (E. Donchin & Coles, 1988; Krigolson et al., 2008; MacLean et al., 2015; Polich, 2007; San Martín, 2012). The P300 is observed ubiquitously in processing task-related feedback, and therefore we expected to detect a P300 in response to both sensory error and reward feedback. In accordance with the hypothesis that the P300 reflects a general processing of task-relevant feedback for behavioral adaptation, we predicted that P300 amplitude would be increased by sensory error and by infrequent reinforcement outcomes regardless of valence.

2.2 Materials and Methods

2.2.1 Experimental Design and Statistical Analysis

Participants made reaching movements toward a visual target and received visual feedback pertaining to reach angle only at movement end point. Neural responses to feedback were recorded by EEG. Participants were instructed that each reach terminating within the target would be rewarded with a small monetary bonus. Participants first performed a block of 50 practice trials. The subsequent behavioral procedure consisted of four blocks of a reward learning task and four blocks of a VMR task. The order of the blocks alternated between the two task types but was otherwise randomized. Participants took self-paced rests between blocks.

In the VMR task, a cursor appeared at movement end point to represent the position of the hand. In randomly selected trials, cursor feedback indicated a reach angle that was rotated relative to the unperturbed feedback. We tested for behavioral adaptation and modulation of ERPs in response to VMR. The perturbations were small relative to the size of the target, such that participants nearly always landed in the target, fulfilling the goal of the task and earning a monetary reward. Thus reward and task error were constant between perturbed and nonperturbed feedback, and by comparing the two conditions we could isolate the neural correlates of sensory error processing.

In the reward learning task, no cursor appeared to indicate the position of the hand. Instead, binary feedback represented whether or not participants succeeded in hitting the target. This allowed us to assess reward-based learning in isolation from sensory error processing, as visual information revealing the position of the hand was not provided. Reward was delivered probabilistically, with a higher probability of reward for reaches in one direction than the other, relative to participants' recent history of reach direction. We compared the neural responses to reward and non-reward feedback to assess the neural correlates of reward processing during adaptation.

Student's *t*-tests were performed with MATLAB R2016b, and the Lilliefors test was used to test the assumption of normality. In the case of nonnormal data, the Wilcoxon signed-rank test was used to test pairwise differences. Repeated-measures analyses of variance (ANOVAs) were conducted with IBM SPSS Statistics version 25. For all ANOVAs, Mauchly's test was used to validate the assumption of sphericity.

2.2.2 Participants

Data from $n = 20$ healthy, right-handed participants were analyzed and reported (23.21 ± 3.09 yr old; 12 women, 8 men). Three additional participants underwent the experimental procedure but were excluded because malfunction of the EEG recording equipment made the data unusable. One additional participant who reported performing movements based on a complex strategy that was unrelated to the experimental task was excluded. Participants provided written informed consent to experimental procedures approved by the Research Ethics Board at Western University.

2.2.3 Apparatus/Behavioral Task

Participants produced reaching movements with their right arm while holding the handle of a robotic arm (InMotion2; Interactive Motion Technologies; Fig. 2-1). Position of the robot handle was sampled at 600 Hz. A semisilvered mirror obscured vision of the arm and displayed visual information related to the task. An air sled supported each participant's right arm.

Participants reached to a white circular target 14 cm away from a circular start position (1-cm diameter) in front of their chest (Fig. 2-1*A*). The start position turned from red to green to cue the onset of each reach once the handle had remained inside it continuously for 750 ms. Participants were instructed that they must wait for the cue to begin each reach but that it was not necessary to react quickly upon seeing the cue.

Participants were instructed to make forward reaches and to stop their hand within the target. An arc-shaped cursor indicated reach extent throughout each movement without revealing reach angle. In only the first five baseline trials of each block, an additional

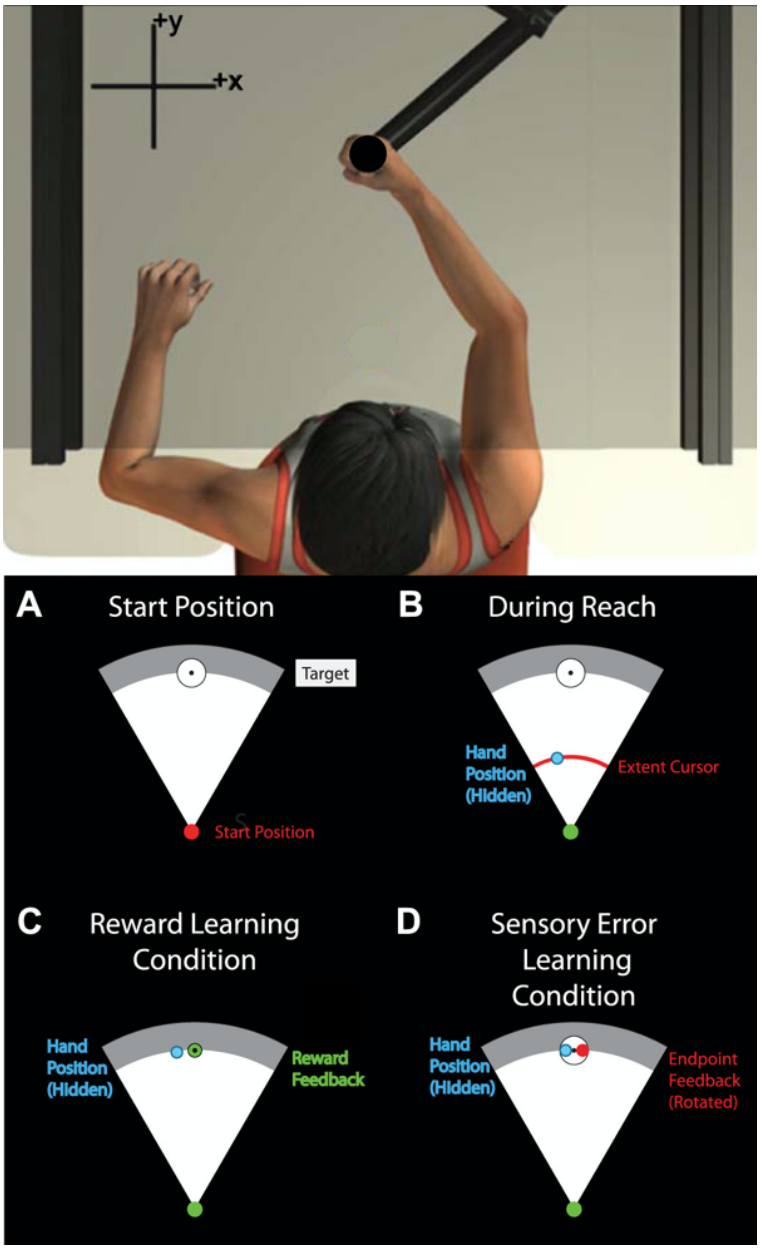

Figure 2-1: Experimental setup. ***A***, Participants (n = 20) reached to visual targets while holding the handle of a robotic arm. Vision of the arm was obscured by a screen that displayed visual information related to the task. ***B***, During reaches, hand position was hidden but an arc-shaped cursor indicated the extent of the reach without revealing reach angle. Feedback was provided at reach end point. ***C***, In the reward learning condition, binary feedback represented whether reaches were successful or unsuccessful in hitting the target by turning green or red, respectively. Reach adaptation was induced by providing reward for movements that did not necessarily correspond to the visual target. ***D***, In the visuomotor rotation condition, feedback represented the end-point position of the hand. Adaptation was induced by rotating the angle of the feedback relative to the actual reach angle.

circular cursor continuously indicated the position of the hand throughout the reach. A viscous force field assisted participants in braking their hand when the reach extent was >14 cm.

The robot ended each movement by fixing the handle position when the hand velocity decreased below 0.03 m/s. The hand was fixed in place for 700 ms, during which time visual feedback of reach angle was provided. Feedback indicated either reach end point position, a binary reward outcome, or feedback of movement speed (see below). Visual feedback was then removed, and the robot guided the hand back to the start position.

Reach end point was defined as the position at which the reach path intersected the perimeter of a circle (14-cm radius) centered at the start position. Reach angle was calculated as the angle between vectors defined by reach end point and the center of the target, each relative to the start position, such that reaching straight ahead corresponds to 0° and counterclockwise reach angles are positive. Feedback about reach angle was provided either in the form of end-point position feedback or binary reward feedback. The type of feedback, as well as various feedback manipulations, varied according to the assigned experimental block type (see *Reward Learning Task* and *Visuomotor Rotation Task*). Participants were told that they would earn additional monetary compensation for reaches that ended within the target, up to a maximum of CAD$10.

Movement duration was defined as the time elapsed between the hand leaving the start position and the moment hand velocity dropped below 0.03 m/s. If movement duration was >700 ms or <450 ms, no feedback pertaining to movement angle was provided. Instead, the gray arc behind the target turned blue or yellow to indicate that the reach was too slow or too fast, respectively. Participants were informed that movements with an incorrect speed would be repeated but would not otherwise affect the experiment.

To minimize the impact of eyeblink-related EEG artifacts, participants were asked to fixate their gaze on a black circular target in the center of the reach target and to refrain from blinking throughout each arm movement and subsequent presentation of feedback.

Practice Block: Each participant first completed a block of practice trials that continued until they achieved 50 movements within the desired range of movement duration. Continuous position feedback was provided during the first 5 trials, and only end-point position feedback was provided for the following 10 trials. Subsequently, no position feedback was provided outside the start position.

Reward Learning Task: Binary reward feedback was provided to induce adaptation of reach angle. Each participant completed four blocks in the reward learning condition. We manipulated feedback with direction of intended learning and reward frequency as factors, using a 2 × 2 design (direction of learning × reward frequency) across blocks. For each direction of intended learning (clockwise and counterclockwise), each participant experienced a block with high reward frequency and a block with low reward frequency. Reward frequency was manipulated to assess effects related to expectation, under the assumption that outcomes that occurred less frequently would violate expectations more strongly. Each block continued until the participant completed 115 reaches with acceptable movement duration. Participants reached toward a circular target 1.2 cm (4.9°) in diameter. The first 15 reaches were baseline trials during which continuous position feedback was provided during the first 5 trials, followed by 10 trials with only end-point position feedback. After these baseline trials no position feedback was provided, and binary reward feedback was instead provided at the end of the movement. Target hits and misses were indicated by the target turning green and red, respectively.

Unbeknownst to participants, reward feedback was delivered probabilistically. The likelihood of reward depended on the difference between the current reach angle and the median reach angle of the previous 10 reaches. In the high-reward frequency condition, reward was delivered at a probability of 100% if the difference between the current reach angle and the running median was in the direction of intended learning and at a probability of 30% otherwise (*Eq. 2.1*). When the running median was at least 6° away from zero in the direction of intended learning, reward was delivered at a fixed probability of 65%. This was intended to minimize conscious awareness of the manipulation by limiting adaptation to ±6°. In the low-reward frequency condition,

reward was similarly delivered at a probability of either 70% or 0% (*Eq. 2.2*). When the running median was at least 6° away from zero in the direction of intended learning, reward was delivered at a fixed probability of 35%. Reach angle and feedback throughout a representative experimental block are shown in Fig. 2-2.

We employed this adaptive, closed-loop reward schedule so that the overall frequency of reward was controlled, as reinforcement-related neural signals are highly sensitive to expectancy effects. The probability of reward on a given is described below:

$$p_{high} = \begin{cases} 1, & z \cdot (\theta_i - \text{median}(\theta_{i-10}, \dots, \theta_{i=1})) > 0 \\ .3, & z \cdot (\theta_i - \text{median}(\theta_{i-10}, \dots, \theta_{i=1})) < 0 \\ .65, & z \cdot \text{median}(\theta_{i-9}, \dots, \theta_i) > 6 \end{cases} \quad (2.1)$$

$$p_{low} = \begin{cases} .7, & z \cdot (\theta_i - \text{median}(\theta_{i-10}, \dots, \theta_{i=1})) > 0 \\ 0, & z \cdot (\theta_i - \text{median}(\theta_{i-10}, \dots, \theta_{i=1})) < 0 \\ .35, & z \cdot \text{median}(\theta_{i-9}, \dots, \theta_i) > 6 \end{cases} \quad (2.2)$$

where p is probability of reward described separately for the high- and low-reward frequency conditions, θ is the reach angle on trial i, $z = 1$ for counterclockwise learning blocks, and $z = -1$ for clockwise learning blocks.

Visuomotor rotation task: End-point feedback was rotated relative to the actual reach angle to induce sensory error-based adaptation. Each participant completed four blocks in the VMR condition. We manipulated feedback with initial rotation direction and perturbation size as factors using a 2 × 2 design across blocks. For each direction of initial rotation (clockwise and counterclockwise) each participant experienced a block with large rotation (1.5°) and a block with small rotation (0.75°). Each block continued until participants completed 125 reaches within acceptable movement duration limits. Participants reached toward a circular target 2.5 cm (10.2°) in diameter. Participants first performed baseline reaches during which cursor feedback reflected veridical reach angle continuously for the first 10 trials and only at movement end point for the subsequent 15 trials. After the baseline reaches the adaptation portion of each block began, unannounced to participants.

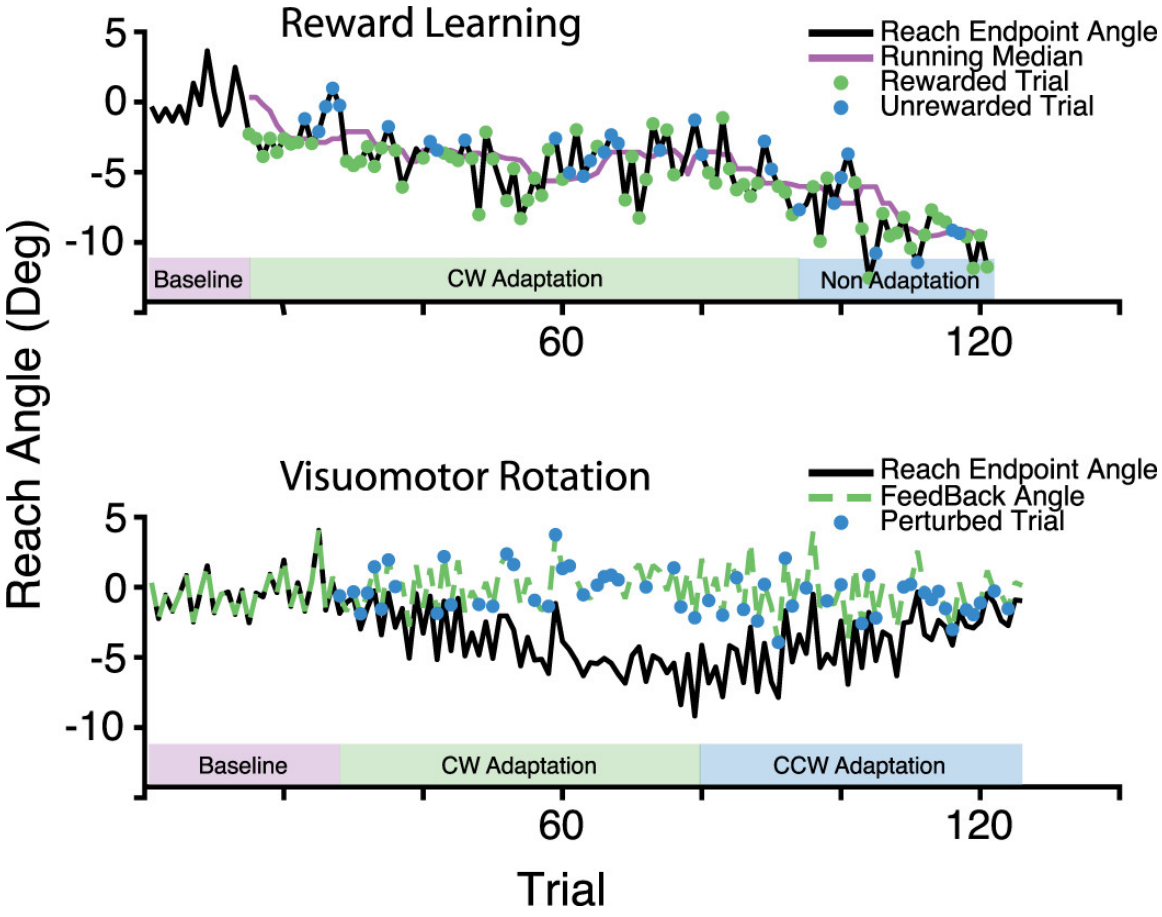

Figure 2-2: Reach angles of a representative participant (*n* = 1). ***Top*****:** the reward learning block assigned to the clockwise (CW) adaptation with the high reward frequency condition. Reaches were rewarded with 100.0% probability for reach angles less than the median of the previous 10 reaches and with 30.0% probability for reach angles greater than this running median. Reward was delivered at a fixed probability of 65.0% when the running median was less than −6°, indicated by the "Non-Adaptation" portion of the block. ***Bottom***: the visuomotor rotation block assigned to the 1.5° rotation condition. The rotation is imposed randomly in 50% of trials. The rotation is initially counterclockwise (CCW) but reverses when the mean of the previous 5 reach angles becomes less than −6.0°.

During the adaptation trials, end-point position feedback was provided that did not necessarily correspond to the true hand position. Participants were instructed that end-point feedback within the target would earn them bonus compensation, but no explicit reward feedback was provided. To determine the feedback angle in the small- and large-perturbation conditions, we added a rotation of 0.75° or 1.5°, respectively, to the true reach angle in a randomly selected 50% of trials. In addition, on every trial we subtracted an estimate of the current state of reach adaptation. Equation 2.3 describes the angle of endpoint cursor feedback provided on each trial.

$$\begin{aligned} X_i &= \theta_i + q - \text{mean}(\theta_{i-5}, \ldots, \theta_{i-1}) + \sum_{k=i-3}^{i-1} (.25 \cdot X_k) \\ q &= z \cdot s \cdot u \\ p(u) &= \begin{cases} .5, & \text{if } u = 1 \\ .5, & \text{if } u = 0 \end{cases} \end{aligned} \tag{2.3}$$

X denotes feedback angle, θ denotes reach angle, and q denotes the perturbation. z denotes the direction of the perturbation ($z = 1$ for counterclockwise perturbations and $z = -1$ for clockwise perturbations). s denotes the size of the perturbation (0.75° or 1.5° in the small- and large-error conditions, respectively). u is a discrete random variable that is realized as either 1 or 0 with equal probability (50%).

If the state of adaptation is accurately estimated and subtracted from the true reach angle, then a reach that reflects the state of adaptation without movement error will result in either unperturbed feedback at 0° or rotated feedback at the angle of the perturbation. The online estimate of adaptation consisted of a running average of the previous five reach angles and a model of reach adaptation that assumed that participants would adapt to a fixed proportion of the reach errors experienced during the previous three trials. A windowed average centered around the current reach angle could estimate the current state of reach adaptation, but the online running average was necessarily centered behind the current reach angle. Thus an online model was necessary to predict the state of adaptation. An adaptation rate of 0.25 was chosen for the online model on the basis of pilot data.

This design allowed us to compare perturbed and unperturbed feedback in randomly intermixed trials. Previous studies have imposed a fixed perturbation throughout a block of trials and compared early trials to late trials in which the error has been reduced through adaptation (MacLean et al., 2015; Tan et al., 2014). In such designs, differences in neural response might be attributed to changes in the state of adaptation or simply habituation to feedback, as opposed to sensory error per se. Alternatively, rotations can be imposed randomly in either direction, but previous work has demonstrated that neural and behavioral responses are larger for consistent perturbations, presumably because the sensorimotor system attributes variability in feedback to noise processes (Tan et al., 2014).

We sought to limit the magnitude of adaptation to 6° in an attempt to minimize awareness of the manipulation. The direction of the perturbation was reversed whenever the average reach angle in the previous five movements differed from zero by at least 6° in the direction of intended reach adaptation. Reach angle and feedback angle throughout a representative experimental block are shown in Fig. 2-2.

2.2.4 EEG Data Acquisition

EEG data were acquired from 16 cap-mounted electrodes with an active electrode system (g.GAMMA; g.tec Medical Engineering) and amplifier (g.USBamp; g.tec Medical Engineering). We recorded from electrodes placed according to the 10-20 System at sites Fz, FCz, Cz, CPz, Pz, POz, FP1, FP2, FT9, FT10, FC1, FC2, F3, F4, F7, and F8, referenced to an electrode placed on participants' left earlobe. Impedances were maintained below 5 kΩ. Data were sampled at 4,800 Hz and filtered online with band-pass (0.1–1,000 Hz) and notch (60 Hz) filters. A photodiode attached to the display monitor was used to synchronize recordings to stimulus onset.

2.2.5 Behavioral Data Analysis

Reward learning task: Motor learning scores were calculated for each participant as the difference between the average reach angle in the counterclockwise learning blocks and the average reach angle in the clockwise learning blocks. We assessed reach angle throughout the entire block primarily because reach direction was often unstable and a smaller window was susceptible to noise. Furthermore, this metric of learning reflected the rate of adaptation throughout the block without assuming a particular function for the time course of learning. Finally, this metric was not dependent on the choice of a particular subset of trials. We excluded baseline trials and trials that did not meet the movement duration criteria, as no feedback related to reach angle was provided on these trials (6.5% of trials in the VMR task, 7.4% of trials in the reward learning task).

Visuomotor rotation task: To quantify trial-by-trial learning we first calculated the change in reach angle between successive trials, as in Eq. 2.4:

$$\Delta\theta_i = \theta_{i+1} - \theta_i \qquad (2.4)$$

We then performed a linear regression on $\Delta\theta_i$ with the rotation imposed on trial i as the predictor variable. The rotation was 0°, ±0.75°, or ±1.5°. This regression was performed on an individual participant basis, separately for each of the four VMR conditions (corresponding to feedback rotations of −1.5°, −0.75°, 0.75°, and 1.5°). For these regressions, we excluded trials that did not meet the duration criteria or that resulted in an absolute visual error of >10° (mean = 2.65 trials per participant, SD= 4.3), as these large errors were thought to reflect execution errors or otherwise atypical movements. We took the average of the resulting slope estimates across blocks, multiplied by −1, as a metric of learning rate for each participant, as it reflects the portion of visual errors that participants corrected with a trial-by-trial adaptive process. Based on simulations of our experimental design using a standard memory updating model (Thoroughman & Shadmehr, 2000; not described here), we found that it was necessary to perform the regression separately for each rotation condition, as collapsing across the different rotation sizes and directions could introduce bias to the estimate of learning rate.

2.2.6 EEG Data Denoising

EEG data were resampled to 480 Hz and filtered off-line between 0.1 and 35 Hz with a second-order Butterworth filter. Continuous data were segmented into 2-s epochs time-locked to feedback stimulus onset at 0 ms (time range: −500 to +1,500 ms). Epochs flagged for containing artifacts as well as any channels with bad recordings were removed after visual inspection. Subsequently, extended infomax independent component analysis was performed on each participant's data (Delorme & Makeig, 2004). Components reflecting eye movements and blink artifacts were identified by visual inspection and subtracted by projection of the remaining components back to the voltage time series.

2.2.7 Event Related Potential Averaging

After artifact removal, we computed ERPs by trial averaging EEG time series epochs for various feedback conditions described in the sections below. ERPs were computed on an individual participant basis separately for recordings from channels FCz and Pz. All ERPs were baseline corrected by subtracting the average voltage in the 75-ms period immediately following stimulus onset. We used a baseline period following stimulus onset because stimuli were presented immediately upon movement termination and the period before stimulus presentation was more likely to be affected by movement-related artifacts. Trials in which reaches did not meet the movement duration criteria were excluded, as feedback relevant to reach adaptation was not provided on these trials (6.5% of trials in the VMR task, 7.4% of trials in the reward learning task).

Reward learning task: We computed ERPs separately for feedback conditions corresponding to "frequent reward," "infrequent reward," ,"frequent nonreward," and "infrequent nonreward." Reward in the high-reward frequency condition and nonreward in the low-reward frequency condition were deemed frequent, whereas reward in the low-reward frequency condition and nonreward in the high-reward frequency condition were deemed infrequent (Holroyd & Krigolson, 2007).

Visuomotor rotation task: We created trial-averaged ERP responses for trials with rotated feedback and trials with nonrotated feedback, separately for the 0.75° and 1.5° rotation

conditions. The resulting ERPs are identified by the conditions "rotated 0.75°," "nonrotated 0.75," "rotated 1.5°," and "nonrotated 1.5°."

To test for effects of absolute end-point error, which is determined not only by VMR but also by movement execution errors, we sorted trials in the adaptation portion of the VMR blocks by the absolute value of the angle of visual feedback relative to the center of the target. We created "most accurate" and "least accurate" ERPs for each participant by selecting the 75 trials with the smallest and largest absolute feedback angle, respectively.

2.2.8 Feedback-Related Negativity/ Reward Positivity Analysis

The FRN/RP was analyzed with a difference wave approach with ERPs recorded from FCz, where it is typically largest (Holroyd & Krigolson, 2007; Miltner et al., 1997; Pfabigan et al., 2011). Although the FRN/RP is classically characterized by a negative voltage peak following nonreward feedback, multiple lines of evidence suggest that a reward-related positivity also contributes to the variance captured by the difference wave approach, despite not producing a distinct peak (Baker & Holroyd, 2011; Becker et al., 2014; Carlson et al., 2011; Proudfit, 2015; Walsh & Anderson, 2012). Furthermore, difference waves can be computed separately for frequent and infrequent outcomes, which subtracts effects of pure surprise while preserving any interaction between feedback valence and reward frequency (Holroyd & Krigolson, 2007). Difference waves were computed for each participant by subtracting ERPs corresponding to unsuccessful outcomes from those corresponding to successful outcomes. FRN/RP amplitude was determined as the mean value of the difference wave between 200 and 350 ms after feedback presentation. This time window was chosen a priori on the basis of previous reports (see Walsh & Anderson, 2012, for meta-analysis). To test for the presence of the FRN/RP for each difference wave, we submitted FRN/RP amplitude to a *t*-test against zero.

Visuomotor rotation task: First, we created difference waves to test whether the rotations imposed on randomly selected trials elicited FRN/RP components. The rotated 0.75° ERPs were subtracted from the nonrotated 0.75° ERPs to create a "small VMR" difference

wave. The rotated 1.5° ERPs were subtracted from the nonrotated 1.5° ERPs to create a "large VMR" difference wave.

Next, we created a difference wave to test whether a FRN/RP was observable by comparing trials where the end-point feedback was furthest from the center of the target to those where feedback was closest to the center of the target. The "least accurate" ERPs were subtracted from the "most accurate" ERPs to create an "end-point error" difference wave.

Reward learning task: The frequent-nonreward ERP was subtracted from the frequent-reward ERP to create a "frequent" difference wave, and the infrequent-nonreward ERP was subtracted from the infrequent-reward EPR to create an "infrequent" difference wave.

2.2.9 P300 Analysis

To analyze the P300 we used ERPs recorded from channel Pz, where it is typically largest (Fabiani et al., 1987; Hajcak et al., 2005; MacLean et al., 2015; Polich, 2007). We calculated P300 amplitude using base-to-peak voltage difference. The temporal regions of interest (ROIs) for the peak and base were determined with grand averages computed across participants and conditions for each task (see *Visuomotor rotation task* and *Reward learning task* below). P300 peak was defined as the maximum peak occurring 250–500 ms after stimulus onset, which always corresponded to the largest peak in the analyzed epoch. P300 base was defined as the minimum preceding peak that occurred at least 100 ms after stimulus onset. For each subject, peak and base voltages were calculated separately for each condition ERP as the average voltage within 50-ms windows centered around the temporal ROIs defined at the group level. P300 amplitude was then determined as the difference between peak and base voltage.

Visuomotor rotation task: P300 amplitude was calculated in four conditions using the rotated 0.75°, nonrotated 0.75°, rotated 1.5°, and nonrotated 1.5° ERPs. Temporal ROIs were determined, as described above, by aggregating all trials across participants and the

four conditions into a single set and averaging to produce an "aggregate grand average from trials" waveform. This approach allows for data-driven ROI selection without inflated type I error rate and has been shown to be insensitive to trial number asymmetry across conditions (J. L. Brooks et al., 2017). We tested for differences in P300 amplitude related to VMR with two-way repeated-measures ANOVA with factors rotation (levels: nonrotated, rotated) and rotation magnitude (levels: 0.75°, 1.5°).

Reward learning task: P300 amplitude was calculated in four conditions using the infrequent reward, frequent reward, infrequent nonreward, and frequent nonreward feedback condition ERPs described above. Because the waveform morphology was considerably different for the ERPs elicited by reward feedback and those elicited by nonreward feedback, we defined temporal ROIs separately for the reward conditions (infrequent reward, frequent reward) and the nonreward conditions (infrequent nonreward and frequent nonreward). In both cases, temporal ROIs were determined by aggregating all trials across participants and the corresponding two conditions into a single set and averaging to produce an "aggregate grand average from trials" waveform. We tested for differences in P300 amplitude between feedback conditions with two-way repeated-measures ANOVA with factors reward (levels: rewarded, nonrewarded) and expectancy (levels: infrequent, frequent).

2.3 Results

2.3.1 Behavioral Results

Reward learning task: In the reward learning task participants adapted their reach angle on the basis of binary reward feedback (Fig. 2-3). We calculated a reward learning score for each subject by subtracting the average reach angle in the clockwise learning condition from that in the counterclockwise learning condition, excluding the baseline trials, such that the average reward learning score would be approximately zero if participants did not respond to the reward feedback in any way. We observed a mean reward learning score of 5.47 (SD 4.66), which is reliably greater than zero [1-sample t-

test; $t(19) = 5.25$, $P < 0.001$]. Participants received reward on 67.0% (SD 4.9) of trials in the high-frequency condition and 38.6% (SD 4.3) of trials in the low-frequency condition.

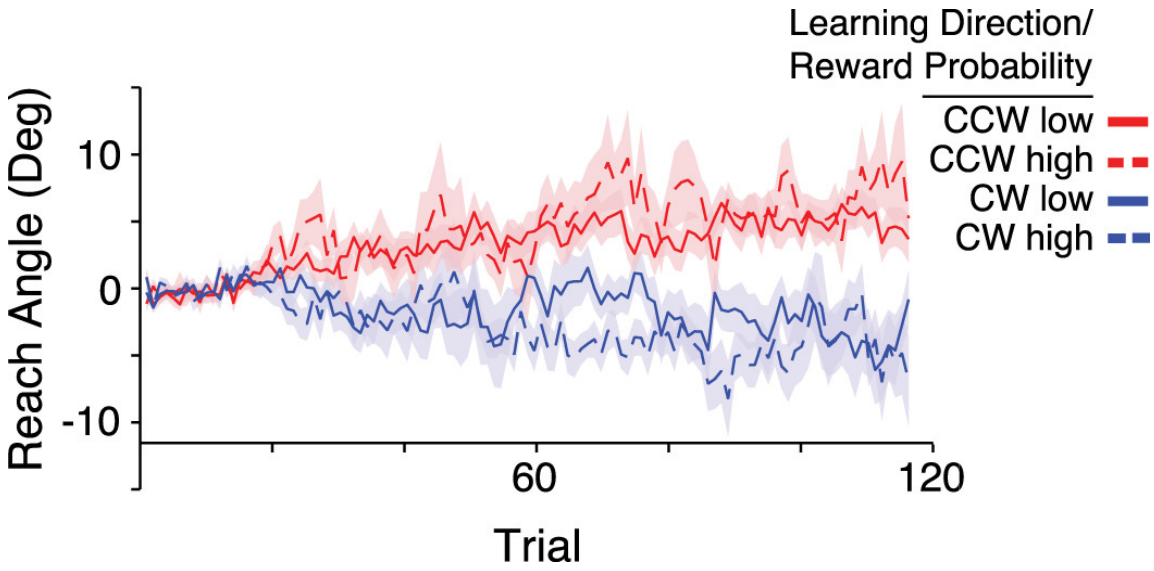

Figure 2-3: Behavior during the reward-based learning task. Participants (n = 20) adapted their reach angle in the reward learning condition. Group average reach angles in the reward learning conditions are plotted. Each participant completed 4 blocks. For each direction of intended learning [clockwise (CW) and counterclockwise (CCW)], each participant completed a block in the high-reward frequency (65%) condition and a block in the low-reward frequency (35%) condition. Shaded regions: ±1 SE.

Visuomotor rotation task: In the VMR task participants received end-point cursor feedback and adapted their reach angles in response to the rotated cursor feedback imposed on randomly selected trials. To estimate trial-by-trial learning rates for individual participants, we quantified the linear relationship between the change in reach angle after each trial with the rotation imposed on the preceding trial as the predictor variable, separately for each rotation condition (−1.5°, −0.75°, 0.75°, and 1.5°). We took the average of the resulting slope estimates and multiplied it by −1 to obtain a measure of learning rate. This metric reflects the proportion of VMR that each participant corrected with a trial-by-trial adaptive process. The mean learning rate was 0.49 (SD 0.46), which was significantly different from zero [1-sample t-test; $t(19) = 4.8$, $P < 0.001$]. This indicates that participants corrected for visual errors on a trial-by-trial basis. Figure 2-4 shows the average change in reach angle for each size and direction of the imposed cursor rotation.

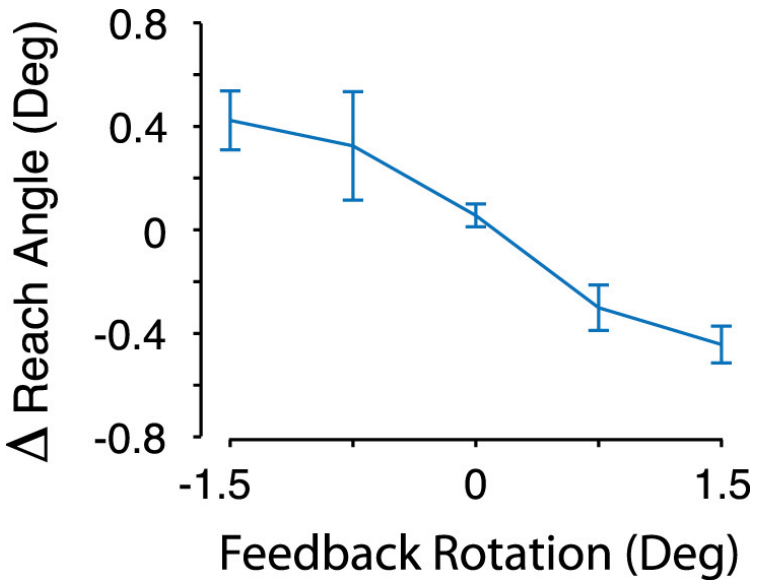

Figure 2-4: Behavior during the visuomotor rotation task. Participants (n = 20) adapted their reach angle on a trial-by-trial basis in the visuomotor rotation condition. The average change (Δ) in reach angle between subsequent pairs of trials is plotted for each size and direction of rotation imposed on the preceding trial. The average change in reach angle is in all cases opposite to the rotation, indicating that participants adapted their reaches to counteract the perturbations.

2.3.2 Feedback-Related Negativity/ Reward Positivity Results

Reward learning task: Figure 2-5*A* shows the ERPs recorded from electrode FCz during the reward learning condition, averaged across participants. The mean value of the "frequent" difference wave recorded from FCz between 200 and 350 ms was significantly different from zero [mean = 5.34 μV (SD 4.11), $t(19) = 5.81$, $P < 0.001$, 1-sample t-test], indicating that frequent feedback elicited a FRN/RP in our reward learning task. The mean value of the "infrequent" difference wave was also significantly larger than zero [mean = 7.09 μV (SD 2.76), $t(19) = 11.47$, $P < 0.001$, 1-sample t-test], indicating that infrequent feedback also elicited a FRN/RP.

The mean amplitude of the "infrequent" difference wave was larger than the mean amplitude of the "frequent" difference wave, although the difference was only marginally significant [$t(19) = 1.66$, $P = 0.056$, paired t-test, 1-tailed; Fig. 2-5*C*].

Visuomotor rotation task: Figure 2-6*A* shows the ERPs recorded from electrode FCz during the VMR condition, averaged across participants. The mean value of the "small VMR" difference wave recorded from FCz between 200 and 350 ms was not significantly different from zero (mean =−0.21 μV (SD 1.29), $Z = -0.67$, $W = 87$, $P = 0.50$, Wilcoxon signed-rank test; Fig. 2-6*C*). Similarly, the mean value of the "large VMR" difference wave recorded from FCz between 200 and 350 ms was not significantly different from zero [mean = −0.26 μV (SD 1.22), $t(19) = -0.97$, $P = 0.34$, 1-sample t-test; Fig. 2-6*C*). These findings indicate that the VMRs imposed in the VMR task did not reliably elicit a FRN/RP.

The mean value of the "end-point error" difference wave recorded from FCz between 200 and 350 ms was not significantly different from zero [mean = 0.61 μV (SD 3.28), $t(19) = 0.82$, $P = 0.42$, 1-sample t-test], indicating that a FRN/RP did not reliably occur on the basis of end-point error feedback. The fact that we were able to detect a FRN/RP in the reward learning task but not in the VMR task is consistent with the notion

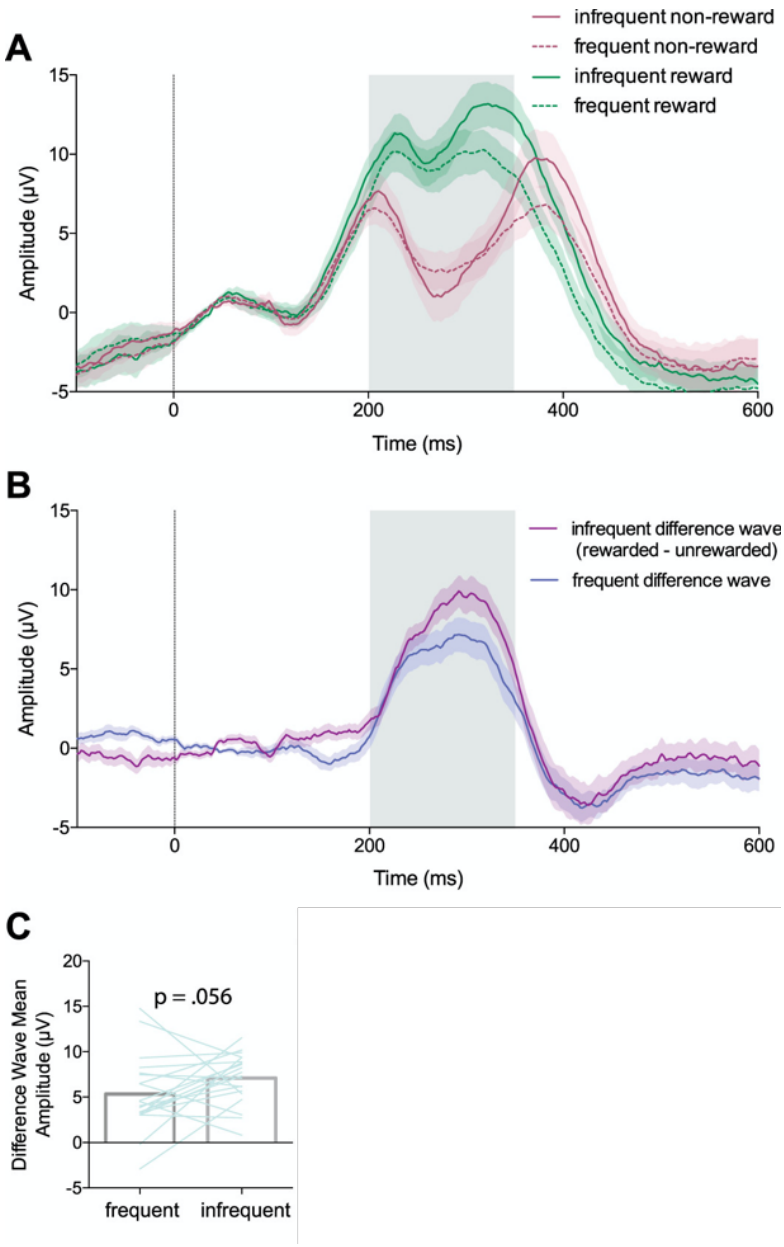

Figure 2-5: **Fronto-central event related potential responses to reinforcement feedback**. ***A***, Trial-averaged event-related potentials (ERPs) recorded from electrode FCz aligned to feedback presentation (0 ms, vertical blue line). Frequent and infrequent reward refer to reward feedback in the high- and low-reward frequency conditions, respectively. Frequent and infrequent nonreward refer to nonreward feedback in the low- and high-reward frequency conditions, respectively. Shaded regions: ±SE (n = 20). The gray shaded box indicates the temporal window of the feedback-related negativity/ reward positivity. ***B***, The difference waves (reward ERP – nonreward ERP) for frequent and infrequent feedback aligned to feedback presentation. ***C***: The mean amplitude of the difference wave (reward ERP – nonreward ERP) between 200 and 350 ms for infrequent and frequent feedback.

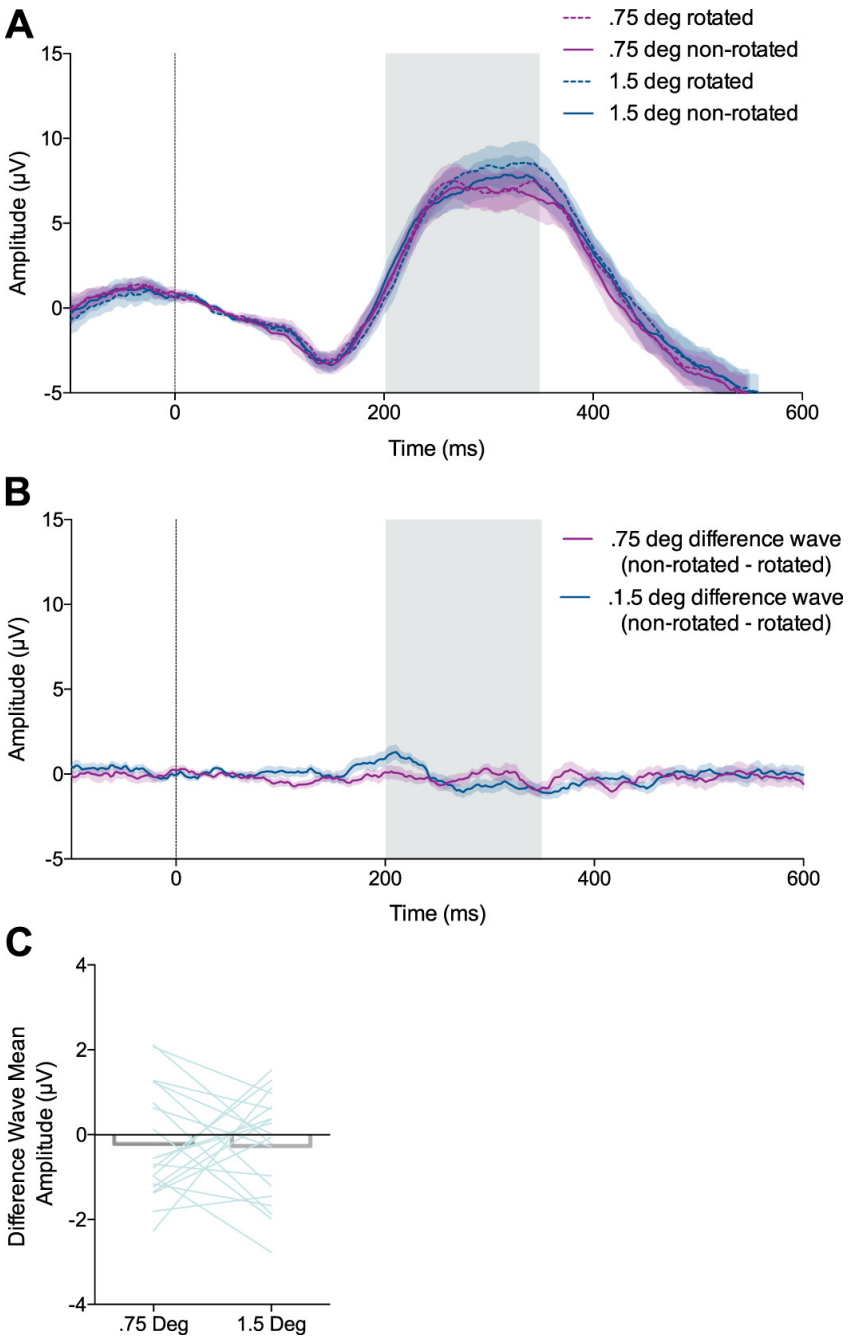

Figure 2-6: Fronto-central event related potential responses to sensory-error feedback. ***A*****,** Trial-averaged event-related potentials (ERPs) recorded from electrode FCz aligned to feedback presentation (0 ms, vertical line). Shaded regions: ±SE (n = 20). The gray shaded box indicates the temporal window of the FRN/RP. ***B*****,** The difference waves (nonrotated ERP − rotated ERP) for the 0.75° and 1.5° rotation conditions aligned to feedback presentation. ***C*****,** The mean amplitude of the difference wave (nonrotated ERP − rotated ERP) between 200 and 350 ms for the 0.75° and 1.5° rotation conditions. Error bars show ±SE.

that the FRN/RP reflects reward processing but not sensory error processing and that our experimental design successfully dissociated the two.

2.3.3 P300 Results

Reward learning task: Figure 2-7*A* shows ERPs recorded from electrode Pz during the reward learning condition, averaged across participants. We performed a 2 × 2 repeated-measures ANOVA on P300 amplitude with factors expectancy and reward. Figure 2-7*B* shows P300 amplitude for each condition, averaged across participants. We found a significant main effect of feedback expectancy [$F(1,19) = 97.16$, $P < 0.001$], indicating that P300 amplitude was significantly larger in the infrequent feedback conditions.

We also found a significant main effect of reward [$F(1,19) = 13.18$, $P = 0.002$], indicating that P300 amplitude was larger after rewarded trials compared with unrewarded trials. We found no reliable interaction between reward and expectancy [$F(1,19) = 0.992$, $P = 0.332$).

Visuomotor rotation task: Figure 2-8*A* shows ERPs recorded from electrode Pz during the VMR task, averaged across participants. We first tested for an effect of the VMR on P300 amplitude by comparing nonrotated feedback trials and rotated feedback trials. We performed a two-way repeated-measures ANOVA with factors presence of rotation and size of rotation (Fig. 2-8*B*). We did not find significant main effects of presence of rotation [$F(1,19) = 2.917$, $P = 0.104$]. We also did not find a main effect of size of rotation [$F(1,19) = 3.087$, $P = 0.095$]. We did find a significant interaction effect between presence of rotation and rotation magnitude [$F(1,19) = 8.728$, $P = 0.008$]. We performed planned pairwise comparisons using Bonferroni corrected t-tests between nonrotated and rotated conditions separately for the small- and large-error conditions. We found that P300 amplitude was significantly greater for rotated compared with nonrotated feedback in the 1.5° rotation condition [$t(19) = 2.83$, $P = 0.021$, Bonferroni corrected] but not the 0.75° rotation condition [$t(19) = 0.09$, $P = 0.93$].

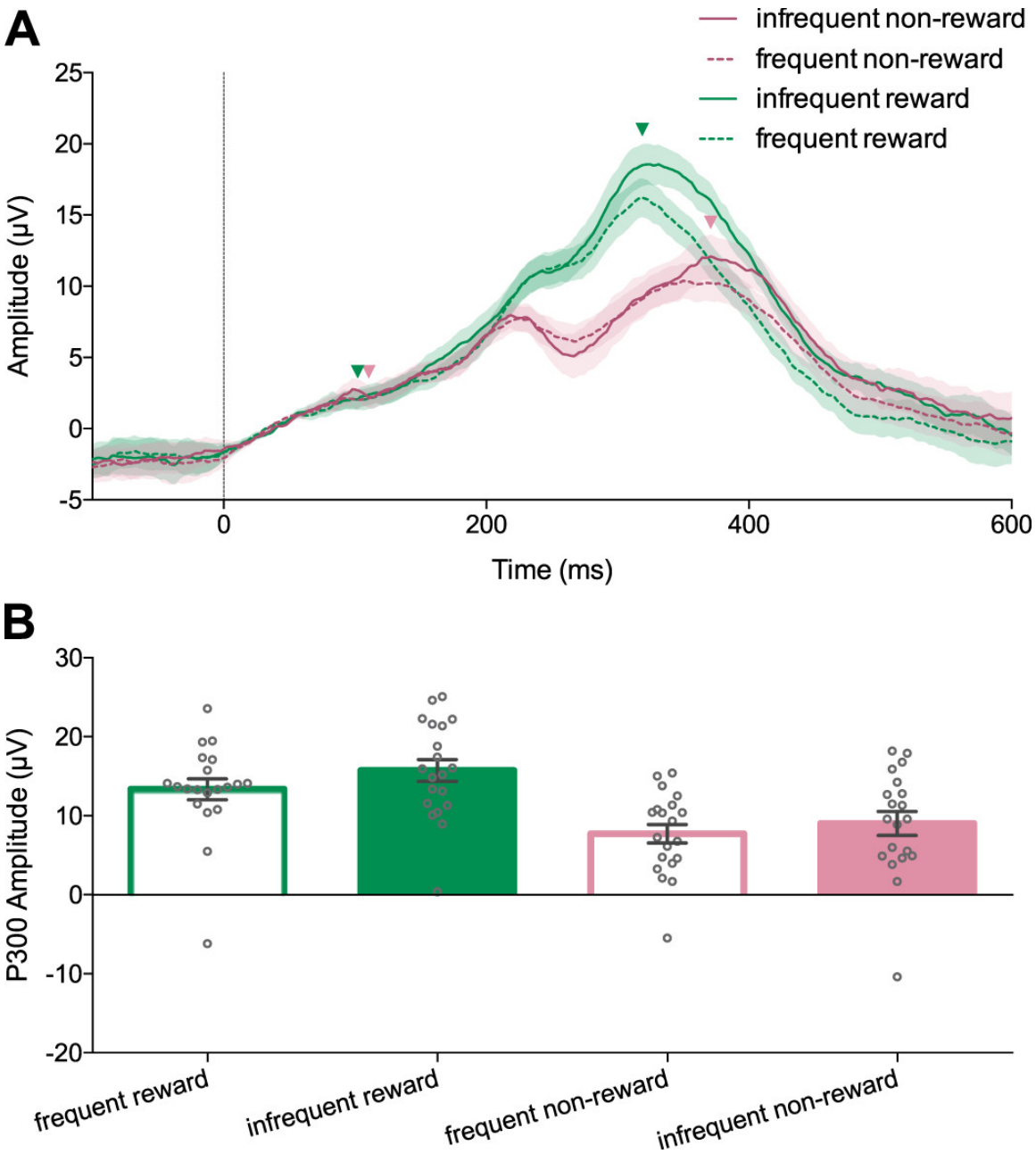

Figure 2-7: Centro-parietal event related potential responses to reinforcement feedback. ***A*****,** trial-averaged event-related potentials (ERPs) recorded from electrode Pz aligned to feedback presentation (0 ms, vertical line). Shaded regions: ±SE (n = 20). Arrowheads indicate the time points for the base and peak of the P300. ***B*****,** P300 amplitude in each feedback condition (error bars: ±SE). P300 amplitude is larger for rewarded feedback relative to unrewarded feedback and for infrequent feedback relative to frequent feedback.

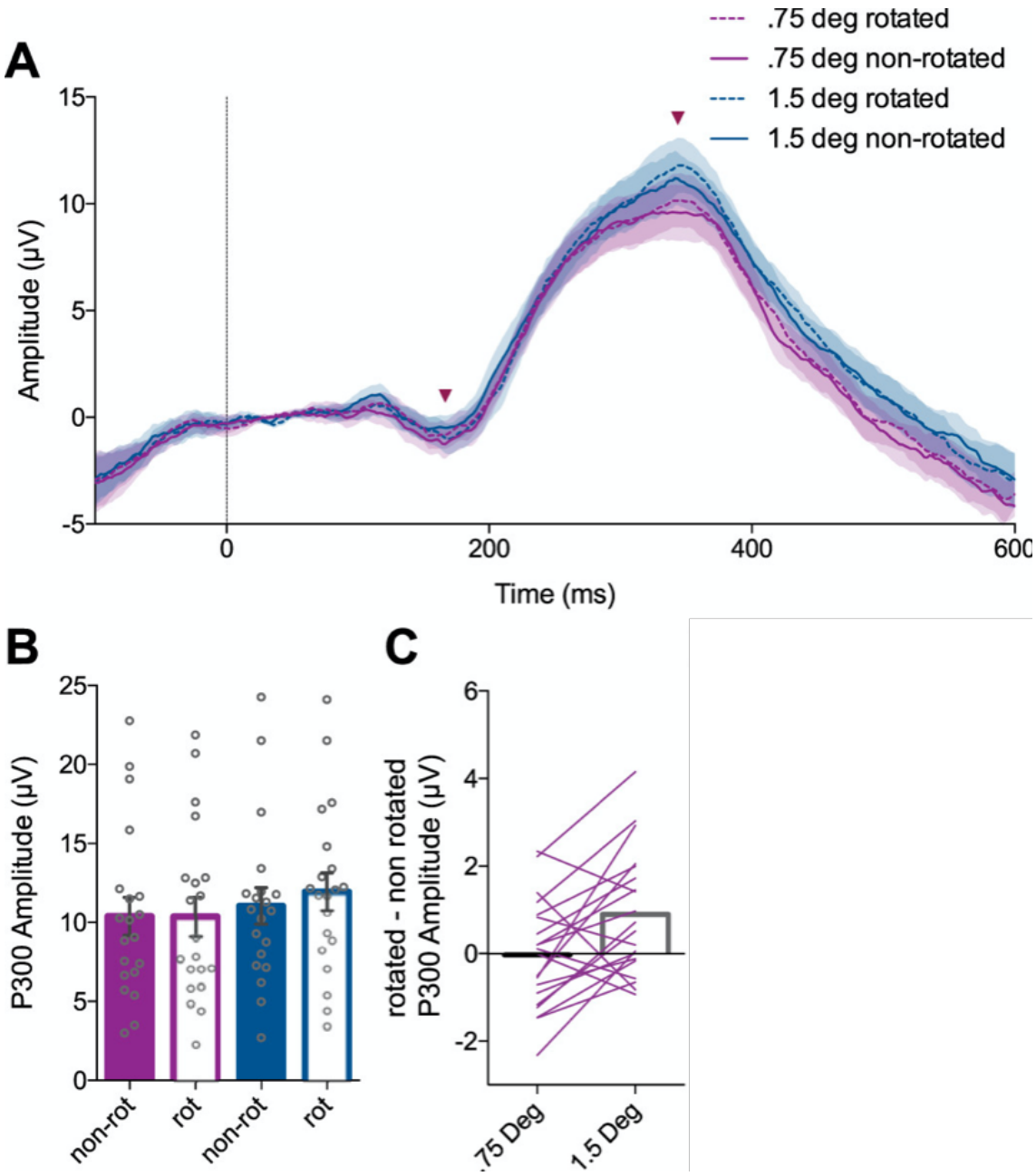

Figure 2-8: Centro-parietal event related potential responses to sensory error feedback. ***A*****,** trial-averaged event related potentials (ERPs) recorded from electrode Pz aligned to feedback presentation (0 ms, vertical line). Shaded regions: ±SE (n = 20). Arrowheads indicate the time points for the base and peak of the P300. ***B*****,** the peak-to-peak amplitude of the P300 during the visuomotor rotation task (error bars: ±SE). ***C*****,** P300 amplitude was larger for rotated than nonrotated trials in the 1.5° rotation condition but not the 0.75° rotation condition.

2.4 Discussion

We observed neural correlates of reward and sensory error feedback processing during motor adaptation. We used reaching tasks that were designed to isolate reward- and sensory error-based learning while producing comparable changes in reach angle. By examining ERPs elicited by feedback delivered at the end of movements we avoided potential confounds caused by neural activity or artifacts related to movement execution, motion of the limb, and online error correction. We observed that the FRN/RP was elicited by binary reward feedback but not by sensory error feedback. Our results support the hypothesis that the processes generating the FRN/RP signal reward prediction error to support reward-based motor learning, and that they are not necessary for sensory error-based learning. The P300 occurred in response to both reward and sensory error feedback, and P300 amplitude was modulated by VMR, reward, and surprise. These findings are consistent with the hypothesis that the P300 reflects a general processing of task-relevant feedback to support behavioral adaptation.

The FRN/RP reflects processing of reward feedback but not sensory error feedback:

Although motor adaptation has traditionally focused on sensory error-based learning, recent work suggests that reward-based learning processes can also contribute to motor adaptation. In the present study, we isolated reward-based learning from sensory error-based learning by providing only binary reward feedback in the absence of visual information indicating the position of the hand relative to the target. This feedback elicited a medial-frontal ERP component known as the FRN/RP.

The FRN/RP was observed in the reward learning task as a difference in voltage between ERPs elicited by nonreward feedback and those elicited by reward feedback. A large body of literature has shown that the FRN/RP is larger for infrequent outcomes than frequent outcomes (see Sambrook & Goslin, 2015, and Walsh & Anderson, 2012, for meta-analyses). Because less frequent outcomes should violate reward predictions more strongly, this finding is taken as support for the theory that the FRN/RP encodes a signed reward prediction error (Cohen & Ranganath, 2007; Eppinger et al., 2008; Holroyd &

Krigolson, 2007; Kreussel et al., 2012). In the present study, the FRN/RP was larger for improbable feedback than for probable feedback, which supports the hypothesis that the FRN/RP encodes a reward prediction error for reward-based learning. However, the statistical reliability of the difference was marginal ($P = 0.056$). This result is potentially due to the relatively small difference in reward frequency experienced between the low- and high-reward frequency conditions (38.6% and 67.0%, respectively) compared with other studies. We decided to avoid using very low or very high reward frequency as we found it to produce highly variable and strategic behavior in the task.

Although a prominent theory of the FRN/RP states that it reflects reward prediction error, other accounts suggest that it reflects error detection more generally or that it encodes salience responses to any unexpected outcomes (W. H. Alexander & Brown, 2011; Hauser, Iannaccone, Stämpfli, et al., 2014; Hird et al., 2018; Soder & Potts, 2018; Talmi et al., 2013; Ullsperger et al., 2014). In line with these ideas, recent work has identified the FRN/RP or the closely related error-related negativity in various motor learning and execution tasks involving sensory error signals. These studies either concluded that reinforcement- and sensory error-based learning processes share common neural resources or they simply did not distinguish between these two processes (Krigolson et al., 2008; MacLean et al., 2015; Savoie et al., 2018; Torrecillos et al., 2014). However, meeting the goals of a task may be rewarding and perturbations that cause sensory-prediction error may disrupt task performance. Thus, sensory prediction error is likely confounded with reward prediction error in most tasks. We designed tasks to dissociate sensory prediction error and reward prediction error, and we found that the FRN/RP was not elicited by sensory error when it did not disrupt task performance. These results support the hypothesis that the FNR/RP specifically reflects reward prediction error during motor adaptation.

The P300 is modulated by sensory error, surprise, and reward:

During the VMR task in the present study, we observed a P300 ERP component in response to reach end-point position feedback, and we found that P300 amplitude was

sensitive to the magnitude of sensory error. P300 amplitude was increased by the larger but not the smaller VMR. These findings suggest that the P300 observed in this task might reflect neural activity that is related to processing of sensory error underlying motor adaptation. It has previously been demonstrated that P300 amplitude decreases along with the magnitude of reach errors during the course of prism adaptation (MacLean et al., 2015). The P300 has also been shown to occur in response to target errors caused by random shifts in target location during reaching (Krigolson et al., 2008).

In the reward learning task, P300 amplitude was larger for reward feedback than nonreward feedback and for infrequent outcomes regardless of valence. Previous findings from cognitive reward-based learning tasks have reported similar effects of outcome valence and frequency (Hajcak et al., 2005, 2007; Leng & Zhou, 2010; Pfabigan et al., 2011; Y. Wu & Zhou, 2009; Zhou et al., 2010). The combined effects of sensory error, outcome valence, and outcome frequency suggest that the P300 amplitude is increased by any feature of feedback that promotes behavioral adaptation. This is consistent with the hypothesis that the P300 reflects activity in diffuse cortical areas that facilitates adaptive behavioral adjustments in response to salient stimuli, likely mediated by norepinephrine release (Dayan & Yu, 2006; Nieuwenhuis et al., 2005; San Martín, 2012; Vazey et al., 2018).

2.5 References

Alexander, W. H., & Brown, J. W. (2011). Medial prefrontal cortex as an action-outcome predictor. *Nature Neuroscience*, *14*(10), 1338–1344. https://doi.org/10.1038/nn.2921

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

Becker, M. P. I., Nitsch, A. M., Miltner, W. H. R., & Straube, T. (2014). A single-trial estimation of the feedback-related negativity and its relation to BOLD responses in a time-estimation task. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *34*(8), 3005–3012. https://doi.org/10.1523/JNEUROSCI.3684-13.2014

Block, H. J., & Bastian, A. J. (2012). Cerebellar involvement in motor but not sensory adaptation. *Neuropsychologia*, *50*(8), 1766–1775. https://doi.org/10.1016/j.neuropsychologia.2012.03.034

Brooks, J. L., Zoumpoulaki, A., & Bowman, H. (2017). Data-driven region-of-interest selection without inflating Type I error rate. *Psychophysiology*, *54*(1), 100–113. https://doi.org/10.1111/psyp.12682

Carlson, J. M., Foti, D., Mujica-Parodi, L. R., Harmon-Jones, E., & Hajcak, G. (2011). Ventral striatal and medial prefrontal BOLD activation is correlated with reward-related electrocortical activity: A combined ERP and fMRI study. *NeuroImage*, *57*(4), 1608–1616. https://doi.org/10.1016/j.neuroimage.2011.05.037

Cashaback, J. G. A., McGregor, H. R., Mohatarem, A., & Gribble, P. L. (2017). Dissociating error-based and reinforcement-based loss functions during sensorimotor learning. *PLoS Computational Biology*, *13*(7), e1005623. https://doi.org/10.1371/journal.pcbi.1005623

Codol, O., Holland, P. J., & Galea, J. M. (2018). The relationship between reinforcement and explicit control during visuomotor adaptation. *Scientific Reports*, *8*(1), 9121. https://doi.org/10.1038/s41598-018-27378-1

Cohen, M. X., & Ranganath, C. (2007). Reinforcement Learning Signals Predict Future Decisions. *Journal of Neuroscience*, *27*(2), 371–378. https://doi.org/10.1523/JNEUROSCI.4421-06.2007

Dayan, P., & Yu, A. J. (2006). Phasic norepinephrine: A neural interrupt signal for unexpected events. *Network: Computation in Neural Systems*, *17*(4), 335–350. https://doi.org/10.1080/09548980601004024

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Dhawale, A. K., Miyamoto, Y. R., Smith, M. A., & Ölveczky, B. P. (2019). Adaptive Regulation of Motor Variability. *Current Biology*, *29*(21), 3551-3562.e7. https://doi.org/10.1016/j.cub.2019.08.052

Diedrichsen, J., Hashambhoy, Y., Rane, T., & Shadmehr, R. (2005). Neural Correlates of Reach Errors. *Journal of Neuroscience*, *25*(43), 9919–9931. https://doi.org/10.1523/JNEUROSCI.1874-05.2005

Donchin, E., & Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences*, *11*(3), 357–374. https://doi.org/10.1017/S0140525X00058027

Eppinger, B., Kray, J., Mock, B., & Mecklinger, A. (2008). Better or worse than expected? Aging, learning, and the ERN. *Neuropsychologia*, *46*(2), 521–539. https://doi.org/10.1016/j.neuropsychologia.2007.09.001

Fabiani, M., Gratton, G., Karis, D., & Donchin, E. (1987). Definition, identification, and reliability of measurement of the P300 component of the event-related brain potential. *Adv Psychophysiol*, *2*, 1–78.

Falkenstein, M., Hohnsbein, J., Hoormann, J., & Blanke, L. (1991). Effects of crossmodal divided attention on late ERP components. II. Error processing in choice reaction tasks. *Electroencephalography and Clinical Neurophysiology*, *78*(6), 447–455. https://doi.org/10.1016/0013-4694(91)90062-9

Folstein, J. R., & Petten, C. V. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology*, *45*(1), 152–170. https://doi.org/10.1111/j.1469-8986.2007.00602.x

Galea, J. M., Mallia, E., Rothwell, J., & Diedrichsen, J. (2015). The dissociable effects of punishment and reward on motor learning. *Nature Neuroscience*, *18*(4), 597–602. https://doi.org/10.1038/nn.3956

Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., & Donchin, E. (1993). A Neural System for Error Detection and Compensation. *Psychological Science*, *4*(6), 385–390.

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

Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2007). It's worse than you thought: The feedback negativity and violations of reward prediction in gambling tasks. *Psychophysiology*, *44*(6), 905–912. https://doi.org/10.1111/j.1469-8986.2007.00567.x

Hauser, T. U., Iannaccone, R., Stämpfli, P., Drechsler, R., Brandeis, D., Walitza, S., & Brem, S. (2014). The feedback-related negativity (FRN) revisited: New insights into the localization, meaning and network organization. *NeuroImage*, *84*, 159–168. https://doi.org/10.1016/j.neuroimage.2013.08.028

Hird, E. J., El-Deredy, W., Jones, A., & Talmi, D. (2018). Temporal dissociation of salience and prediction error responses to appetitive and aversive taste. *Psychophysiology*, *55*(2), e12976. https://doi.org/10.1111/psyp.12976

Holland, P., Codol, O., & Galea, J. M. (2018). Contribution of explicit processes to reinforcement-based motor learning. *Journal of Neurophysiology*, *119*(6), 2241–2255. https://doi.org/10.1152/jn.00901.2017

Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*(4), 679–709. https://doi.org/10.1037/0033-295X.109.4.679

Holroyd, C. B., & Krigolson, O. E. (2007). Reward prediction error signals associated with a modified time estimation task. *Psychophysiology*, *44*(6), 913–917. https://doi.org/10.1111/j.1469-8986.2007.00561.x

Huang, V. S., Haith, A., Mazzoni, P., & Krakauer, J. W. (2011). Rethinking Motor Learning and Savings in Adaptation Paradigms: Model-Free Memory for Successful Actions Combines with Internal Models. *Neuron*, *70*(4), 787–801. https://doi.org/10.1016/j.neuron.2011.04.012

Inoue, K., Kawashima, R., Satoh, K., Kinomura, S., Sugiura, M., Goto, R., Ito, M., & Fukuda, H. (2000). A PET study of visuomotor learning under optical rotation. *NeuroImage*, *11*(5 Pt 1), 505–516. https://doi.org/10.1006/nimg.2000.0554

Izawa, J., & Shadmehr, R. (2011). Learning from Sensory and Reward Prediction Errors during Motor Adaptation. *PLOS Computational Biology*, *7*(3), e1002012. https://doi.org/10.1371/journal.pcbi.1002012

Kooij, K. van der, & Smeets, J. B. J. (2019). Reward-based motor adaptation can generalize across actions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(1), 71–81. https://doi.org/10.1037/xlm0000573

Krakauer, J. W. (2009). Motor Learning and Consolidation: The Case of Visuomotor Rotation. In D. Sternad (Ed.), *Progress in Motor Control: A Multidisciplinary*

Perspective (pp. 405–421). Springer US. https://doi.org/10.1007/978-0-387-77064-2_21

Krakauer, J. W., Ghez, C., & Ghilardi, M. F. (2005). Adaptation to Visuomotor Transformations: Consolidation, Interference, and Forgetting. *Journal of Neuroscience*, *25*(2), 473–478. https://doi.org/10.1523/JNEUROSCI.4218-04.2005

Kreussel, L., Hewig, J., Kretschmer, N., Hecht, H., Coles, M. G. H., & Miltner, W. H. R. (2012). The influence of the magnitude, probability, and valence of potential wins and losses on the amplitude of the feedback negativity. *Psychophysiology*, *49*(2), 207–219. https://doi.org/10.1111/j.1469-8986.2011.01291.x

Krigolson, O. E., Holroyd, C. B., Van Gyn, G., & Heath, M. (2008). Electroencephalographic correlates of target and outcome errors. *Experimental Brain Research*, *190*(4), 401–411. https://doi.org/10.1007/s00221-008-1482-x

Leng, Y., & Zhou, X. (2010). Modulation of the brain activity in outcome evaluation by interpersonal relationship: An ERP study. *Neuropsychologia*, *48*(2), 448–455. https://doi.org/10.1016/j.neuropsychologia.2009.10.002

MacLean, S. J., Hassall, C. D., Ishigami, Y., Krigolson, O. E., & Eskes, G. A. (2015). Using brain potentials to understand prism adaptation: The error-related negativity and the P300. *Frontiers in Human Neuroscience*, *9*. https://doi.org/10.3389/fnhum.2015.00335

Mazzoni, P., & Krakauer, J. W. (2006). An Implicit Plan Overrides an Explicit Strategy during Visuomotor Adaptation. *Journal of Neuroscience*, *26*(14), 3642–3645. https://doi.org/10.1523/JNEUROSCI.5317-05.2006

Miltner, W. H., Braun, C. H., & Coles, M. G. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: Evidence for a "generic" neural system for error detection. *Journal of Cognitive Neuroscience*, *9*(6), 788–798. https://doi.org/10.1162/jocn.1997.9.6.788

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

Nikooyan, A. A., & Ahmed, A. A. (2014). Reward feedback accelerates motor learning. *Journal of Neurophysiology*, *113*(2), 633–646. https://doi.org/10.1152/jn.00032.2014

Pfabigan, D. M., Alexopoulos, J., Bauer, H., & Sailer, U. (2011). Manipulation of feedback expectancy and valence induces negative and positive reward prediction error signals manifest in event-related brain potentials. *Psychophysiology*, *48*(5), 656–664. https://doi.org/10.1111/j.1469-8986.2010.01136.x

Polich, J. (2007). Updating P300: An Integrative Theory of P3a and P3b. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, *118*(10), 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019

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

Rand, M. K., & Heuer, H. (2019). Visual and proprioceptive recalibrations after exposure to a visuomotor rotation. *European Journal of Neuroscience*, *50*(8), 3296–3310. https://doi.org/10.1111/ejn.14433

Sambrook, T. D., & Goslin, J. (2015). A neural reward prediction error revealed by a meta-analysis of ERPs using great grand averages. *Psychological Bulletin*, *141*(1), 213–235. http://dx.doi.org.proxy1.lib.uwo.ca/10.1037/bul0000006

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

Savoie, F.-A., Thénault, F., Whittingstall, K., & Bernier, P.-M. (2018). Visuomotor Prediction Errors Modulate EEG Activity Over Parietal Cortex. *Scientific Reports*, *8*(1), 12513. https://doi.org/10.1038/s41598-018-30609-0

Shmuelof, L., Huang, V. S., Haith, A. M., Delnicki, R. J., Mazzoni, P., & Krakauer, J. W. (2012). Overcoming motor "forgetting" through reinforcement of learned actions. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *32*(42), 14617–14621. https://doi.org/10.1523/JNEUROSCI.2184-12.2012

Smith, M. A., & Shadmehr, R. (2005). Intact Ability to Learn Internal Models of Arm Dynamics in Huntington's Disease But Not Cerebellar Degeneration. *Journal of Neurophysiology*, *93*(5), 2809–2821. https://doi.org/10.1152/jn.00943.2004

Soder, H. E., & Potts, G. F. (2018). Medial frontal cortex response to unexpected motivationally salient outcomes. *International Journal of Psychophysiology*, *132*, 268–276. https://doi.org/10.1016/j.ijpsycho.2017.11.003

Synofzik, M., Lindner, A., & Thier, P. (2008). The cerebellum updates predictions about the visual consequences of one's behavior. *Current Biology: CB*, *18*(11), 814–818. https://doi.org/10.1016/j.cub.2008.04.071

Talmi, D., Atkinson, R., & El-Deredy, W. (2013). The Feedback-Related Negativity Signals Salience Prediction Errors, Not Reward Prediction Errors. *Journal of Neuroscience*, *33*(19), 8264–8269. https://doi.org/10.1523/JNEUROSCI.5695-12.2013

Tan, H., Jenkinson, N., & Brown, P. (2014). Dynamic Neural Correlates of Motor Error Monitoring and Adaptation during Trial-to-Trial Learning. *Journal of Neuroscience*, *34*(16), 5678–5688. https://doi.org/10.1523/JNEUROSCI.4739-13.2014

Tanaka, H., Sejnowski, T. J., & Krakauer, J. W. (2009). Adaptation to Visuomotor Rotation Through Interaction Between Posterior Parietal and Motor Cortical Areas. *Journal of Neurophysiology*, *102*(5), 2921–2932. https://doi.org/10.1152/jn.90834.2008

Therrien, A. S., Wolpert, D. M., & Bastian, A. J. (2016). Effective reinforcement learning following cerebellar damage requires a balance between exploration and motor noise. *Brain*, *139*(1), 101–114. https://doi.org/10.1093/brain/awv329

Thoroughman, K. A., & Shadmehr, R. (2000). Learning of action through adaptive combination of motor primitives. *Nature*, *407*(6805), 742–747. https://doi.org/10.1038/35037588

Torrecillos, F., Albouy, P., Brochier, T., & Malfait, N. (2014). Does the processing of sensory and reward-prediction errors involve common neural resources? Evidence from a frontocentral negative potential modulated by movement execution errors. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *34*(14), 4845–4856. https://doi.org/10.1523/JNEUROSCI.4390-13.2014

Ullsperger, M., Danielmeier, C., & Jocham, G. (2014). Neurophysiology of performance monitoring and adaptive behavior. *Physiological Reviews*, *94*(1), 35–79. https://doi.org/10.1152/physrev.00041.2012

Vazey, E. M., Moorman, D. E., & Aston-Jones, G. (2018). Phasic locus coeruleus activity regulates cortical encoding of salience information. *Proceedings of the National Academy of Sciences*, *115*(40), E9439–E9448. https://doi.org/10.1073/pnas.1803716115

Walsh, M. M., & Anderson, J. R. (2012). Learning from experience: Event-related potential correlates of reward processing, neural adaptation, and behavioral choice. *Neuroscience and Biobehavioral Reviews*, *36*(8), 1870–1884. https://doi.org/10.1016/j.neubiorev.2012.05.008

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

Zhou, Z., Yu, R., & Zhou, X. (2010). To do or not to do? Action enlarges the FRN and P300 effects in outcome evaluation. *Neuropsychologia*, *48*(12), 3606–3613. https://doi.org/10.1016/j.neuropsychologia.2010.08.010

Chapter 3

Effects of levodopa on neural and behavioral responses to feedback during motor adaptation

3.1 Introduction

Human motor control is adaptive to changes of the environment and the body through multiple mechanisms. Two important processes for adaptation are reward-based learning through reinforcement of successful actions and error-based learning through recalibration of internal mappings between motor commands and sensory outcomes (Huang et al., 2011; Izawa & Shadmehr, 2011; Taylor et al., 2014; Wolpert et al., 1995). Chapter 2 describes experimental findings supporting the hypothesis that an event related potential (ERP) component called the feedback-related negativity/ reward positivity (FRN/RP) encodes reward-prediction error signals during reward-based motor adaptation. These findings also demonstrated that the process generating the FRN/RP is specific to reward-based learning, as these signals were not elicited by feedback conveying sensory error information. In theories of biological reinforcement learning, phasic dopamine release within fronto-striatal circuits is thought modulate synaptic plasticity underlying reward-based learning through signaling of reward prediction error (Izawa & Shadmehr, 2011; Sutton & Barto, 2011; Watabe-Uchida et al., 2017). The role of dopamine in reward-based learning has mostly been studied in tasks involving the selection of discrete stimuli or actions, and it is not yet known whether motor adaptation is also mediated by dopaminergic learning mechanisms. Here, we test the hypothesis that dopamine release contributes to reward-based motor adaptation through signaling of reward-prediction error.

Most research into motor adaptation to date has focused on sensory-error based learning, for which visuomotor rotation (VMR) tasks are a common experimental model (Inoue et al., 2000; Krakauer, 2009; Krakauer et al., 2005; Rand & Heuer, 2019; H. Tanaka et al., 2009). In VMR tasks, a cursor on a digital display represents the position of the hand, and the mapping between the actual reach angle and the position of the cursor is rotated about the movement start position. Participants reliably compensate for a proportion of VMR perturbations following even single movements. Learning involves the cerebellum as well as parietal, sensory, and motor cortical areas (Diedrichsen et al., 2005; M. Ito, 2000; Krakauer et al., 2004; M. A. Smith & Shadmehr, 2005; H. Tanaka et al., 2009; Taylor

et al., 2010; Wong et al., 2019). It is thought that these neural circuits predict the sensory consequences of motor commands, and that adaptation occurs in response to sensory prediction error when sensory afference violates these predictions (Adams et al., 2013; Bhanpuri et al., 2013; Izawa & Shadmehr, 2011; Miall et al., 2007; Shadmehr et al., 2010; Synofzik et al., 2008; Therrien & Bastian, 2015; Tseng et al., 2007; Wolpert et al., 1995).

While sensory error-based learning mechanisms are dominant in typical motor adaptation paradigms, reward-based learning processes can also contribute to motor adaptation (Bernardi et al., 2015; Cashaback et al., 2019; Izawa & Shadmehr, 2011; Kim et al., 2019; Kooij et al., 2018; Kooij & Smeets, 2019; McDougle et al., 2016; Mehler et al., 2017; Nikooyan & Ahmed, 2014; Palidis et al., 2019; Sidarta et al., 2016, 2018). When sensory error-based learning cannot occur due to limited sensory feedback or cerebellar damage, reward-based learning can produce comparable behavioral adaptation (Cashaback et al., 2017; Izawa & Shadmehr, 2011; Therrien et al., 2016). Reward-based adaptation can be isolated experimentally in tasks analogous to VMR paradigms in which cursor feedback indicating the position of the hand is not provided, and instead only binary reinforcement feedback signals whether a target movement is successfully produced (Izawa & Shadmehr, 2011; Shmuelof et al., 2012).

Levodopa is a dopamine precursor commonly used to treat motor symptoms in patients with Parkinson's disease. In cognitive tasks, levodopa has been shown to impair reward-based learning in both patients and healthy participants (Cools et al., 2001, 2007; Feigin et al., 2003; Frank et al., 2004; Graef et al., 2010; Hiebert et al., 2014; Jahanshahi et al., 2010; Kwak et al., 2010; MacDonald et al., 2011; Swainson et al., 2000; Torta et al., 2009; Vo et al., 2016, 2018). According to the "dopamine overdose" hypothesis, dopamine levels affect performance in tasks that depend on the ventral striatum according to an inverted-u function (Cools et al., 2001). In early-stage Parkinson's disease, the dorsal striatum is significantly depleted of dopamine whereas the ventral striatum is comparatively spared. Dopaminergic therapy is thus predicted to ameliorate deficits caused by dopamine-depletion in the dorsal striatum but to worsen functions ascribed to the ventral striatum by overstimulating dopamine release in this region. In line with this view, reward-based

learning is thought to rely on dopamine signaling in ventral striatum and is impaired by levodopa (Hiebert et al., 2019).

Although dopamine is widely implicated in cognitive forms of reward-based learning, it is not clear whether this role of dopamine extends to reward-based motor adaptation. We administered levodopa to healthy young participants to test for effects on motor adaptation. Participants received levodopa and placebo in separate sessions using a repeated measures design. Both sessions included a reward-based learning task and a sensory error-based VMR task. In the reward-based learning task, adaptation was induced through binary reinforcement feedback at the end of each movement. We measured changes in the mean reach angle due to reinforcement as well as modulations in trial-by-trial variability of reach angle as a response to reward outcomes. Previous research has shown that motor variability increases following unrewarded outcomes compared to rewarded outcomes (Dhawale et al., 2019; Holland et al., 2018; Kooij & Smeets, 2019; Mastrigt et al., 2020; Pekny et al., 2015). This could indicate reinforcement of rewarded actions as well as exploration in response to unrewarded outcomes (Cashaback et al., 2019; Dhawale et al., 2019). This variance modulation is impaired in individuals with Parkinson's disease who are medicated, but it remains unclear whether this deficit is caused by the disease process itself or by side-effects of dopaminergic medication (Pekny et al., 2015).

It is thought that reward prediction error drives biological reinforcement learning when an action results in an outcome that is better or worse than expected (Daw & Tobler, 2014; Sambrook & Goslin, 2015; Schultz, 2016; Sutton & Barto, 2011; Walsh & Anderson, 2012). Phasic changes in the firing rate of midbrain dopamine neurons match reward prediction error signals predicted by computational models of reinforcement learning (Bayer & Glimcher, 2005; García-García et al., 2017; Jocham & Ullsperger, 2009; Schultz et al., 1997; Watabe-Uchida et al., 2017). Dopaminergic projections to the striatum and frontal cortex are thought to mediate synaptic plasticity in these regions underlying reward-based learning (Otani et al., 2003; Reynolds & Wickens, 2002; J. X. Wang et al., 2018). Reward-prediction error signals are consistently observed throughout the medial

frontal cortex and striatum (Diederen et al., 2017; Oyama et al., 2010; Pagnoni et al., 2002; Rutledge et al., 2010; Schultz et al., 1998; Silvetti et al., 2014; Watabe-Uchida et al., 2017).

We recorded EEG to measure the FRN/RP ERP component in response to reinforcement feedback. A prominent theory of the FRN/RP states that it reflects reward prediction error signals in the medial frontal cortex driven by dopamine release (Becker et al., 2014; Carlson et al., 2011; Emeric et al., 2008; Foti et al., 2011; Gehring & Willoughby, 2002; Hauser, Iannaccone, Stämpfli, et al., 2014; Holroyd et al., 2008; Holroyd & Coles, 2002; Mathewson et al., 2008; Miltner et al., 1997; Sambrook & Goslin, 2015, 2016; Vezoli & Procyk, 2009; Walsh & Anderson, 2012; Warren et al., 2015). However, direct evidence for a link between dopamine and the FRN/RP is fairly limited, and no studies have investigated this link in the context of motor adaptation (Enge et al., 2017; Forster et al., 2017; Marco-Pallarés et al., 2009; Mueller et al., 2014; Santesso et al., 2009; Schutte et al., 2020).

We tested the hypothesis that dopaminergic signaling of reward-prediction error mediates reward-based motor adaptation. In accordance with the "dopamine overdose hypothesis", we predicted that levodopa would diminish the magnitude of the FRN/RP, a measure of neural reward prediction error signaling. We predicted that levodopa-induced reductions in FRN/RP magnitude would correspond with impairments in both reward-based motor adaptation and modulation of trial-by-trial variability in response to reward feedback. We included a sensory error-based learning task to test the specificity of the role of dopamine in reward-based motor adaptation. We predicted that levodopa would not affect the neural or behavioral responses to sensory error.

3.2 Materials and methods

3.2.1 Participants

Data from a total of n=21 [12 female, Age: 20.99 years (SD 3.26)] healthy, right-handed participants were analyzed and reported. Two additional participants were excluded from analysis due to malfunction of the robot that prevented the experiment from being

completed, and two participants were excluded who did not return for the second testing session. All participants were screened for neurological and psychiatric illness, history of drug or alcohol abuse, and contraindications for levodopa. Participants provided written informed consent to experimental procedures approved by the Research Ethics Board at Western University.

3.2.2 Experimental design

Drug administration: All participants underwent two experimental sessions, with levodopa and placebo being administered in separate sessions using a randomized, double-blind, crossover design. The two sessions were separated by a washout period of at least one week. In one session, a capsule was ingested that contained 100 mg of levodopa (L-3,4-dihydroxyphenylalanine) and 25 mg of carbidopa. Levodopa is a dopamine precursor, and carbidopa is a decarboxylase inhibitor given to reduce conversion of levodopa to dopamine in the periphery. This dose has been shown to produce various behavioral effects in healthy young adults (Flöel et al., 2005; Knecht et al., 2004; Onur et al., 2011; Vo et al., 2016, 2017, 2018). In the other session, an equal volume of placebo was administered in an identical capsule. The order of administration was counterbalanced. After administration of the capsule, the robot was calibrated, the EEG cap was placed on the participant's head, and participants performed a practice block of the behavioral task. Subsequently, the experimental tasks began 45 minutes after ingestion of the capsule to coincide with peak plasma levels of levodopa (Olanow et al., 2000). We measured heart rate, blood pressure, and subjective alertness immediately prior to ingestion of placebo or levodopa and again at the end of each session. Alertness was assessed using the Bond-Lader visual analog scale (Bond & Lader, 1974).

Overview of behavioral tasks: Each participant underwent the same experimental tasks in both sessions. Participants made reaching movements toward a visual target and received visual feedback pertaining to reach angle only at movement end point (figure 3-1). Neural responses to feedback were recorded using EEG. Participants were instructed that each reach terminating within the target would be rewarded with a small monetary bonus.

Participants first performed a block of 50 practice trials. The subsequent behavioral procedure consisted of two blocks of a reward learning task and two blocks of a visuomotor rotation (VMR) task. The order of the blocks alternated between the two task types but was otherwise randomized. Participants took self-paced rests between blocks.

In the VMR task, a cursor appeared at movement end point to represent the position of the hand (Figure 3-1D). In unperturbed trials, the cursor was displayed directly over the occluded robot handle. In randomly selected trials, the cursor's position was decoupled from the robot handle position such that the cursor indicated a reach endpoint position that was rotated about the start position relative to the actual reach endpoint position. This was intended to produce sensory prediction error and trial-by-trial compensatory changes in reach angle opposite the direction of the rotations. The rotations were small relative to the size of the target, such that participants nearly always landed in the target, fulfilling the goal of the task and earning a monetary reward (the cursor feedback was within the target on 95.5% of trials, SD: 2%). Thus, reward and task error were constant between perturbed and unperturbed feedback, and by comparing the two conditions we could isolate the neural correlates of sensory error processing.

In the reward learning task, no cursor appeared to indicate the position of the hand. Instead, binary feedback represented whether or not participants succeeded in hitting the target (Figure 3-1C). This allowed us to assess reward-based learning in isolation from sensory error processing, as visual information revealing the position of the hand was not provided. In separate blocks, reward feedback was tailored to produce adaptation towards increasingly clockwise and counterclockwise reach angles. Reward was delivered when the difference between the current reach angle and the median of the previous 10 reach angles was in the direction of intended learning. We compared the neural responses to reward and non-reward feedback to assess the neural correlates of reward processing.

3.2.3 Apparatus/Behavioral Task

Participants produced reaching movements with their right arm while holding the handle of a robotic arm (InMotion2; Interactive Motion Technologies; figure 3-1). Position of

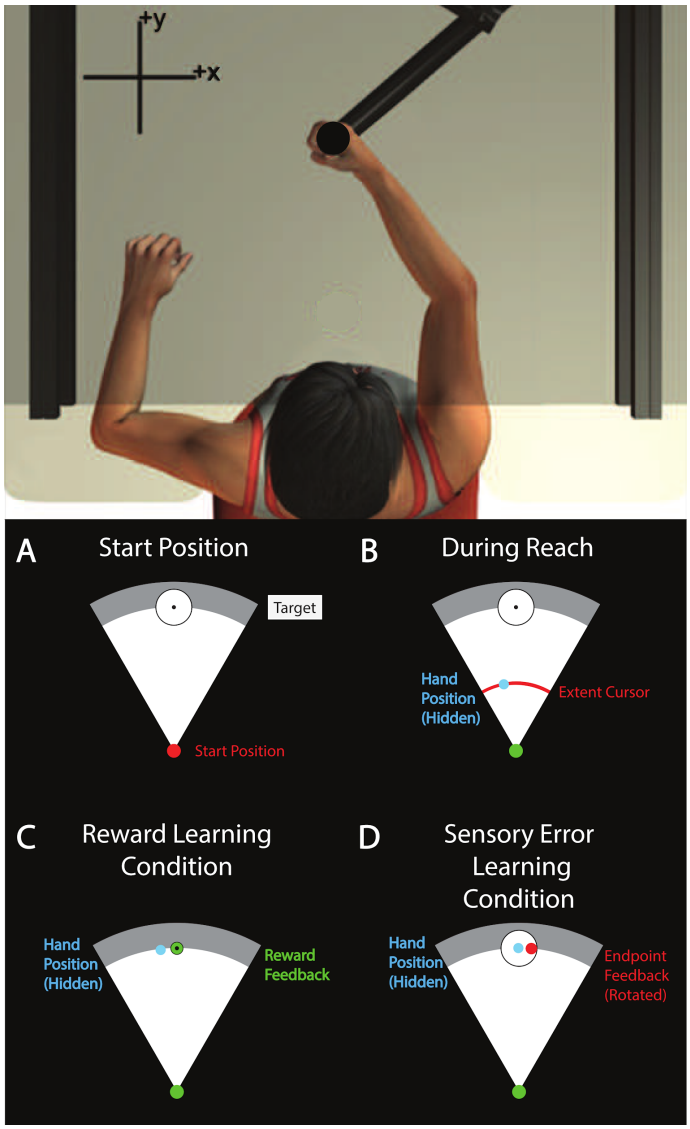

Figure 3-1: Experimental setup. Top: The experimental apparatus. Participants reached to visual targets while holding the handle of a robotic arm. Vision of the arm was obscured by a screen that displayed visual information related to the task. **Bottom:** Illustrations of the visual display. **A,** Participants made outward reaching movements from a start position at body midline to a visual target. **B,** During reaches, hand position was hidden but an arc-shaped cursor indicated the extent of the reach without revealing reach angle. Feedback was provided at reach end point. ***C,*** In the reward learning task, binary feedback represented whether reaches were successful or unsuccessful in hitting the target by turning green or red, respectively. Reach adaptation was induced by providing reward for movements that did not necessarily correspond to the visual target. ***D,*** In the visuomotor rotation task, cursor feedback represented the end-point position of the hand. Adaptation was induced by rotating the angle of the feedback relative to the actual reach angle.

the robot handle was sampled at 600 Hz. A semi-silvered mirror obscured vision of the arm and displayed visual information related to the task. An air sled supported each participant's right arm. Participants reached towards a white circular target 14 cm away from a circular start position in front of their chest. The start position turned from red to green to cue the onset of each reach once the handle had remained inside it continuously for 750 ms. Participants were instructed that they must wait for the cue to begin each reach but that it was not necessary to react quickly upon seeing the cue. Participants were instructed to make forward reaches and to stop their hand within the target. An arc-shaped cursor indicated reach extent throughout each movement without revealing reach angle. In only the first five baseline trials of each block, an additional circular cursor continuously indicated the position of the hand throughout the reach. A viscous force field assisted participants in braking their hand when the reach extent was 14 cm. The robot ended each movement by fixing the handle position when the hand velocity decreased below 0.03 m/s. The hand was fixed in place for 700 ms, during which time visual feedback of reach angle was provided. Feedback indicated either reach end point position, a binary reward outcome, or feedback of movement speed (see below). Visual feedback was then removed, and the robot guided the hand back to the start position. Reach end point was defined as the position at which the reach path intersected the perimeter of a circle (14-cm radius) centered at the start position. Reach angle was calculated as the angle between vectors defined by reach end point and the center of the target, each relative to the start position, such that reaching straight ahead corresponds to 0° and counterclockwise reach angles are positive.

Feedback about reach angle was provided either in the form of end-point position feedback or binary reward feedback. The type of feedback, as well as various feedback manipulations, varied according to the assigned experimental block type (see Reward Learning Task and Visuomotor Rotation Task). Participants were told that they would earn additional monetary compensation for reaches that ended within the target, up to a maximum of $10 CAD. Movement duration was defined as the time elapsed between the hand leaving the start position and the moment hand velocity dropped below 0.03 m/s. If

movement duration was >700 ms or <450 ms, no feedback pertaining to movement angle was provided. Instead, a gray arc behind the target turned blue or yellow to indicate that the reach was too slow or too fast, respectively. Participants were informed that movements with an incorrect speed would be repeated but would not otherwise affect the experiment. To minimize the impact of eyeblink-related EEG artifacts, participants were asked to fixate their gaze on a black circular target in the center of the reach target and to refrain from blinking throughout each arm movement and subsequent presentation of feedback.

Practice block: Each participant first completed a block of practice trials that continued until they achieved 50 movements within the desired range of movement duration. Continuous position feedback was provided during the first 5 trials, and only end-point position feedback was provided for the following 10 trials. Subsequently, no position feedback was provided outside the start position.

Reward Learning task: Binary reward feedback was provided to induce adaptation of reach angle (figure 3-1C). Each session included two blocks in the reward learning condition. The direction of intended learning was clockwise in one block and counterclockwise in the other. Each block continued until the participant completed 125 reaches with acceptable movement duration. Participants reached toward a circular target 1.2 cm in diameter. The first 11 reaches were baseline trials during which continuous position feedback was provided during the first 5 trials, followed by 6 trials with only end-point cursor feedback. After these baseline trials no cursor feedback was provided, and binary reward feedback was instead provided at the end of the movement. Target hits and misses were indicated by the target turning green and red, respectively. Unbeknownst to participants, reward feedback did not necessarily correspond to the visual target. Instead, reward was delivered if the difference between the current reach angle and the median angle of the previous 10 reaches was in the direction of intended learning. When the running median was at least 6° away from zero in the direction of intended learning, reward was delivered at a fixed probability of 50%. This was intended to minimize conscious awareness of the manipulation by limiting adaptation to 6°. Reward was never

delivered when the absolute value of the reach angle was greater than 10°, for the same reason. We employed this adaptive, closed-loop reward schedule so that the overall frequency of reward was controlled.

Visuomotor rotation task: End-point feedback was rotated relative to the actual reach angle to induce sensory error-based adaptation (figure 3-1D). Each session included two blocks in the VMR condition. Each block continued until participants completed 124 reaches within acceptable movement duration limits. Participants reached toward a circular target 3.5 cm in diameter. Participants first performed baseline reaches during which cursor feedback reflected veridical reach angle continuously for the first 5 trials and only at movement end point for the subsequent 5 trials. After the baseline reaches the adaptation portion of each block began, unannounced to participants. During the adaptation trials, end-point position feedback was provided indicating a reach angle that was rotated relative to the true reach angle. There were 114 total adaptation trials (38 with 0° rotation, and 19 each with ±2° and ±4° rotations). Participants were instructed that end-point feedback within the target would earn them bonus compensation, but no explicit reward feedback was provided.

3.2.4 EEG data acquisition

EEG data were acquired from 16 cap-mounted electrodes with an active electrode system (g.GAMMA; g.tec Medical Engineering) and amplifier (g.USBamp; g.tec Medical Engineering). We recorded from electrodes placed according to the 10-20 System at sites Fp1, Fp2, F3, F4, F7, F8, FT9, FT10, FCz, Cz, C3, C4, CPz, CP3, CP4, and Pz referenced to an electrode placed on participants' left earlobe. Impedances were maintained below 5 kΩ. Data were sampled at 4,800 Hz and filtered online with band-pass (0.1–1,000 Hz) and notch (60 Hz) filters. A photodiode attached to the display monitor was used to synchronize recordings to stimulus onset.

3.2.5 Behavioral data analysis

Reward learning task. We computed learning scores in each drug condition by subtracting the average reach angle in the clockwise condition from the average reach angle in the counterclockwise condition. As such, positive scores indicate learning. We excluded baseline trials and trials that did not meet the movement duration criteria, as no feedback related to reach angle was provided on these trials. Each block continued until 114 trials after the baseline period met the movement duration criteria, so equal numbers of trials were analyzed for each participant. We tested for the presence of learning by submitting learning scores to 1-sample T-Tests against zero, and we compared learning scores in the placebo and levodopa conditions using paired T-Tests.

We also analyzed trial-by-trial variability in reach angle in response to reinforcement feedback using an approach similar to Pekny et al. (2015). First, we calculated trial-by-trial changes in reach angle as in *Eq. 3.1:*

$$\Delta\theta_i = \theta_{i+1} - \theta_i \quad (3.1)$$

We then multiplied $\Delta\theta_i$ by -1 for trials in the clockwise learning condition, so that positive values for $\Delta\theta_i$ corresponded to changes in reach angle in the direction of intended learning, and any biases in $\Delta\theta$ related to the direction of intended learning would have the same sign in the CW and CCW learning conditions. Next, we conditioned $\Delta\theta_i$ on the reinforcement outcome of trial i and the drug condition to obtain trial-by-trial changes in reach angle following reward and non-reward after both placebo and levodopa administration. Next, we quantified trial by trial variability in each condition as the natural logarithm of the sample variance of $\Delta\theta_i$. Our dependent variable is an estimate of variance. This estimate of variance itself has variance due to sampling. For a normal distribution, the variance of a sample variance is proportional to the square of the true population variance. A log transformation is appropriate for linear modeling when the variance of the dependent measure is proportional to the square of its expectation (Montgomery et al., 2021). We then performed 2x2 repeated measures

ANOVA on Log(var($\Delta\theta_i$)). The factors were drug (levels: placebo, levodopa), and reward outcome on trial i (levels: non-reward, reward).

Visuomotor rotation task. To quantify trial-by-trial learning we first calculated the change in reach angle between successive trials, as in *Eq. 3.1.* We then performed a linear regression on $\Delta\theta_i$ with the rotation imposed on trial i as the predictor variable. The rotation was 0°, ±2°, or ±4°. This regression was performed on an individual participant basis, separately for placebo and levodopa conditions. We excluded trials that did not meet the duration criteria as no visual feedback was provided on these trials. We took the resulting slope estimates multiplied by -1 as a metric of learning rate for each participant, as it reflects the portion of visual errors that participants corrected with a trial-by-trial adaptive process. We tested for the presence of adaptation in each condition by submitting learning rates to 1-sample t-tests against zero. We tested for an effect of levodopa vs placebo on learning rates using a paired t-test.

3.2.6 EEG preprocessing

EEG data were resampled to 480 Hz and filtered off-line between 0.1 and 35 Hz with a second-order Butterworth filter. Continuous data were segmented into 2-s epochs time-locked to feedback stimulus onset at 0 ms (time range: -500 to +1,500 ms). Epochs flagged for containing artifacts as well as any channels with bad recordings were removed after visual inspection. One participant was excluded entirely from the EEG analysis due to excessive muscle artifacts. Subsequently, extended infomax independent component analysis was performed on each participant's data (Delorme & Makeig, 2004). Components reflecting eye movements and blink artifacts were identified by visual inspection and subtracted by projection of the remaining components back to the voltage time series.

3.2.7 EEG data analysis

After artifact removal, we computed ERPs by trial averaging EEG time series epochs for various feedback conditions described in the sections below. ERPs were computed on an

individual participant basis separately for recordings from channels FCz and Pz. We selected FCz and Pz a priori because these electrodes typically correspond to the peaks of the scalp distributions for the feedback related negativity/reward positivity and the P300 ERP components, respectively. We found this to be true in the data presented in chapter 2, which were collected using very similar experimental paradigms. All ERPs were baseline corrected by subtracting the average voltage in the 75-ms period immediately following stimulus onset. We used a baseline period following stimulus onset because stimuli were presented immediately upon movement termination and the period before stimulus presentation was more likely to be affected by movement related artifacts. Trials in which reaches did not meet the movement duration criteria were excluded, as feedback relevant to reach adaptation was not provided on these trials. Finally, ERPs were low-pass filtered with a cutoff frequency of 30 Hz.

We computed ERPs separately following administration of placebo and levodopa. In the reward learning task, we computed ERPs separately for feedback indicating non-reward (placebo: 107.2 ±9.7 trials, levodopa: 104.0 ±8.3 trials) and feedback indicating reward (placebo: 118.4 ±9.6 trials, levodopa: 118.0 ±8.1 trials). In the visuomotor rotation task, we computed ERPs separately for veridical endpoint feedback (placebo: 72.6 ± 3.5 trials, levodopa: 72.9 ± 3.1 trials), ±2° rotated feedback (placebo: 70.8 ± 5.2 trials, levodopa: 72.1 ± 3.8 trials), and ±4° rotated feedback (placebo: 64.5 ± 4.7 trials, levodopa: 66.3 ± 4.1 trials). We excluded trials in which the cursor did not land within the target.

We selected time windows of interest for ERP analysis using independent data from a previous experiment with very similar procedures, which were reported in chapter 2. We analyzed the amplitudes of FRN/RP and P300 components within 50 ms time windows centered around the latencies of the FRN/RP and P300 peaks observed in the previous study. The FNR/RP peak was taken as the maximum value of the difference between ERPs elicited by reward and non-reward feedback recorded from electrode FCz (latency: 292ms). For completeness, we used the same time window to test for FRN/RP effects in the visuomotor rotation task of the current study although we did not observe an FRN/RP component in our previous visuomotor rotation task. The P300 peak latencies

were determined separately for reward and non-reward feedback as the times of maximal amplitude of ERPs recorded from electrode Pz (reward: 319ms, non-reward: 371ms). The peak latencies selected for the FRN/RP and P300 components in the reward learning task corresponded very closely to the peaks observed in the current data. However, the P300 peak in the visuomotor rotation task of the current study was earlier than that in our previous experiment. This difference in latency may be due to changes in the nature of the feedback. Thus, we determined the latency of the P300 peak in the visuomotor rotation task of the current study using a data-driven method that does not bias comparisons between conditions (J. L. Brooks et al., 2017). We aggregated all trials across conditions and participants and computed a trial averaged ERP using recordings from electrode Pz. The P300 peak was determined as the maximal amplitude of this averaged waveform (latency: 317ms).

We tested for effects of feedback manipulations on FRN/RP components using the average amplitudes of ERPs recorded from electrode FCz within the FRN/RP time window. We tested for effects on P300 ERP components using the average amplitudes of ERPs recorded from electrode Pz within the P300 time window corresponding to a given condition. For the reward learning task, we used 2x2 repeated measures ANOVA with factors drug (levels: placebo, levodopa) and reinforcement outcome (levels: reward, non-reward). For the visuomotor rotation task, we used 2x3 repeated measures ANOVA with factors drug (levels: placebo, Levodopa), and rotation (levels: 0°, ±2°, ±4°).

3.2.8 Statistics

Statistical tests were implemented using JASP v0.14.1. We performed comparisons on sample means using 1 sample T-Tests, paired sample T-Tests, or independent sample T-Tests. These comparisons allowed us to compute one-tailed Bayes factors representing $p(data|H_+) \,/\, p(data|H_0)$, where H_0 represents the null hypothesis corresponding to the standard *t*-distribution for an effect size of 0, and H_+ represents the alternative hypothesis corresponding to a *t*-distribution constructed using a one-tailed prior distribution of effect sizes. The use of 1-tailed priors is recommended in the case of directional

hypotheses to provide "a fairer balance between the ability to provide evidence for H_0 and H_+" (Keysers et al., 2020). We used the default effect size priors implemented in JASP (Cauchy scale 0.707). These priors are generally appropriate for effect sizes typical of neuroscience research, and the use of default priors is recommended for standardized and objective analysis (Keysers et al., 2020; Rouder et al., 2012; Wetzels et al., 2011). Bayesian estimates of effect size are reported as median posterior Cohen's δ with 95% credibility interval using 2-tailed priors for H1 to avoid biasing the estimate in the expected direction. We also report T-statistics, p-values, and 95% confidence intervals generated using 2-tailed frequentist T-Tests. For factorial analyses, we conducted frequentist and Bayesian repeated measures ANOVAs using JASP with default priors. Bayes factors were computed for the inclusion of each effect as the ratio of the data likelihood under the model containing that effect vs equivalent models stripped of that effect. Bayes factors >3 and >10 were taken as moderate and strong evidence in favor of the alternative hypothesis, respectively. Bayes factors <1/3 and <1/10 were taken as moderate and strong evidence in favor of the null hypothesis, respectively. Bayes factors between 1/3 and 3 were taken as inconclusive evidence (Keysers et al., 2020).

Directional priors used for alternative hypotheses specified our predictions that learning metrics would be greater than zero (Reward learning score, VMR learning rate). In comparing placebo and levodopa conditions, our alternative hypotheses specified that learning metrics would be lower in levodopa conditions than placebo conditions, in accordance with the "dopamine overdose" hypothesis. All other Bayes factors are computed with 2-tailed priors, as they were conducted without directional a priori hypotheses (control measures, etc.).

3.3 Results

3.3.1 Control measures

Control measures: Participants' judgments at the end of the second session as to whether they received placebo or drug were correct at near chance level (47.62%). Table 3-1 shows the values for heart rate, blood pressure, and alertness recorded at the beginning and end

of each experimental session for both the placebo and levodopa conditions. We computed the percent change in heart rate and blood pressure recorded at the beginning and end of each session. There were no reliable differences between the levodopa and placebo conditions in the percent change of heart rate ($t(18) = 0.70$, $p=0.49$, 95%CI for difference = [-0.03 0.07], BF = 0.30, posterior δ: median = 0.139, 95%CI = [-0.278 0.565]), systolic blood pressure ($t(18) = -0.39$, $p=0.70$, 95%CI for difference = [-0.06 0.04], BF = 0.25, posterior δ: median = -0.077, 95%CI = [-0.498 0.338]), or diastolic blood pressure ($t(18) = -0.88$, $p=0.39$, 95%CI for difference = [-0.07 0.03], BF = 0.33, posterior δ: median = -0.173, 95%CI = [-0.603 0.245]). We did observe a significant difference between levodopa and placebo in the percent change of alertness ($t(20) = 2.46$, $p=0.023$, 95%CI for difference = [0.02 0.19], BF = 2.53, posterior δ: median = 0.477, 95%CI = [0.044 0.930]). However, this effect was likely due to chance as alertness was only different between the two drug conditions at the time point pre-administration of the capsule ($t(20) = 2.18$, $p=0.042$), but not post-administration ($t(20) = -0.068$, $p=0.95$). We also tested for effects of levodopa on the median response time (the latency between the go cue and the robot handle leaving the home position), and the median movement time. We observed no reliable differences in response time between the placebo and levodopa conditions in either the reward learning task ($t(20)=0.72$, $p=0.48$, 95%CI for difference = [-37.49 77.34], BF = 0.29, posterior δ: median = 0.137, 95%CI = [-0.261 0.545]), or the VMR task ($t(20)=0.62$, $p=0.54$, 95%CI for difference = [-33.91 62.56], BF = 0.27, posterior δ: median = 0.118, 95%CI = [-0.280 0.523]). We also observed no reliable difference in movement time between the placebo and levodopa conditions in either the reward learning task ($t(20)=-0.11$, $p=0.91$, 95%CI for difference = [-20.75 18.69], BF = 0.23, posterior δ: median = -0.021, 95%CI = [-0.420 0.377]), or the VMR task ($t(20)=-0.21$, $p=0.84$, 95%CI for difference = [-16.21 13.27], BF = 0.23, posterior δ: median = -0.039, 95%CI = [-0.44 0.358]).

Measure	Placebo	Levodopa
Heart Rate	Pre: 76.24 (SD: 11.29)	Pre: 77.55 (SD: 8.41)
	Post: 69.60 (SD: 7.27)	Post: 71.53 (SD: 6.92)
Systolic	Pre: 104.43 (SD: 9.01)	Pre: 103.95 (SD: 8.34)
	Post: 104.20 (SD: 6.47)	Post: 102.79 (SD: 8.70)
Diastolic	Pre: 72.14 (SD: 5.14)	Pre: 70.55 (SD: 6.81)
	Post: 73.20 (SD: 4.55)	Post: 69.74 (SD: 6.04)
Alertness	Pre: 64.58 (SD: 8.38)	Pre: 58.20 (SD: 11.79)
	Post: 47.99 (SD: 15.43)	Post: 48.16 (SD: 15.33)
Response Time	RL: 464.09 (SD: 140.05)	RL: 484.01 (SD: 149.00)
	VMR: 445.91 (SD: 120.96)	VMR: 460.24 (SD: 133.04)
Movement Time	RL: 548.17 (SD: 37.04)	RL: 547.14 (SD: 35.28)
	VMR: 547.90 (SD:34.92)	VMR: 546.43 (SD: 40.52)

Table 3-1: Control measurements. Heart rate (bpm). Systolic blood pressure (mm Hg). Diastolic blood pressure (mm Hg). Alertness, Bond-Lader visual analog scale alertness measure. Response Time, latency between go cue and hand exiting the start position (ms). Movement Time, duration of movement (ms). RL, reward learning task. VMR, visuomotor rotation task.

3.3.2 Behavioral results

Reward learning task. Behavioral data from the reward learning task are shown in Figure 3-2. Learning scores were reliably greater than zero in both the placebo condition (mean = 6.03, SD = 3.58, t(20) = 7.72, p = 2.02e-7, 95%CI = [4.40 7.66], BF = 1.56e5, posterior δ: median = 1.58 95%CI = [0.92 2.28]), and the levodopa condition (mean = 6.93, SD = 3.86, t(20) = 8.23, p = 7.49e-8, 95%CI = [5.17 8.69], BF = 3.9e5, posterior δ: median = 1.69, 95%CI = [1.00 2.41]) conditions. Learning scores were slightly higher in the levodopa condition, though this difference was not statistically reliable. This result provided strong evidence against our hypothesis of reduced learning in the levodopa group (t(20) = -1.58, p = 0.13, 95%CI for difference = [-2.09 0.29], BF = 0.10, posterior δ: median = -0.30, 95%CI = [-0.73 0.11]). We observed similar evidence against the hypothesized effect of levodopa when learning scores were computed using only the final 20 trials in each block (t(20) = -1.60, p = 0.13, 95%CI for difference = [-3.05 0.40], BF = 0.10, posterior δ: median = -0.31, 95%CI = [-0.73 0.10]).

The variability of trial-by-trial changes in reach angle following reward and non-reward outcomes is shown in Figure 3-3. We found a reliable main effect of reinforcement outcome on the log transformed variance of trial-by-trial changes in reach angle (F(1,20) = 74.84 , p = 3.41e-8, η_p^2= 0.79, BF = 3.02e14). This indicates an increase in trial-by-trial variance of reach angle following non-reward outcomes relative to reward. We found moderate evidence against effects of drug condition (F(1,20) = 0.0072 , p = 0.93, η_p^2= 3.86e-4, BF = 0.22) and reward by drug interaction (F(1,20) = 0.0478 , p = 0.829,η_p^2=2.38e-3, BF = 0.30).

Visuomotor rotation task. Mean trial-by-trial changes in reach angle after the different feedback rotations are shown in Figure 3-4. Learning rates were reliably greater than zero following administration of both placebo (mean: 0.313, SD: 0.133, t(20) = 10.77, p = 8.93e-10, 95%CI = [0.25 0.37], BF = 2.4e7, posterior δ: median = 2.22, 95%CI = [1.40 3.10])) and levodopa (mean: 0.294, SD: 0.102, t(20) = 13.18, p = 2.54e-11, 95%CI = [0.25 0.34], BF = 6.75e8). Learning rates were not reliably different in the two conditions (t(20)

= 0.703, p=0.491, 95%CI for difference = [-0.04 0.07], BF = 0.42, posterior δ: median = 0.134, 95%CI = [-0.265 0.540])).

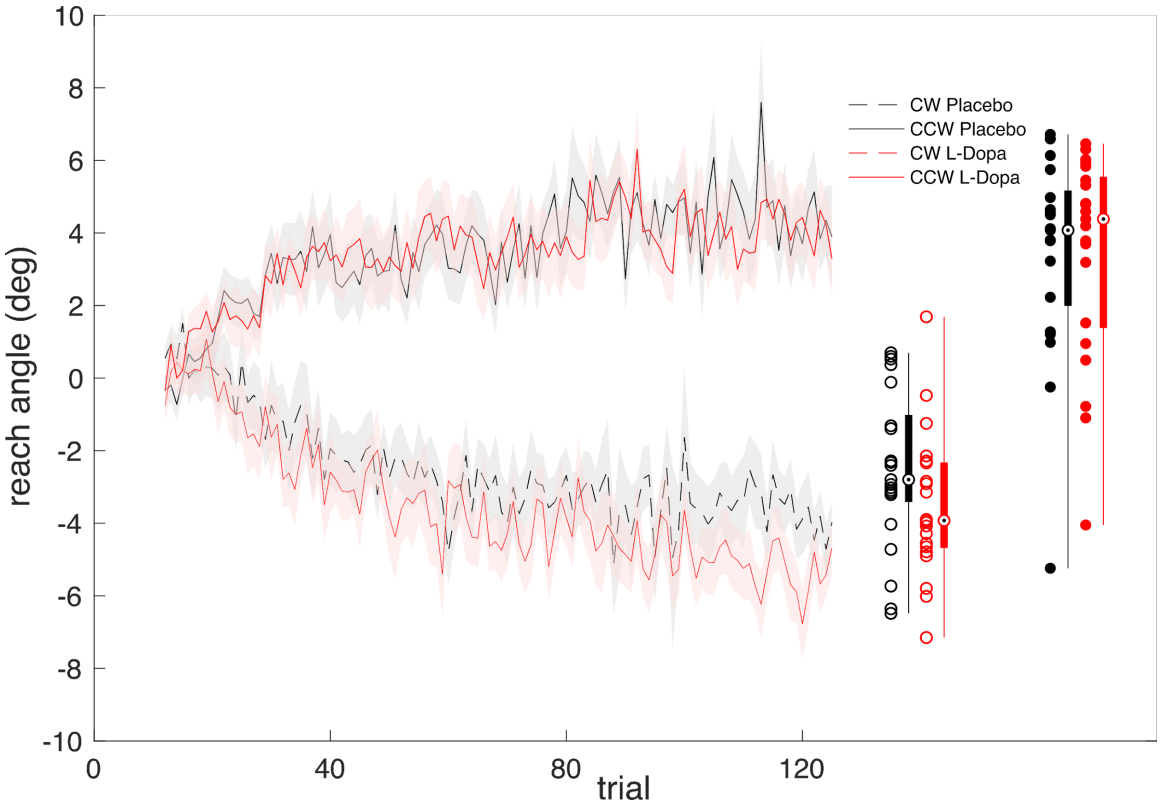

Figure 3-2: Behavior during the reward-based motor learning task (n=21). The time series show group average reach angles in the reward learning task across trials (Shaded region: ± SEM). After both placebo and levodopa administration, participants completed a block in each direction of intended learning condition [clockwise (CW) and counterclockwise (CCW)]. Trials 1-11 were baseline trials without reinforcement feedback, and are not shown. Individual data points on the right show the average reach angles across trials in each condition for each participant (CCW: solid markers, CW: open markers, black: placebo, red: L-Dopa). Box plots summarize the distributions of individual data using circular markers to indicate the medians, thick lines to indicate interquartile ranges, and thin lines to indicate full ranges.

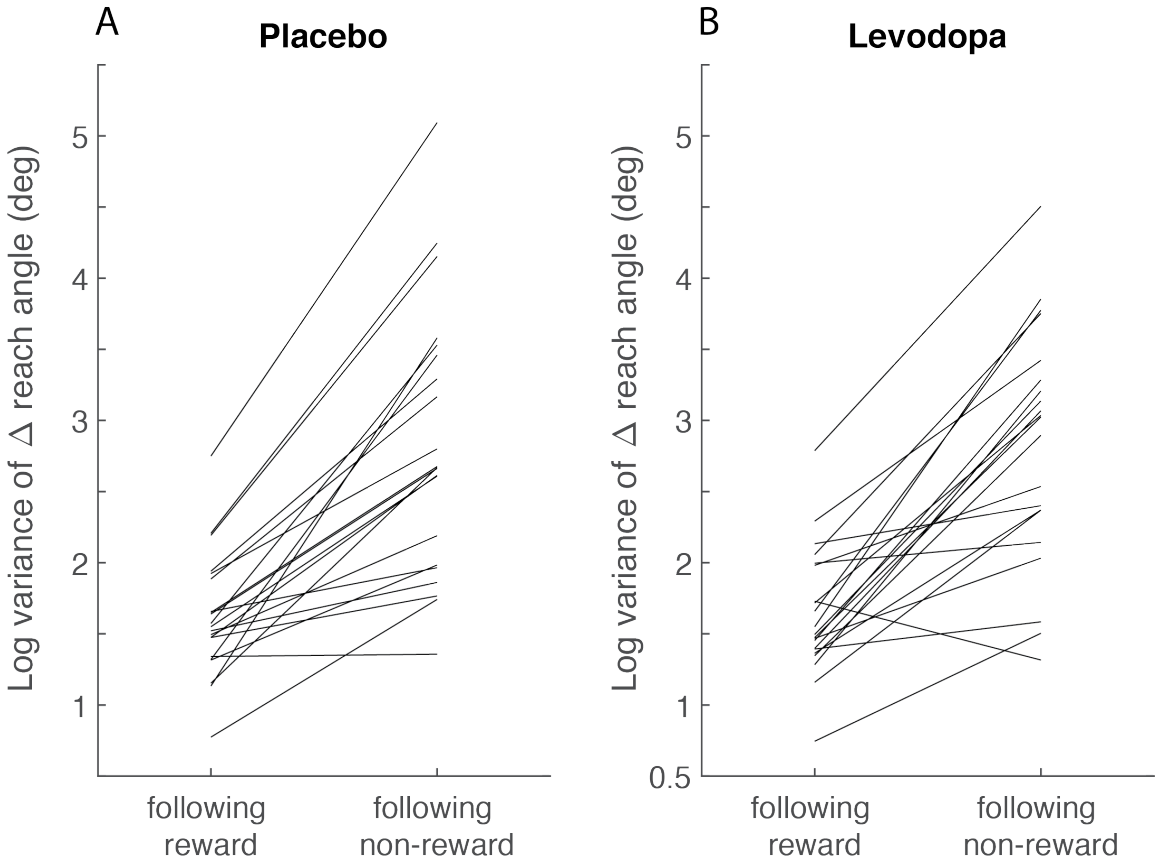

Figure 3-3: Reward induced modulation of trial-by-trial variability of reach angle (n=21). The log transformed variance of trial-by-trial changes in reach angle (deg) following reward and non-reward are plotted for each participant following administration of levodopa (**A**) and placebo (**B**).

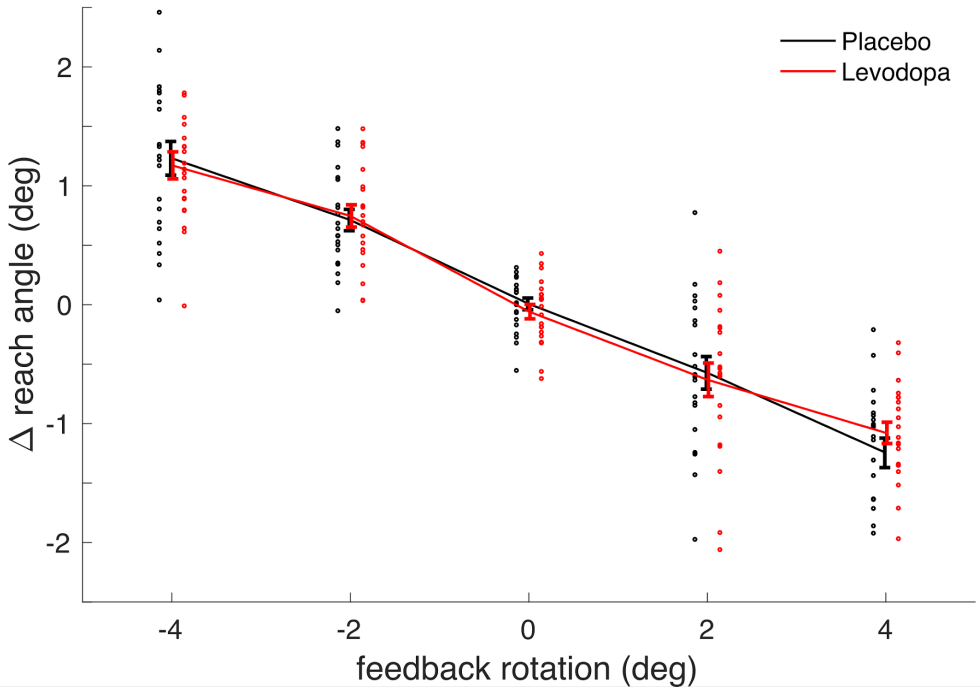

Figure 3-4: Behavior during the visuomotor rotation task (n=21). The average change in reach angle between subsequent pairs of trials is plotted for each size and direction of rotation imposed on the preceding trial. The average change in reach angle is in all cases opposite to the rotation, indicating that participants adapted their reaches to counteract the perturbations. Individual data points show average changes in reach angle across trials for each participant. Lines show average change in reach angle across participants (Error bars: ± SEM).

3.3.3 Event-related potential results

3.3.3.1 Reward learning task

Feedback-related negativity/Reward positivity: Event-related potentials (ERPs) elicited by reinforcement feedback at electrode FCz are shown in Figure 3-5A. We analyzed the FRN/RP by submitting the average ERP amplitude at electrode FCz between 267-317ms to frequentist and Bayesian repeated measures ANOVAs (figure 3-5B). We found a reliable main effect of reward outcome on FRN/RP amplitude ($F(1,19) = 42.25$, $p = 3.16e\text{-}6$, η_p^2=0.69, BF = 8.89e8). We observed moderate evidence both against effects of drug ($F(1,19) = 0.13$, $p = 0.73$, η_p^2=6.56e-3, BF = 0.24) and against a reward by drug interaction ($F(1,19) = 0.2$, $p = 0.66$, η_p^2=0.01, BF = 0.30) on FRN/RP amplitude.

P300: ERPs elicited by reinforcement feedback at electrode Pz are shown in Figure 3-5C. We analyzed the P300 by submitting the average ERP amplitudes at electrode Pz during the P300 time windows (Reward: 294-344ms, Non-reward: 346-396ms) to frequentist and Bayesian repeated measures ANOVAs (figure 3-5D). We found a reliable main effect of reward outcome on P300 amplitude ($F(1,19) = 35.83$, $p = 9.26e\text{-}6$, η_p^2=0.65, BF = 3.5e5). We observed moderate evidence both against an effect of drug ($F(1,19) = 0.20$, $p = 0.66$, η_p^2=0.01, BF = 0.26) and against a reward by drug interaction ($F(1,19) = 0.13$, $p = 0.73$, η_p^2= 6.56e-3, BF = 0.29) on P300 amplitudes.

3.3.3.2 Visuomotor rotation task

Feedback-related negativity/Reward positivity: ERPs elicited by endpoint cursor feedback at electrode FCz are shown in Figure 3-6A. We analyzed the FRN/RP by submitting the average ERP amplitude at electrode FCz between 267-317ms to repeated measures ANOVAs (figure 3-6B). We did not find reliable main effects of drug ($F(1,19) = 1.37$, $p = 0.26$, η_p^2=0.07), or feedback rotation ($F(2,38) = 0.1$, $p = 0.86$ (Greenhouse-Geisser corrected), η_p^2= 5.12e-3). We did observe a reliable drug by rotation interaction effect ($F(2,38) = 4.75$, $p = 0.02$ (Greenhouse-Geisser corrected), η_p^2= 0.2). Simple main effects

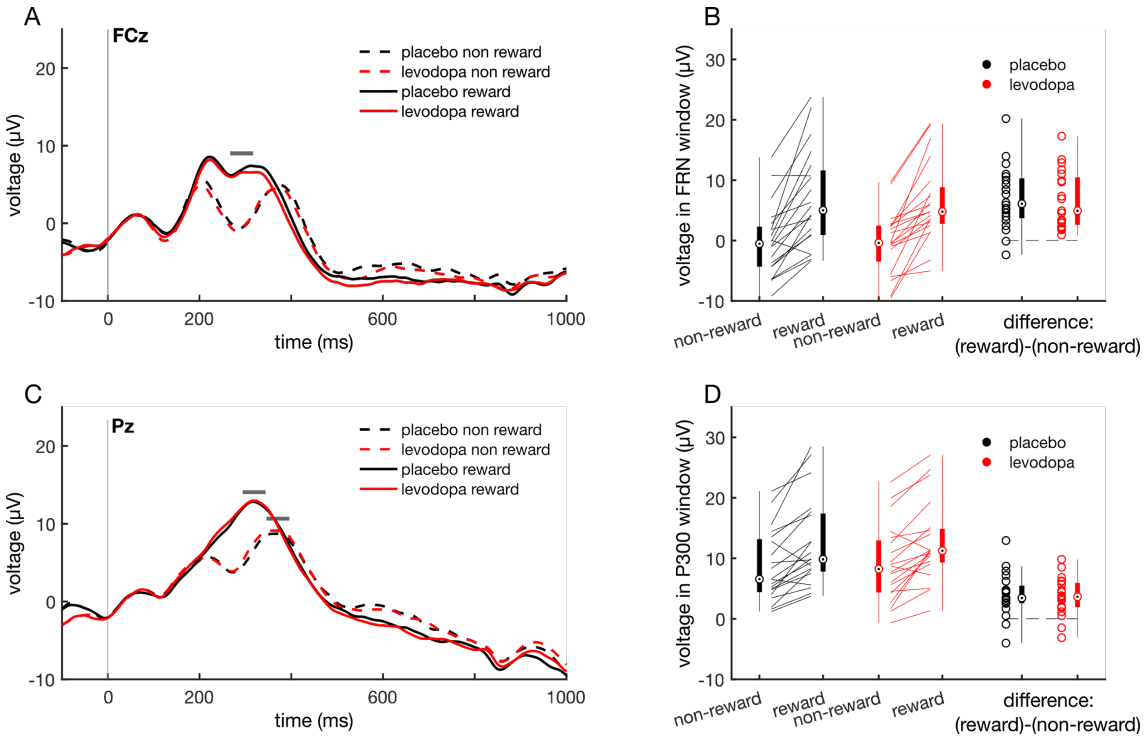

Figure 3-5: Event-related potentials elicited by reinforcement feedback (n=20). **A**, Grand averaged ERPs recorded from electrode FCz. ERPs are aligned to reinforcement feedback presentation (0 ms: vertical grey line). Horizontal grey bar indicates time window for FRN/RP analysis (267-317ms). Trials were selected by reinforcement outcome (reward or non-reward) and drug condition (levodopa or placebo) for ERP averaging. **B**, ERP amplitude during the FRN/RP time window. Individual participants' data show amplitude following reward, non-reward, and the difference [(reward) - (non-reward)]. Boxplots indicate the median (circular markers), the interquartile range (thick bars) and the range (thin lines). **C**, Trial averaged ERPs recorded from electrode Pz. Horizontal grey bars indicate time windows for P300 analysis (Reward: 294-344ms, Non-reward: 346-396ms). **D**, ERP amplitudes during the P300 time windows, as in B.

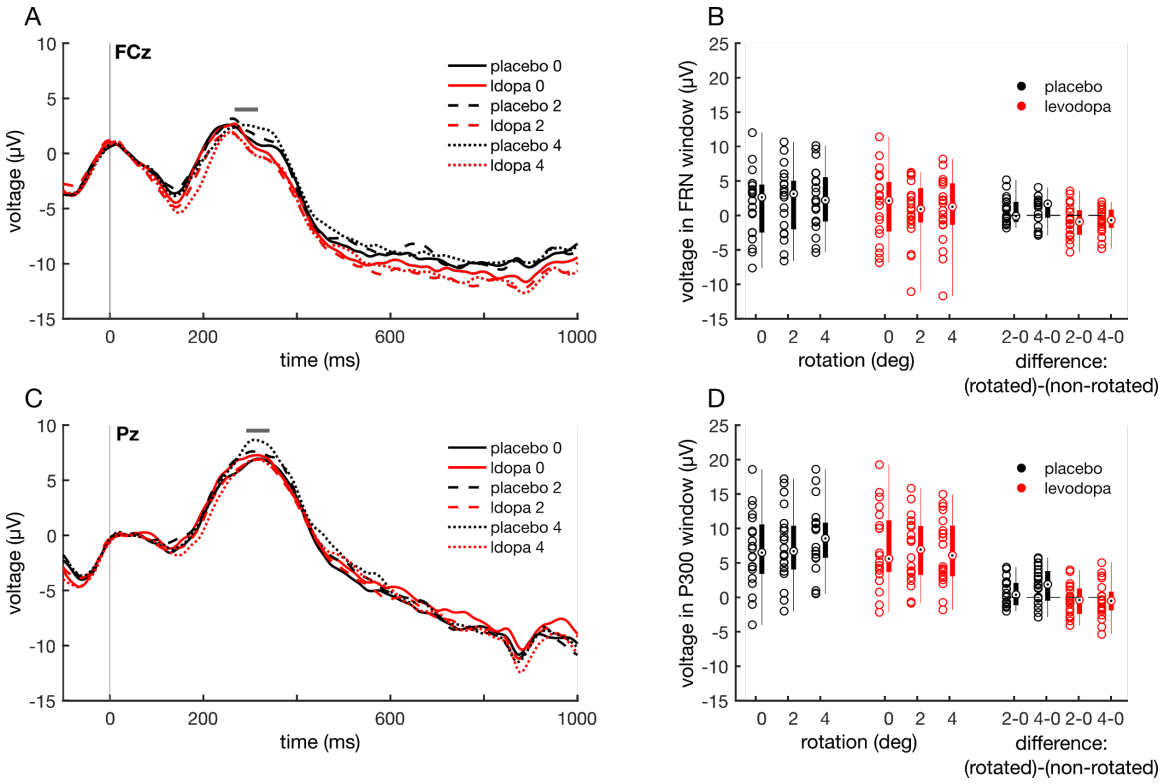

Figure 3-6: Event-related potentials elicited by sensory error feedback (n=20). **A**, Grand averaged ERPs recorded from electrode FCz. ERPs are aligned to endpoint cursor feedback presentation (0 ms: vertical gray line). Horizontal grey bar indicates time window for FRN/RP analysis (267-317ms). Trials were selected by feedback rotation (0°, ±2°, or ±4°) and drug condition (levodopa or placebo) for averaging. **B**, ERP amplitude during the FRN/RP time window. Individual participants' data show amplitude following unrotated feedback as well as feedback rotated by ±2°, and ±4°. Differences in ERP amplitude between rotated and unrotated feedback are also shown for each participant. Boxplots indicate the median (circular markers), the interquartile range (thick bars) and the range (thin lines). **C**, Trial averaged ERPs recorded from electrode Pz. Horizontal grey bars indicate time window for P300 analysis (292-342 ms). **D**, ERP amplitudes during the P300 time windows, as in B.

did not show reliable main effects of rotation in either the placebo ($F(2,38) = 2.17$, $p=0.13$) or levodopa ($F(2,38) = 2.06$, $p=0.14$) conditions on FRN/RP amplitudes.

P300: ERPs elicited by endpoint cursor feedback at electrode Pz are shown in Figure 3-6C. We analyzed the P300 by submitting the average ERP amplitude at electrode Pz between 292-342 ms to repeated measures ANOVAs (figure 3-6D). We did not find reliable main effects of drug ($F(1,19) = 0.43$, $p = 0.52$, η_p^2=0.02), or feedback rotation ($F(2,38) = 1.31$, $p = 0.28$ (Greenhouse-Geisser corrected), η_p^2= 0.06). We did observe a reliable drug by rotation interaction effect ($F(2,38) = 7.46$, $p = 2.24e\text{-}3$ (Greenhouse-Geisser corrected), η_p^2= 0.28). Simple main effects revealed a reliable main effect of rotation in the placebo ($F(2,38) = 5.72$, $p=6.72e\text{-}3$) but not the levodopa ($F(2,38) = 0.51$, $p=0.60$) condition on P300 amplitude.

3.4 Discussion

We tested for effects of levodopa, a dopamine precursor, on motor adaptation and event-related potential (ERP) responses to feedback in reward- and sensory-error based motor adaptation tasks. We hypothesized that levodopa would selectively impair neural and behavioral responses to reinforcement feedback in the reward-based learning task, while neural and behavioral responses to sensory error would be not be affected by levodopa. However, the only reliable influence of levodopa was in modulating the effect of visuomotor rotation on the P300 event-related potential component.

Visuomotor rotation task: During the VMR task, a cursor appeared at the endpoint of each reach to represent the position of the hand, and this feedback was perturbed through random rotations. We observed robust trial-by-trial adaptation to these perturbations. We did not find evidence that adaptation was affected by levodopa. This was expected, as trial-by-trial error correction induced by relatively small visuomotor rotations is thought to be driven primarily by sensory error-based learning mechanisms as opposed to dopaminergic reinforcement learning circuits (Diedrichsen et al., 2005; M. Ito, 2000; Krakauer et al., 2004; H. Tanaka et al., 2009; Taylor et al., 2010; Wong et al., 2019).

In the experiment described in chapter 2, we found that visuomotor rotation increases the amplitude of the P300 ERP component, a centro-parietal ERP deflection peaking approximately 300-400ms following feedback presentation (Aziz et al., 2020; MacLean et al., 2015; Palidis et al., 2019). In the present study, we observed an interaction effect between feedback rotation and drug condition on the P300 amplitude. P300 amplitude increased in response to visuomotor rotations in the placebo condition but not in the levodopa condition. This result replicated our previous finding that visuomotor rotations increase the amplitude of P300 responses to feedback, and additionally suggests that this effect is dependent on dopaminergic signaling. The modulation of P300 amplitude by sensory error is clearly not essential for adaptation, as disruption of this effect by levodopa did not correspond with any behavioral changes. Previous findings have also suggested a possible relationship between dopamine function and the P300 response, however the neural mechanisms and functional significance of the P300 in relation to motor adaptation remain unclear (Chu et al., 2018; Hansenne et al., 1995; Mulert et al., 2006; Noble et al., 1994; Sohn et al., 1998; Stanzione et al., 1990, 1991; Takeshita & Ogura, 1994). Variants of the P300 are elicited by many types of task-relevant stimuli, and have been localized to diffuse cortical areas including parietal, frontal, and motor regions, which have been implicated in processing prediction error (Bledowski et al., 2004; Calhoun et al., 2006; Johnson et al., 2019; Y. Li et al., 2009; Mantini et al., 2009; Polich, 2007; Ragazzoni et al., 2019; Sabeti et al., 2016; Soltani & Knight, 2000). We observed a similar interaction effect between rotation and drug condition in recordings from electrode FCz during the FRN/RP time window. This appeared to be largely attributable to the P300 effect described above, as the time windows were largely overlapping and the P300 was clearly measured at FCz as well.

Reward learning task: Participants adapted reliably to manipulations of binary reinforcement feedback intended to produce either progressively clockwise or counterclockwise reach angles. However, we found no effects of levodopa on adaptation. One explanation of our findings is that the behavioral and neural processes measured in the current study do not depend on dopaminergic reward learning mechanisms. Another

possibility is that the drug manipulation was not sufficiently powerful to disrupt these processes. The former interpretation depends on previous findings that levodopa impairs cognitive forms of reward learning using the same drug administration protocols in similar populations. However, the current study is limited by the lack of a positive control task demonstrating known behavioral effects of levodopa. Quattrocchi et al. (2018) found no effect of levodopa or a dopamine antagonist haloperidol on modulation of sensory error-based learning by additional reinforcement feedback. Holland et al. (2019) found no association between dopamine-related gene polymorphisms on adaptation through binary reinforcement feedback in a task similar to that used in the current study. Together, these findings suggest that reward-based motor adaptation may not rely on dopamine function, or at least that additional mechanisms may compensate for differences in dopamine function.

The "dopamine overdose" hypothesis states that dopaminergic medications such as levodopa might disrupt learning processes mediated by the ventral striatum by overstimulating dopamine signaling in this brain region. The ventral striatum may specifically mediate stimulus-based reinforcement learning, while action-based reinforcement learning in the current study may be subserved by the dorsal striatum (Rothenhoefer et al., 2017). Furthermore, levodopa may specifically impair learning from unfavorable outcomes as opposed to rewarding outcomes (Cools et al., 2006, 2007; Frank et al., 2004; Vo et al., 2018). Non-reward outcomes in the current task may not contribute significantly to learning as they do not instruct the correct response, unlike in binary response tasks.

Another important distinction is between model-free and model-based reinforcement learning processes (Babayan et al., 2018; Daw et al., 2011; Deserno et al., 2015; Dolan & Dayan, 2013; Doll et al., 2016; Gardner et al., 2018; Gläscher et al., 2010; Russek et al., 2017; Sambrook et al., 2018; Shahar et al., 2019; Sharp et al., 2016). Model-free reinforcement learning is characterized by reinforcement of simple stimulus-response associations that facilitate habitual, reflexive responding. Model-based learning allows for flexible planning according to a mental representation of the task, and can be limited by

working memory processes. Levodopa has been shown to impair reward-based learning in healthy controls and people with Parkinson's disease, but to improve model-based learning and related cognitive functions such as working memory, cognitive flexibility, and attention (Beato et al., 2008; Cools et al., 2001, 2003; Cooper et al., 1992; Costa et al., 2003; Kulisevsky, 2000; Lange et al., 1992; Lewis et al., 2005; Marini et al., 2003; Moustafa et al., 2008; Sharp et al., 2016; Torta et al., 2009; Wunderlich et al., 2012). It is possible that "dopamine overdose" by levodopa selectively impairs model-free learning. It may be that reward-based motor adaptation in the current study relies on processes other than model-free learning that are not affected by levodopa. Reward-based motor adaptation tasks similar to that in the current study have been shown to primarily involve strategic aiming that can be influenced by explicit instructions and cognitive load, characteristics that are inconsistent with model-free learning (Codol et al., 2018; Holland et al., 2018).

We also analyzed the variability of trial-by-trial changes in reach angle as a function of reward outcomes. Reward related modulation of motor variability has been shown to be impaired in medicated Parkinson's disease in a very similar task (Pekny et al., 2015). We hypothesized that this effect may be due to side-effects of dopaminergic medication, and that we would observe similar impairments in healthy participants after levodopa administration. However, we observed no effect of levodopa on reward-related modulation of motor variability. Reward-based modulation of exploratory variance may therefore not depend on the ventral striatum, which is relatively spared in early stage Parkinson's disease and therefore vulnerable to "dopamine overdose" in patients and healthy controls alike. Instead, it may depend on the dorsal striatum, which is more closely related to movement planning and is primarily impacted by early stage Parkinson's disease.

Reinforcement feedback elicited a very reliable FRN/RP ERP component. Meta-analyses have shown that the FRN/RP encodes a quantitative reward prediction error across multiple different tasks (Sambrook & Goslin, 2015; Walsh & Anderson, 2012). Reports have linked the FRN/RP signal to behavioral adjustments in response to feedback (Arbel et al., 2013; Frank et al., 2005; Holroyd & Krigolson, 2007; van der Helden et al., 2010).

These findings support a prominent theory purporting that the FRN/RP is a reflection of reinforcement learning processes in the anterior cingulate cortex driven by phasic dopamine reward prediction error signals (Holroyd & Coles, 2002; Walsh & Anderson, 2012). Contrary to our hypothesis, we observed no effects of levodopa on the FRN/RP in response to reinforcement feedback. Previous studies have supported a link between dopamine and the FRN/RP, although results have been mixed. FRN/RP amplitude has been shown to be impaired in Parkinson's disease patients with apathy (Martínez-Horta et al., 2014). Brown et al. (2020) found that the reward positivity was impaired in Parkinson's disease patients relative to controls ON levodopa but not OFF levodopa, consistent with the dopamine overdose hypothesis. In healthy participants, the dopamine antagonist haloperidol has shown mixed results in reducing the amplitude of the reward positivity (Forster et al., 2017; Schutte et al., 2020). Mueller et al. (2014) found that the D2 receptor dopamine antagonist sulpiride had opposite effects on FRN/RP amplitude depending on a genotype variant that regulates prefrontal dopamine levels. They suggested a u-shaped relationship between dopamine release in the prefrontal cortex and FRN/RP amplitude mediated by the balance between D1 and D2 receptor activation. Because the effect of dopamine manipulation on the FRN/RP seems to depend on genetic differences in baseline dopamine release, one possibility is that levodopa in the current study had inconsistent effects on different subgroups of participants that cancelled each other in the group average.

Conclusions: We tested the hypothesis that reward-based motor adaptation is mediated by dopaminergic signaling of reward prediction error. Contrary to our hypotheses, we did not observe effects of levodopa on reward-based motor learning or the FRN/RP ERP component, which have both been theorized to depend on dopaminergic signaling of reward prediction error. The dopamine overdose hypothesis suggests that levodopa impairs stimulus-response reinforcement learning processes in the ventral striatum. Reward-based motor adaptation may instead depend on distinct reinforcement learning circuits that are not disrupted by levodopa such as cortical reward learning mechanisms or dopaminergic projections to the dorsal striatum.

3.5 References

Adams, R. A., Shipp, S., & Friston, K. J. (2013). Predictions not commands: Active inference in the motor system. *Brain Structure & Function*, *218*(3), 611–643. https://doi.org/10.1007/s00429-012-0475-5

Arbel, Y., Goforth, K., & Donchin, E. (2013). The Good, the Bad, or the Useful? The Examination of the Relationship between the Feedback-related Negativity (FRN) and Long-term Learning Outcomes. *Journal of Cognitive Neuroscience*, *25*(8), 1249–1260. https://doi.org/10.1162/jocn_a_00385

Aziz, J. R., MacLean, S. J., Krigolson, O. E., & Eskes, G. A. (2020). Visual Feedback Modulates Aftereffects and Electrophysiological Markers of Prism Adaptation. *Frontiers in Human Neuroscience*, *14*. https://doi.org/10.3389/fnhum.2020.00138

Babayan, B. M., Uchida, N., & Gershman, S. J. (2018). Belief state representation in the dopamine system. *Nature Communications*, *9*(1), 1891. https://doi.org/10.1038/s41467-018-04397-0

Bayer, H. M., & Glimcher, P. W. (2005). Midbrain Dopamine Neurons Encode a Quantitative Reward Prediction Error Signal. *Neuron*, *47*(1), 129–141. https://doi.org/10.1016/j.neuron.2005.05.020

Beato, R., Levy, R., Pillon, B., Vidal, C., du Montcel, S. T., Deweer, B., Bonnet, A.-M., Houeto, J.-L., Dubois, B., & Cardoso, F. (2008). Working memory in Parkinson's disease patients: Clinical features and response to levodopa. *Arquivos De Neuro-Psiquiatria*, *66*(2A), 147–151. https://doi.org/10.1590/s0004-282x2008000200001

Becker, M. P. I., Nitsch, A. M., Miltner, W. H. R., & Straube, T. (2014). A single-trial estimation of the feedback-related negativity and its relation to BOLD responses in a time-estimation task. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *34*(8), 3005–3012. https://doi.org/10.1523/JNEUROSCI.3684-13.2014

Bernardi, N. F., Darainy, M., & Ostry, D. J. (2015). Somatosensory Contribution to the Initial Stages of Human Motor Learning. *Journal of Neuroscience*, *35*(42), 14316–14326. https://doi.org/10.1523/JNEUROSCI.1344-15.2015

Bhanpuri, N. H., Okamura, A. M., & Bastian, A. J. (2013). Predictive Modeling by the Cerebellum Improves Proprioception. *Journal of Neuroscience*, *33*(36), 14301–14306. https://doi.org/10.1523/JNEUROSCI.0784-13.2013

Bledowski, C., Prvulovic, D., Hoechstetter, K., Scherg, M., Wibral, M., Goebel, R., & Linden, D. E. J. (2004). Localizing P300 generators in visual target and distractor processing: A combined event-related potential and functional magnetic resonance imaging study. *The Journal of Neuroscience: The Official Journal of the*

Society for Neuroscience, *24*(42), 9353–9360. https://doi.org/10.1523/JNEUROSCI.1897-04.2004

Bond, A., & Lader, M. (1974). The use of analogue scales in rating subjective feelings. *British Journal of Medical Psychology*, *47*(3), 211–218. https://doi.org/10.1111/j.2044-8341.1974.tb02285.x

Brooks, J. L., Zoumpoulaki, A., & Bowman, H. (2017). Data-driven region-of-interest selection without inflating Type I error rate. *Psychophysiology*, *54*(1), 100–113. https://doi.org/10.1111/psyp.12682

Brown, D. R., Richardson, S. P., & Cavanagh, J. F. (2020). An EEG marker of reward processing is diminished in Parkinson's disease. *Brain Research*, *1727*, 146541. https://doi.org/10.1016/j.brainres.2019.146541

Calhoun, V. D., Adali, T., Pearlson, G. D., & Kiehl, K. A. (2006). Neuronal chronometry of target detection: Fusion of hemodynamic and event-related potential data. *NeuroImage*, *30*(2), 544–553. https://doi.org/10.1016/j.neuroimage.2005.08.060

Carlson, J. M., Foti, D., Mujica-Parodi, L. R., Harmon-Jones, E., & Hajcak, G. (2011). Ventral striatal and medial prefrontal BOLD activation is correlated with reward-related electrocortical activity: A combined ERP and fMRI study. *NeuroImage*, *57*(4), 1608–1616. https://doi.org/10.1016/j.neuroimage.2011.05.037

Cashaback, J. G. A., Lao, C. K., Palidis, D. J., Coltman, S. K., McGregor, H. R., & Gribble, P. L. (2019). The gradient of the reinforcement landscape influences sensorimotor learning. *PLOS Computational Biology*, *15*(3), e1006839. https://doi.org/10.1371/journal.pcbi.1006839

Cashaback, J. G. A., McGregor, H. R., Mohatarem, A., & Gribble, P. L. (2017). Dissociating error-based and reinforcement-based loss functions during sensorimotor learning. *PLoS Computational Biology*, *13*(7), e1005623. https://doi.org/10.1371/journal.pcbi.1005623

Chu, C.-L., Lee, I. H., Chi, M. H., Chen, K. C., Chen, P. S., Yao, W. J., Chiu, N. T., & Yang, Y. K. (2018). Availability of dopamine transporters and auditory P300 abnormalities in adults with attention-deficit hyperactivity disorder: Preliminary results. *CNS Spectrums*, *23*(4), 264–270. https://doi.org/10.1017/S1092852917000049

Codol, O., Holland, P. J., & Galea, J. M. (2018). The relationship between reinforcement and explicit control during visuomotor adaptation. *Scientific Reports*, *8*(1), 9121. https://doi.org/10.1038/s41598-018-27378-1

Cools, R., Altamirano, L., & D'Esposito, M. (2006). Reversal learning in Parkinson's disease depends on medication status and outcome valence. *Neuropsychologia*, *44*(10), 1663–1673. https://doi.org/10.1016/j.neuropsychologia.2006.03.030

Cools, R., Barker, R. A., Sahakian, B. J., & Robbins, T. W. (2001). Enhanced or impaired cognitive function in Parkinson's disease as a function of dopaminergic medication and task demands. *Cerebral Cortex (New York, N.Y.: 1991)*, *11*(12), 1136–1143. https://doi.org/10.1093/cercor/11.12.1136

Cools, R., Barker, R. A., Sahakian, B. J., & Robbins, T. W. (2003). L-Dopa medication remediates cognitive inflexibility, but increases impulsivity in patients with Parkinson's disease. *Neuropsychologia*, *41*(11), 1431–1441. https://doi.org/10.1016/s0028-3932(03)00117-9

Cools, R., Lewis, S. J. G., Clark, L., Barker, R. A., & Robbins, T. W. (2007). L-DOPA disrupts activity in the nucleus accumbens during reversal learning in Parkinson's disease. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, *32*(1), 180–189. https://doi.org/10.1038/sj.npp.1301153

Cooper, J. A., Sagar, H. J., Doherty, S. M., Jordan, N., Tidswell, P., & Sullivan, E. V. (1992). Different effects of dopaminergic and anticholinergic therapies on cognitive and motor function in Parkinson's disease. A follow-up study of untreated patients. *Brain: A Journal of Neurology*, *115 (Pt 6)*, 1701–1725. https://doi.org/10.1093/brain/115.6.1701

Costa, A., Peppe, A., Dell'Agnello, G., Carlesimo, G. A., Murri, L., Bonuccelli, U., & Caltagirone, C. (2003). Dopaminergic modulation of visual-spatial working memory in Parkinson's disease. *Dementia and Geriatric Cognitive Disorders*, *15*(2), 55–66. https://doi.org/10.1159/000067968

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, *69*(6), 1204–1215. https://doi.org/10.1016/j.neuron.2011.02.027

Daw, N. D., & Tobler, P. N. (2014). Chapter 15 - Value Learning through Reinforcement: The Basics of Dopamine and Reinforcement Learning. In P. W. Glimcher & E. Fehr (Eds.), *Neuroeconomics (Second Edition)* (pp. 283–298). Academic Press. https://doi.org/10.1016/B978-0-12-416008-8.00015-2

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Deserno, L., Huys, Q. J. M., Boehme, R., Buchert, R., Heinze, H.-J., Grace, A. A., Dolan, R. J., Heinz, A., & Schlagenhauf, F. (2015). Ventral striatal dopamine reflects behavioral and neural signatures of model-based control during sequential decision making. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(5), 1595–1600. https://doi.org/10.1073/pnas.1417219112

Dhawale, A. K., Miyamoto, Y. R., Smith, M. A., & Ölveczky, B. P. (2019). Adaptive Regulation of Motor Variability. *Current Biology*, *29*(21), 3551-3562.e7. https://doi.org/10.1016/j.cub.2019.08.052

Diederen, K. M. J., Ziauddeen, H., Vestergaard, M. D., Spencer, T., Schultz, W., & Fletcher, P. C. (2017). Dopamine Modulates Adaptive Prediction Error Coding in the Human Midbrain and Striatum. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *37*(7), 1708–1720. https://doi.org/10.1523/JNEUROSCI.1979-16.2016

Diedrichsen, J., Hashambhoy, Y., Rane, T., & Shadmehr, R. (2005). Neural Correlates of Reach Errors. *Journal of Neuroscience*, *25*(43), 9919–9931. https://doi.org/10.1523/JNEUROSCI.1874-05.2005

Dolan, R. J., & Dayan, P. (2013). Goals and Habits in the Brain. *Neuron*, *80*(2), 312–325. https://doi.org/10.1016/j.neuron.2013.09.007

Doll, B. B., Bath, K. G., Daw, N. D., & Frank, M. J. (2016). Variability in Dopamine Genes Dissociates Model-Based and Model-Free Reinforcement Learning. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *36*(4), 1211–1222. https://doi.org/10.1523/JNEUROSCI.1901-15.2016

Emeric, E. E., Brown, J. W., Leslie, M., Pouget, P., Stuphorn, V., & Schall, J. D. (2008). Performance monitoring local field potentials in the medial frontal cortex of primates: Anterior cingulate cortex. *Journal of Neurophysiology*, *99*(2), 759–772. https://doi.org/10.1152/jn.00896.2006

Enge, S., Mothes, H., Fleischhauer, M., Reif, A., & Strobel, A. (2017). Genetic variation of dopamine and serotonin function modulates the feedback-related negativity during altruistic punishment. *Scientific Reports*, *7*(1), 2996. https://doi.org/10.1038/s41598-017-02594-3

Feigin, A., Ghilardi, M. F., Carbon, M., Edwards, C., Fukuda, M., Dhawan, V., Margouleff, C., Ghez, C., & Eidelberg, D. (2003). Effects of levodopa on motor sequence learning in Parkinson's disease. *Neurology*, *60*(11), 1744–1749. https://doi.org/10.1212/01.wnl.0000072263.03608.42

Flöel, A., Breitenstein, C., Hummel, F., Celnik, P., Gingert, C., Sawaki, L., Knecht, S., & Cohen, L. G. (2005). Dopaminergic influences on formation of a motor memory. *Annals of Neurology*, *58*(1), 121–130. https://doi.org/10.1002/ana.20536

Forster, S. E., Zirnheld, P., Shekhar, A., Steinhauer, S. R., O'Donnell, B. F., & Hetrick, W. P. (2017). Event-related potentials reflect impaired temporal interval learning following haloperidol administration. *Psychopharmacology*, *234*(17), 2545–2562. https://doi.org/10.1007/s00213-017-4645-2

Foti, D., Weinberg, A., Dien, J., & Hajcak, G. (2011). Event-related potential activity in the basal ganglia differentiates rewards from nonrewards: Temporospatial principal components analysis and source localization of the feedback negativity. *Human Brain Mapping*, *32*(12), 2207–2216. https://doi.org/10.1002/hbm.21182

Frank, M. J., Seeberger, L. C., & O'reilly, R. C. (2004). By carrot or by stick: Cognitive reinforcement learning in parkinsonism. *Science (New York, N.Y.)*, *306*(5703), 1940–1943. https://doi.org/10.1126/science.1102941

Frank, M. J., Woroch, B. S., & Curran, T. (2005). Error-Related Negativity Predicts Reinforcement Learning and Conflict Biases. *Neuron*, *47*(4), 495–501. https://doi.org/10.1016/j.neuron.2005.06.020

García-García, I., Zeighami, Y., & Dagher, A. (2017). Reward Prediction Errors in Drug Addiction and Parkinson's Disease: From Neurophysiology to Neuroimaging. *Current Neurology and Neuroscience Reports*, *17*(6), 46. https://doi.org/10.1007/s11910-017-0755-9

Gardner, M. P. H., Schoenbaum, G., & Gershman, S. J. (2018). Rethinking dopamine as generalized prediction error. *Proceedings. Biological Sciences*, *285*(1891). https://doi.org/10.1098/rspb.2018.1645

Gehring, W. J., & Willoughby, A. R. (2002). The Medial Frontal Cortex and the Rapid Processing of Monetary Gains and Losses. *Science*, *295*(5563), 2279–2282. https://doi.org/10.1126/science.1066893

Gläscher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). States versus rewards: Dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*, *66*(4), 585–595. https://doi.org/10.1016/j.neuron.2010.04.016

Graef, S., Biele, G., Krugel, L. K., Marzinzik, F., Wahl, M., Wotka, J., Klostermann, F., & Heekeren, H. R. (2010). Differential influence of levodopa on reward-based learning in Parkinson's disease. *Frontiers in Human Neuroscience*, *4*, 169. https://doi.org/10.3389/fnhum.2010.00169

Hansenne, M., Pitchot, W., Gonzalez Moreno, A., Papart, P., Timsit-Berthier, M., & Ansseau, M. (1995). Catecholaminergic function and P300 amplitude in major depressive disorder (P300 and catecholamines). *Electroencephalography and Clinical Neurophysiology*, *96*(2), 194–196. https://doi.org/10.1016/0168-5597(94)00317-8

Hauser, T. U., Iannaccone, R., Stämpfli, P., Drechsler, R., Brandeis, D., Walitza, S., & Brem, S. (2014). The feedback-related negativity (FRN) revisited: New insights into the localization, meaning and network organization. *NeuroImage*, *84*, 159–168. https://doi.org/10.1016/j.neuroimage.2013.08.028

Hiebert, N. M., Owen, A. M., Ganjavi, H., Mendonça, D., Jenkins, M. E., Seergobin, K. N., & MacDonald, P. A. (2019). Dorsal striatum does not mediate feedback-based, stimulus-response learning: An event-related fMRI study in patients with Parkinson's disease tested on and off dopaminergic therapy. *NeuroImage*, *185*, 455–470. https://doi.org/10.1016/j.neuroimage.2018.10.045

Hiebert, N. M., Seergobin, K. N., Vo, A., Ganjavi, H., & MacDonald, P. A. (2014). Dopaminergic therapy affects learning and impulsivity in Parkinson's disease. *Annals of Clinical and Translational Neurology*, *1*(10), 833–843. https://doi.org/10.1002/acn3.128

Holland, P., Codol, O., & Galea, J. M. (2018). Contribution of explicit processes to reinforcement-based motor learning. *Journal of Neurophysiology*, *119*(6), 2241–2255. https://doi.org/10.1152/jn.00901.2017

Holland, P., Codol, O., Oxley, E., Taylor, M., Hamshere, E., Joseph, S., Huffer, L., & Galea, J. M. (2019). Domain-Specific Working Memory, But Not Dopamine-Related Genetic Variability, Shapes Reward-Based Motor Learning. *The Journal of Neuroscience*, *39*(47), 9383–9396. https://doi.org/10.1523/JNEUROSCI.0583-19.2019

Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*(4), 679–709. https://doi.org/10.1037/0033-295X.109.4.679

Holroyd, C. B., & Krigolson, O. E. (2007). Reward prediction error signals associated with a modified time estimation task. *Psychophysiology*, *44*(6), 913–917. https://doi.org/10.1111/j.1469-8986.2007.00561.x

Holroyd, C. B., Pakzad-Vaezi, K. L., & Krigolson, O. E. (2008). The feedback correct-related positivity: Sensitivity of the event-related brain potential to unexpected positive feedback. *Psychophysiology*, *45*(5), 688–697. https://doi.org/10.1111/j.1469-8986.2008.00668.x

Huang, V. S., Haith, A., Mazzoni, P., & Krakauer, J. W. (2011). Rethinking Motor Learning and Savings in Adaptation Paradigms: Model-Free Memory for Successful Actions Combines with Internal Models. *Neuron*, *70*(4), 787–801. https://doi.org/10.1016/j.neuron.2011.04.012

Inoue, K., Kawashima, R., Satoh, K., Kinomura, S., Sugiura, M., Goto, R., Ito, M., & Fukuda, H. (2000). A PET study of visuomotor learning under optical rotation. *NeuroImage*, *11*(5 Pt 1), 505–516. https://doi.org/10.1006/nimg.2000.0554

Ito, M. (2000). Mechanisms of motor learning in the cerebellum1. *Brain Research*, *886*(1–2), 237–245. https://doi.org/10.1016/S0006-8993(00)03142-5

Izawa, J., & Shadmehr, R. (2011). Learning from Sensory and Reward Prediction Errors during Motor Adaptation. *PLOS Computational Biology*, *7*(3), e1002012. https://doi.org/10.1371/journal.pcbi.1002012

Jahanshahi, M., Wilkinson, L., Gahir, H., Dharmaindra, A., Dharmarinda, A., Dharminda, A., & Lagnado, D. A. (2010). Medication impairs probabilistic classification learning in Parkinson's disease. *Neuropsychologia*, *48*(4), 1096–1103. https://doi.org/10.1016/j.neuropsychologia.2009.12.010

Jocham, G., & Ullsperger, M. (2009). Neuropharmacology of performance monitoring. *Neuroscience and Biobehavioral Reviews*, *33*(1), 48–60. https://doi.org/10.1016/j.neubiorev.2008.08.011

Johnson, J. F., Belyk, M., Schwartze, M., Pinheiro, A. P., & Kotz, S. A. (2019). The role of the cerebellum in adaptation: ALE meta-analyses on sensory feedback error. *Human Brain Mapping*, *40*(13), 3966–3981. https://doi.org/10.1002/hbm.24681

Keysers, C., Gazzola, V., & Wagenmakers, E.-J. (2020). Using Bayes factor hypothesis testing in neuroscience to establish evidence of absence. *Nature Neuroscience*, *23*(7), 788–799. https://doi.org/10.1038/s41593-020-0660-4

Kim, H. E., Parvin, D. E., & Ivry, R. B. (2019). The influence of task outcome on implicit motor learning. *ELife*, *8*. https://doi.org/10.7554/eLife.39882

Knecht, S., Breitenstein, C., Bushuven, S., Wailke, S., Kamping, S., Flöel, A., Zwitserlood, P., & Ringelstein, E. B. (2004). Levodopa: Faster and better word learning in normal humans. *Annals of Neurology*, *56*(1), 20–26. https://doi.org/10.1002/ana.20125

Kooij, K. van der, & Smeets, J. B. J. (2019). Reward-based motor adaptation can generalize across actions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(1), 71–81. https://doi.org/10.1037/xlm0000573

Kooij, K. van der, Wijdenes, L. O., Rigterink, T., Overvliet, K. E., & Smeets, J. B. J. (2018). Reward abundance interferes with error-based learning in a visuomotor adaptation task. *PLOS ONE*, *13*(3), e0193002. https://doi.org/10.1371/journal.pone.0193002

Krakauer, J. W. (2009). Motor Learning and Consolidation: The Case of Visuomotor Rotation. In D. Sternad (Ed.), *Progress in Motor Control: A Multidisciplinary Perspective* (pp. 405–421). Springer US. https://doi.org/10.1007/978-0-387-77064-2_21

Krakauer, J. W., Ghez, C., & Ghilardi, M. F. (2005). Adaptation to Visuomotor Transformations: Consolidation, Interference, and Forgetting. *Journal of Neuroscience*, *25*(2), 473–478. https://doi.org/10.1523/JNEUROSCI.4218-04.2005

Krakauer, J. W., Ghilardi, M.-F., Mentis, M., Barnes, A., Veytsman, M., Eidelberg, D., & Ghez, C. (2004). Differential Cortical and Subcortical Activations in Learning Rotations and Gains for Reaching: A PET Study. *Journal of Neurophysiology*, *91*(2), 924–933. https://doi.org/10.1152/jn.00675.2003

Kulisevsky, J. (2000). Role of dopamine in learning and memory: Implications for the treatment of cognitive dysfunction in patients with Parkinson's disease. *Drugs & Aging*, *16*(5), 365–379. https://doi.org/10.2165/00002512-200016050-00006

Kwak, Y., Müller, M. L. T. M., Bohnen, N. I., Dayalu, P., & Seidler, R. D. (2010). Effect of dopaminergic medications on the time course of explicit motor sequence learning in Parkinson's disease. *Journal of Neurophysiology*, *103*(2), 942–949. https://doi.org/10.1152/jn.00197.2009

Lange, K. W., Robbins, T. W., Marsden, C. D., James, M., Owen, A. M., & Paul, G. M. (1992). L-dopa withdrawal in Parkinson's disease selectively impairs cognitive performance in tests sensitive to frontal lobe dysfunction. *Psychopharmacology*, *107*(2–3), 394–404. https://doi.org/10.1007/BF02245167

Lewis, S. J. G., Slabosz, A., Robbins, T. W., Barker, R. A., & Owen, A. M. (2005). Dopaminergic basis for deficits in working memory but not attentional set-shifting in Parkinson's disease. *Neuropsychologia*, *43*(6), 823–832. https://doi.org/10.1016/j.neuropsychologia.2004.10.001

Li, Y., Wang, L.-Q., & Hu, Y. (2009). Localizing P300 generators in high-density event-related potential with fMRI. *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research*, *15*(3), MT47-53.

MacDonald, P. A., MacDonald, A. A., Seergobin, K. N., Tamjeedi, R., Ganjavi, H., Provost, J.-S., & Monchi, O. (2011). The effect of dopamine therapy on ventral and dorsal striatum-mediated cognition in Parkinson's disease: Support from functional MRI. *Brain: A Journal of Neurology*, *134*(Pt 5), 1447–1463. https://doi.org/10.1093/brain/awr075

MacLean, S. J., Hassall, C. D., Ishigami, Y., Krigolson, O. E., & Eskes, G. A. (2015). Using brain potentials to understand prism adaptation: The error-related negativity and

the P300. *Frontiers in Human Neuroscience*, *9*. https://doi.org/10.3389/fnhum.2015.00335

Mantini, D., Corbetta, M., Perrucci, M. G., Romani, G. L., & Del Gratta, C. (2009). Large-scale brain networks account for sustained and transient activity during target detection. *NeuroImage*, *44*(1), 265–274. https://doi.org/10.1016/j.neuroimage.2008.08.019

Marco-Pallarés, J., Cucurell, D., Cunillera, T., Krämer, U. M., Càmara, E., Nager, W., Bauer, P., Schüle, R., Schöls, L., Münte, T. F., & Rodriguez-Fornells, A. (2009). Genetic variability in the dopamine system (dopamine receptor D4, catechol-O-methyltransferase) modulates neurophysiological responses to gains and losses. *Biological Psychiatry*, *66*(2), 154–161. https://doi.org/10.1016/j.biopsych.2009.01.006

Marini, P., Ramat, S., Ginestroni, A., & Paganini, M. (2003). Deficit of short-term memory in newly diagnosed untreated parkinsonian patients: Reversal after L-dopa therapy. *Neurological Sciences: Official Journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology*, *24*(3), 184–185. https://doi.org/10.1007/s10072-003-0121-3

Martínez-Horta, S., Riba, J., Bobadilla, R. F. de, Pagonabarraga, J., Pascual-Sedano, B., Antonijoan, R. M., Romero, S., Mañanas, M. À., García-Sanchez, C., & Kulisevsky, J. (2014). Apathy in Parkinson's Disease: Neurophysiological Evidence of Impaired Incentive Processing. *Journal of Neuroscience*, *34*(17), 5918–5926. https://doi.org/10.1523/JNEUROSCI.0251-14.2014

Mastrigt, N. M. van, Smeets, J. B. J., & Kooij, K. van der. (2020). Quantifying exploration in reward-based motor learning. *PLOS ONE*, *15*(4), e0226789. https://doi.org/10.1371/journal.pone.0226789

Mathewson, K. J., Dywan, J., Snyder, P. J., Tays, W. J., & Segalowitz, S. J. (2008). Aging and electrocortical response to error feedback during a spatial learning task. *Psychophysiology*, *45*(6), 936–948. https://doi.org/10.1111/j.1469-8986.2008.00699.x

McDougle, S. D., Ivry, R. B., & Taylor, J. A. (2016). Taking Aim at the Cognitive Side of Learning in Sensorimotor Adaptation Tasks. *Trends in Cognitive Sciences*, *20*(7), 535–544. https://doi.org/10.1016/j.tics.2016.05.002

Mehler, D. M. A., Reichenbach, A., Klein, J., & Diedrichsen, J. (2017). Minimizing endpoint variability through reinforcement learning during reaching movements involving shoulder, elbow and wrist. *PloS One*, *12*(7), e0180803. https://doi.org/10.1371/journal.pone.0180803

Miall, R. C., Christensen, L. O. D., Cain, O., & Stanley, J. (2007). Disruption of State Estimation in the Human Lateral Cerebellum. *PLOS Biology*, *5*(11), e316. https://doi.org/10.1371/journal.pbio.0050316

Miltner, W. H., Braun, C. H., & Coles, M. G. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: Evidence for a "generic" neural system for error detection. *Journal of Cognitive Neuroscience*, *9*(6), 788–798. https://doi.org/10.1162/jocn.1997.9.6.788

Montgomery, D. C., Peck, E. A., & Vining, G. G. (2021). *Introduction to Linear Regression Analysis*. John Wiley & Sons.

Moustafa, A. A., Sherman, S. J., & Frank, M. J. (2008). A dopaminergic basis for working memory, learning and attentional shifting in Parkinsonism. *Neuropsychologia*, *46*(13), 3144–3156. https://doi.org/10.1016/j.neuropsychologia.2008.07.011

Mueller, E. M., Burgdorf, C., Chavanon, M.-L., Schweiger, D., Hennig, J., Wacker, J., & Stemmler, G. (2014). The COMT Val158Met polymorphism regulates the effect of a dopamine antagonist on the feedback-related negativity. *Psychophysiology*, *51*(8), 805–809. https://doi.org/10.1111/psyp.12226

Mulert, C., Juckel, G., Giegling, I., Pogarell, O., Leicht, G., Karch, S., Mavrogiorgou, P., Möller, H.-J., Hegerl, U., & Rujescu, D. (2006). A Ser9Gly polymorphism in the dopamine D3 receptor gene (DRD3) and event-related P300 potentials. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, *31*(6), 1335–1344. https://doi.org/10.1038/sj.npp.1300984

Nikooyan, A. A., & Ahmed, A. A. (2014). Reward feedback accelerates motor learning. *Journal of Neurophysiology*, *113*(2), 633–646. https://doi.org/10.1152/jn.00032.2014

Noble, E. P., Berman, S. M., Ozkaragoz, T. Z., & Ritchie, T. (1994). Prolonged P300 latency in children with the D2 dopamine receptor A1 allele. *American Journal of Human Genetics*, *54*(4), 658–668.

Olanow, W., Schapira, A. H., & Rascol, O. (2000). Continuous dopamine-receptor stimulation in early Parkinson's disease. *Trends in Neurosciences*, *23*(10 Suppl), S117-126. https://doi.org/10.1016/s1471-1931(00)00030-6

Onur, Ö. A., Piefke, M., Lie, C.-H., Thiel, C. M., & Fink, G. R. (2011). Modulatory effects of levodopa on cognitive control in young but not in older subjects: A pharmacological fMRI study. *Journal of Cognitive Neuroscience*, *23*(10), 2797–2810. https://doi.org/10.1162/jocn.2011.21603

Otani, S., Daniel, H., Roisin, M.-P., & Crepel, F. (2003). Dopaminergic modulation of long-term synaptic plasticity in rat prefrontal neurons. *Cerebral Cortex (New York, N.Y.: 1991)*, *13*(11), 1251–1256. https://doi.org/10.1093/cercor/bhg092

Oyama, K., Hernádi, I., Iijima, T., & Tsutsui, K.-I. (2010). Reward Prediction Error Coding in Dorsal Striatal Neurons. *Journal of Neuroscience*, *30*(34), 11447–11457. https://doi.org/10.1523/JNEUROSCI.1719-10.2010

Pagnoni, G., Zink, C. F., Montague, P. R., & Berns, G. S. (2002). Activity in human ventral striatum locked to errors of reward prediction. *Nature Neuroscience*, *5*(2), 97–98. https://doi.org/10.1038/nn802

Palidis, D. J., Cashaback, J. G. A., & Gribble, P. L. (2019). Neural signatures of reward and sensory error feedback processing in motor learning. *Journal of Neurophysiology*, *121*(4), 1561–1574. https://doi.org/10.1152/jn.00792.2018

Pekny, S. E., Izawa, J., & Shadmehr, R. (2015). Reward-dependent modulation of movement variability. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *35*(9), 4015–4024. https://doi.org/10.1523/JNEUROSCI.3244-14.2015

Polich, J. (2007). Updating P300: An Integrative Theory of P3a and P3b. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, *118*(10), 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019

Quattrocchi, G., Monaco, J., Ho, A., Irmen, F., Strube, W., Ruge, D., Bestmann, S., & Galea, J. M. (2018). Pharmacological Dopamine Manipulation Does Not Alter Reward-Based Improvements in Memory Retention during a Visuomotor Adaptation Task. *ENeuro*, *5*(3). https://doi.org/10.1523/ENEURO.0453-17.2018

Ragazzoni, A., Di Russo, F., Fabbri, S., Pesaresi, I., Di Rollo, A., Perri, R. L., Barloscio, D., Bocci, T., Cosottini, M., & Sartucci, F. (2019). "Hit the missing stimulus". A simultaneous EEG-fMRI study to localize the generators of endogenous ERPs in an omitted target paradigm. *Scientific Reports*, *9*(1), 3684. https://doi.org/10.1038/s41598-019-39812-z

Rand, M. K., & Heuer, H. (2019). Visual and proprioceptive recalibrations after exposure to a visuomotor rotation. *European Journal of Neuroscience*, *50*(8), 3296–3310. https://doi.org/10.1111/ejn.14433

Reynolds, J. N., & Wickens, J. R. (2002). Dopamine-dependent plasticity of corticostriatal synapses. *Neural Networks*, *15*(4), 507–521. https://doi.org/10.1016/S0893-6080(02)00045-X

Rothenhoefer, K. M., Costa, V. D., Bartolo, R., Vicario-Feliciano, R., Murray, E. A., & Averbeck, B. B. (2017). Effects of Ventral Striatum Lesions on Stimulus-Based versus Action-Based Reinforcement Learning. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *37*(29), 6902–6914. https://doi.org/10.1523/JNEUROSCI.0631-17.2017

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, *56*(5), 356–374. https://doi.org/10.1016/j.jmp.2012.08.001

Russek, E. M., Momennejad, I., Botvinick, M. M., Gershman, S. J., & Daw, N. D. (2017). Predictive representations can link model-based reinforcement learning to model-free mechanisms. *PLoS Computational Biology*, *13*(9), e1005768. https://doi.org/10.1371/journal.pcbi.1005768

Rutledge, R. B., Dean, M., Caplin, A., & Glimcher, P. W. (2010). Testing the reward prediction error hypothesis with an axiomatic model. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *30*(40), 13525–13536. https://doi.org/10.1523/JNEUROSCI.1747-10.2010

Sabeti, M., Katebi, S. D., Rastgar, K., & Azimifar, Z. (2016). A multi-resolution approach to localize neural sources of P300 event-related brain potential. *Computer Methods and Programs in Biomedicine*, *133*, 155–168. https://doi.org/10.1016/j.cmpb.2016.05.013

Sambrook, T. D., & Goslin, J. (2015). A neural reward prediction error revealed by a meta-analysis of ERPs using great grand averages. *Psychological Bulletin*, *141*(1), 213–235. http://dx.doi.org.proxy1.lib.uwo.ca/10.1037/bul0000006

Sambrook, T. D., & Goslin, J. (2016). Principal components analysis of reward prediction errors in a reinforcement learning task. *NeuroImage*, *124*, 276–286. https://doi.org/10.1016/j.neuroimage.2015.07.032

Sambrook, T. D., Hardwick, B., Wills, A. J., & Goslin, J. (2018). Model-free and model-based reward prediction errors in EEG. *NeuroImage*, *178*, 162–171. https://doi.org/10.1016/j.neuroimage.2018.05.023

Santesso, D. L., Evins, A. E., Frank, M. J., Schetter, E. C., Bogdan, R., & Pizzagalli, D. A. (2009). Single dose of a dopamine agonist impairs reinforcement learning in humans: Evidence from event-related potentials and computational modeling of striatal-cortical function. *Human Brain Mapping*, *30*(7), 1963–1976.

Schultz, W. (2016). Dopamine reward prediction error coding. *Dialogues in Clinical Neuroscience*, *18*(1), 23–32.

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

Schultz, W., Tremblay, L., & Hollerman, J. R. (1998). Reward prediction in primate basal ganglia and frontal cortex. *Neuropharmacology*, *37*(4), 421–429. https://doi.org/10.1016/S0028-3908(98)00071-9

Schutte, I., Deschamps, P. K. H., van Harten, P. N., & Kenemans, J. L. (2020). Dopaminergic and noradrenergic manipulation of anticipatory reward and probability event-related potentials. *Psychopharmacology*, *237*(7), 2019–2030. https://doi.org/10.1007/s00213-020-05515-x

Shadmehr, R., Smith, M. A., & Krakauer, J. W. (2010). Error correction, sensory prediction, and adaptation in motor control. *Annual Review of Neuroscience*, *33*, 89–108. https://doi.org/10.1146/annurev-neuro-060909-153135

Shahar, N., Moran, R., Hauser, T. U., Kievit, R. A., McNamee, D., Moutoussis, M., NSPN Consortium, & Dolan, R. J. (2019). Credit assignment to state-independent task representations and its relationship with model-based decision making. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(32), 15871–15876. https://doi.org/10.1073/pnas.1821647116

Sharp, M. E., Foerde, K., Daw, N. D., & Shohamy, D. (2016). Dopamine selectively remediates "model-based" reward learning: A computational approach. *Brain: A Journal of Neurology*, *139*(Pt 2), 355–364. https://doi.org/10.1093/brain/awv347

Shmuelof, L., Huang, V. S., Haith, A. M., Delnicki, R. J., Mazzoni, P., & Krakauer, J. W. (2012). Overcoming motor "forgetting" through reinforcement of learned actions. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *32*(42), 14617–14621. https://doi.org/10.1523/JNEUROSCI.2184-12.2012

Sidarta, A., Vahdat, S., Bernardi, N. F., & Ostry, D. J. (2016). Somatic and Reinforcement-Based Plasticity in the Initial Stages of Human Motor Learning. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *36*(46), 11682–11692. https://doi.org/10.1523/JNEUROSCI.1767-16.2016

Sidarta, A., van Vugt, F. T., & Ostry, D. J. (2018). Somatosensory working memory in human reinforcement-based motor learning. *Journal of Neurophysiology*, *120*(6), 3275–3286. https://doi.org/10.1152/jn.00442.2018

Silvetti, M., Nuñez Castellar, E., Roger, C., & Verguts, T. (2014). Reward expectation and prediction error in human medial frontal cortex: An EEG study. *NeuroImage*, *84*, 376–382. https://doi.org/10.1016/j.neuroimage.2013.08.058

Smith, M. A., & Shadmehr, R. (2005). Intact Ability to Learn Internal Models of Arm Dynamics in Huntington's Disease But Not Cerebellar Degeneration. *Journal of Neurophysiology*, *93*(5), 2809–2821. https://doi.org/10.1152/jn.00943.2004

Sohn, Y. H., Kim, G. W., Huh, K., & Kim, J. S. (1998). Dopaminergic influences on the P300 abnormality in Parkinson's disease. *Journal of the Neurological Sciences*, *158*(1), 83–87. https://doi.org/10.1016/s0022-510x(98)00102-6

Soltani, M., & Knight, R. T. (2000). Neural origins of the P300. *Critical Reviews in Neurobiology*, *14*(3–4), 199–224.

Stanzione, P., Fattapposta, F., Giunti, P., D'Alessio, C., Tagliati, M., Affricano, C., & Amabile, G. (1991). P300 variations in parkinsonian patients before and during dopaminergic monotherapy: A suggested dopamine component in P300. *Electroencephalography and Clinical Neurophysiology*, *80*(5), 446–453. https://doi.org/10.1016/0168-5597(91)90093-d

Stanzione, P., Fattapposta, F., Tagliati, M., D'Alessio, C., Marciani, M. G., Foti, A., & Amabile, G. (1990). Dopaminergic pharmacological manipulations in normal humans confirm the specificity of the visual (PERG-VEP) and cognitive (P300) electrophysiological alterations in Parkinson's disease. *Electroencephalography and Clinical Neurophysiology. Supplement*, *41*, 216–220. https://doi.org/10.1016/b978-0-444-81352-7.50026-1

Sutton, R. S., & Barto, A. G. (2011). *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press.

Swainson, R., Rogers, R. D., Sahakian, B. J., Summers, B. A., Polkey, C. E., & Robbins, T. W. (2000). Probabilistic learning and reversal deficits in patients with Parkinson's disease or frontal or temporal lobe lesions: Possible adverse effects of dopaminergic medication. *Neuropsychologia*, *38*(5), 596–612. https://doi.org/10.1016/s0028-3932(99)00103-7

Synofzik, M., Lindner, A., & Thier, P. (2008). The cerebellum updates predictions about the visual consequences of one's behavior. *Current Biology: CB*, *18*(11), 814–818. https://doi.org/10.1016/j.cub.2008.04.071

Takeshita, S., & Ogura, C. (1994). Effect of the dopamine D2 antagonist sulpiride on event-related potentials and its relation to the law of initial value. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *16*(1), 99–106. https://doi.org/10.1016/0167-8760(94)90046-9

Tanaka, H., Sejnowski, T. J., & Krakauer, J. W. (2009). Adaptation to Visuomotor Rotation Through Interaction Between Posterior Parietal and Motor Cortical Areas. *Journal of Neurophysiology*, *102*(5), 2921–2932. https://doi.org/10.1152/jn.90834.2008

Taylor, J. A., Klemfuss, N. M., & Ivry, R. B. (2010). An explicit strategy prevails when the cerebellum fails to compute movement errors. *Cerebellum (London, England)*, *9*(4), 580–586. https://doi.org/10.1007/s12311-010-0201-x

Taylor, J. A., Krakauer, J. W., & Ivry, R. B. (2014). Explicit and Implicit Contributions to Learning in a Sensorimotor Adaptation Task. *Journal of Neuroscience*, *34*(8), 3023–3032. https://doi.org/10.1523/JNEUROSCI.3619-13.2014

Therrien, A. S., & Bastian, A. J. (2015). Cerebellar damage impairs internal predictions for sensory and motor function. *Current Opinion in Neurobiology*, *33*, 127–133. https://doi.org/10.1016/j.conb.2015.03.013

Therrien, A. S., Wolpert, D. M., & Bastian, A. J. (2016). Effective reinforcement learning following cerebellar damage requires a balance between exploration and motor noise. *Brain*, *139*(1), 101–114. https://doi.org/10.1093/brain/awv329

Torta, D. M. E., Castelli, L., Zibetti, M., Lopiano, L., & Geminiani, G. (2009). On the role of dopamine replacement therapy in decision-making, working memory, and reward in Parkinson's disease: Does the therapy-dose matter? *Brain and Cognition*, *71*(2), 84–91. https://doi.org/10.1016/j.bandc.2009.04.003

Tseng, Y., Diedrichsen, J., Krakauer, J. W., Shadmehr, R., & Bastian, A. J. (2007). Sensory Prediction Errors Drive Cerebellum-Dependent Adaptation of Reaching. *Journal of Neurophysiology*, *98*(1), 54–62. https://doi.org/10.1152/jn.00266.2007

van der Helden, J., Boksem, M. A. S., & Blom, J. H. G. (2010). The importance of failure: Feedback-related negativity predicts motor learning efficiency. *Cerebral Cortex (New York, N.Y.: 1991)*, *20*(7), 1596–1603. https://doi.org/10.1093/cercor/bhp224

Vezoli, J., & Procyk, E. (2009). Frontal feedback-related potentials in nonhuman primates: Modulation during learning and under haloperidol. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *29*(50), 15675–15683. https://doi.org/10.1523/JNEUROSCI.4943-09.2009

Vo, A., Seergobin, K. N., & MacDonald, P. A. (2017). Effects of levodopa on stimulus-response learning versus response selection in healthy young adults. *Behavioural Brain Research*, *317*, 553–561. https://doi.org/10.1016/j.bbr.2016.10.019

Vo, A., Seergobin, K. N., & MacDonald, P. A. (2018). Independent effects of age and levodopa on reversal learning in healthy volunteers. *Neurobiology of Aging*, *69*, 129–139. https://doi.org/10.1016/j.neurobiolaging.2018.05.014

Vo, A., Seergobin, K. N., Morrow, S. A., & MacDonald, P. A. (2016). Levodopa impairs probabilistic reversal learning in healthy young adults. *Psychopharmacology*, *233*(14), 2753–2763. https://doi.org/10.1007/s00213-016-4322-x

Walsh, M. M., & Anderson, J. R. (2012). Learning from experience: Event-related potential correlates of reward processing, neural adaptation, and behavioral choice. *Neuroscience and Biobehavioral Reviews*, *36*(8), 1870–1884. https://doi.org/10.1016/j.neubiorev.2012.05.008

Wang, J. X., Kurth-Nelson, Z., Kumaran, D., Tirumala, D., Soyer, H., Leibo, J. Z., Hassabis, D., & Botvinick, M. M. (2018). Prefrontal cortex as a meta-reinforcement learning system. *Nature Neuroscience*, *21*(6), 860–868. https://doi.org/10.1038/s41593-018-0147-8

Warren, C. M., Hyman, J. M., Seamans, J. K., & Holroyd, C. B. (2015). Feedback-related negativity observed in rodent anterior cingulate cortex. *Journal of Physiology-Paris*, *109*(1), 87–94. https://doi.org/10.1016/j.jphysparis.2014.08.008

Watabe-Uchida, M., Eshel, N., & Uchida, N. (2017). Neural Circuitry of Reward Prediction Error. *Annual Review of Neuroscience*, *40*, 373–394. https://doi.org/10.1146/annurev-neuro-072116-031109

Wetzels, R., Matzke, D., Lee, M. D., Rouder, J. N., Iverson, G. J., & Wagenmakers, E.-J. (2011). Statistical Evidence in Experimental Psychology: An Empirical Comparison Using 855 t Tests. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *6*(3), 291–298. https://doi.org/10.1177/1745691611406923

Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science*, *269*(5232), 1880–1882. https://doi.org/10.1126/science.7569931

Wong, A. L., Marvel, C. L., Taylor, J. A., & Krakauer, J. W. (2019). Can patients with cerebellar disease switch learning mechanisms to reduce their adaptation deficits? *Brain: A Journal of Neurology*, *142*(3), 662–673. https://doi.org/10.1093/brain/awy334

Wunderlich, K., Smittenaar, P., & Dolan, R. J. (2012). Dopamine enhances model-based over model-free choice behavior. *Neuron*, *75*(3), 418–424. https://doi.org/10.1016/j.neuron.2012.03.042

Chapter 4

Effects of levodopa on force-field adaptation, savings and interference

4.1 Introduction

Human motor control reliably adapts to changes of the body or the environment. For example, neural motor commands must be altered when an athlete uses a new piece of equipment or when the responses of their muscles change due to fatigue. Force field learning is a common experimental model used to study motor adaptation in response to alterations to the dynamics of reaching movements (Shadmehr & Mussa-Ivaldi, 1994). In typical studies of force field adaptation, a robot applies velocity-dependent forces to the hand during reaches to targets. Motor output reliably adapts to reduce the errors caused by these mechanical perturbations, even following a single movement (Diedrichsen et al., 2005; O. Donchin et al., 2003; M. A. Smith et al., 2006; Thoroughman & Shadmehr, 2000). Adaptation quickly washes out when force-field perturbations are removed. However, savings causes adaptation to occur more quickly during re-exposure to a previously encountered perturbation after loss of initial learning (Coltman et al., 2019; Herzfeld et al., 2014; M. A. Smith et al., 2006; Zarahn et al., 2008). Previous adaptation can also cause anterograde interference which slows learning during subsequent exposure to an opposite direction perturbation (Krakauer, 2009; Miall et al., 2004; Shadmehr & Brashers-Krug, 1997; M. A. Smith et al., 2006). Here, we used levodopa to manipulate dopamine release in healthy participants to test the hypothesis that savings and interference depend on dopamine signaling.

Initial adaptation to force field perturbations is thought occur through error-based updating of a neural model of limb dynamics (O. Donchin et al., 2003; Kawato, 1999; McNamee & Wolpert, 2019; Shadmehr et al., 2010). Adaptation depends heavily on the cerebellum (Block & Bastian, 2012; Izawa et al., 2012; Maschke et al., 2004; M. A. Smith & Shadmehr, 2005; Tseng et al., 2007). Activity in both the cerebellum and sensory-motor cortical regions reflects errors induced by force field perturbations and adaptive changes corresponding with behavioral learning (Diedrichsen et al., 2005; C.-S. R. Li et al., 2001; Padoa-Schioppa et al., 2004; Perich et al., 2018; Xiao et al., 2006). One account of savings and interference suggests that these effects are not produced by updating of internal sensory-motor models, but rather by additional influences of a reinforcement learning

process (Huang et al., 2011; Orban de Xivry & Lefèvre, 2015). This hypothesis states that when error-based adaptation counteracts the errors induced by a perturbation, the adapted motor commands are reinforced due to successful performance in the task. When the same perturbation is encountered in the future, those previously reinforced actions may be recalled, resulting in savings. According to this hypothesis, recall of previously successful actions also occurs when an opposite perturbation is introduced. In this case the previous motor solution is incompatible with the novel perturbation, resulting in interference. Unlike the error-based learning process that produces initial adaptation, this hypothesized reinforcement learning process is thought to be "model-free" in the sense that adaptive responses are not planned according to informational content of the error such as direction. Instead, actions that were previously associated with success are simply recalled in response to any perturbation.

In cognitive reward-based learning tasks, model-free reinforcement learning is thought to occur though plasticity at cortico-striatal synapses caused by dopaminergic reward signaling (Daw et al., 2005; Dolan & Dayan, 2013; Doll et al., 2016; Gläscher et al., 2010; Glimcher, 2011; Steinberg et al., 2013). Parkinson's disease (PD) is primarily characterized by loss of dopaminergic projections to the striatum, and people with PD show reduced savings and anterograde interference for motor adaptation despite intact initial learning (Bédard & Sanes, 2011; Leow et al., 2012, 2013; Marinelli et al., 2009). Although these findings suggest that savings and interference may be mediated by dopaminergic reinforcement learning, numerous features of PD could potentially explain the observed deficits. PD is accompanied by significant structural and functional changes in primary motor cortex (Burciu & Vaillancourt, 2018). Learning impairments may also be a side-effect of medication.

Levodopa, a dopamine precursor used to treat PD, has been shown to impair cognitive reward-based learning in both patients and healthy participants (Cools et al., 2001, 2007; Feigin et al., 2003; Frank et al., 2004; Graef et al., 2010; Hiebert et al., 2014; Jahanshahi et al., 2010; Kwak et al., 2010; MacDonald et al., 2011; Swainson et al., 2000; Torta et al., 2009; Vo et al., 2016, 2018). In early-stage Parkinson's disease, the dorsal striatum is

significantly depleted of dopamine whereas the ventral striatum is comparatively spared. According to the "dopamine overdose" hypothesis, levodopa overstimulates dopamine release in the ventral striatum of both healthy controls and people with early stage PD, resulting in learning deficits (Cools et al., 2007). Although levodopa ameliorates deficits caused by dopamine-depletion in the dorsal striatum, the ventral striatum is less dopamine-depleted in early-stage PD and thus may be susceptible to impairment by excessive dopamine release due to levodopa. In line with this view, cognitive reward-based learning is thought to rely on dopamine signaling in the ventral striatum and is impaired by levodopa (Hiebert et al., 2019). Here, we administered levodopa to healthy participants to provide a more specific and controlled test for a role of dopamine in savings and interference. According to the "dopamine overdose" hypothesis, we predicted that levodopa would reduce savings and interference while leaving initial adaptation unaffected.

In chapter 3, we describe findings that levodopa did not affect motor adaptation induced by explicit binary reinforcement feedback. One possible explanation is that levodopa impairs model-free reinforcement learning while reward-based motor learning depends on different processes. Reward-based motor adaptation is abolished by cognitive load and depends on executive functions such as explicit strategy, planning, and working memory (Holland et al., 2018, 2019; Sidarta et al., 2018). These features of reward-based motor adaptation are generally inconsistent with model-free learning which is thought to produce automatic, habitual response tendencies that do not depend on executive function. Because savings and anterograde interference are hypothesized to depend on model-free reinforcement learning, we predicted that they might be impaired by levodopa despite the fact that reward-based motor adaptation is not.

Participants ingested either levodopa or placebo prior to performing a force field adaptation task. Participants made reaching movements towards visual targets while holding a handle attached to a robotic manipulandum. During reaches when a force-field was active, the robot applied forces perpendicular to the direction of movement and proportional to the velocity of the hand. Participants first performed a block of baseline

reaches during which the robot applied no forces. A clockwise force field was applied in the second block of reaches, followed by a third block of reaches with no force field to allow for washout of initial adaptation. A clockwise force field was applied again during the fourth block to test for savings. We expected that adaptation to compensate for the forces produced by the robot would be facilitated during the second exposure to the clockwise force field due to savings. We hypothesized that levodopa would reduce savings by disrupting dopaminergic learning mechanisms. In the final block, participants adapted to a counterclockwise force field. We expected that adaptation would be impaired in this block by anterograde interference due to the previous exposure to an opposite force field. We hypothesized that levodopa would disrupt anterograde interference and thus improve adaptation in this final block.

4.2 Materials and Methods

4.2.1 Participants

A total of 38 participants were tested (Table 4-1). All participants were screened for neurological and psychiatric illness, history of drug or alcohol abuse, and contraindications for levodopa. Participants provided written informed consent to experimental procedures approved by the Research Ethics Board at Western University.

4.2.2 Procedure

Drug administration: Participants were administered either levodopa or placebo in a randomized double-blind design. A capsule was ingested that contained 100 mg of levodopa (L-3,4-dihydroxyphenylalanine) and 25 mg of carbidopa or an equal volume of placebo. The experimental tasks began 45 minutes after ingestion of the capsule to coincide with peak plasma levels of levodopa. We measured subjective alertness using the Bond-Lader visual analog scale (Bond & Lader, 1974) as well as heart rate and blood pressure immediately prior to ingesting the capsule and again at the end of each session.

Measure	Placebo	Levodopa
n	19	19
n female	9	10
Age	21.2 (SD: 2.5)	22.2 (SD: 3.4 years)
Heart Rate	Pre: 75.1 (SD: 9.5)	Pre: 71.6842 (SD: 12.8)
	Post: 66.2 (SD: 10.2)	Post: 65.7 (SD: 11.3)
Systolic	Pre: 109.2 (SD: 15.4)	Pre: 108.4 (SD: 11.4)
	Post: 104.8 (SD: 14.5)	Post: 99.7 (SD: 10.1)
Diastolic	Pre: 72.0 (SD: 10.2)	Pre: 73.2 (SD: 15.5)
	Post: 70.1 (SD: 10.2)	Post: 67.0 (SD: 8.2)
Alertness	Pre: 31.3 (SD: 15.3)	Pre: 27.1 (SD: 11.0)
	Post: 39.4 (SD: 17.0)	Post: 43.4 (SD: 12.7)
Peak Velocity	0.43 (SD = 0.01)	0.43 (SD = 0.02)

Table 4-1: Control measurements. Heart rate (bpm). Systolic blood pressure (mm Hg). Diastolic blood pressure (mm Hg). Alertness, Bond-Lader visual analog scale alertness measure. Peak Velocity, maximum tangential velocity of the hand averaged across trials (m/s).

Force field adaptation task: Participants produced reaching movements with their right arm while holding the handle of a robotic arm (InMotion2; Interactive Motion Technologies). The position of the robot handle was sampled at 600 Hz. A semi-silvered mirror obscured vision of the arm and displayed visual information related to the task. An air sled supported each participant's right arm.

On each trial, participants reached from a central home position (blue circle 20 mm in diameter) to one of 8 circular targets (24 mm in diameter) arranged around the home position at a distance of 10 cm. The target angles were 0°, 45°, 90°, 135°, 180°, 225°, 270°, and 315°. A 5-mm pink circular cursor represented the position of the robot handle. When the cursor reached the target on each trial, the target either turned blue to indicate that the movement duration was satisfactory (375 ± 100 ms), green to indicate that the movement was too slow, or red to indicate that the movement was too fast. The subject moved the robot handle back to the home position at the end of each reach.

In null field blocks, the robot motors did not apply any external forces to the hand. In force field blocks, the robot applied forces to the hand that were perpendicular to the direction of movement and proportional to the velocity of the hand (*eq. 4.1)*. The direction of the force field was either clockwise or counterclockwise, in separate blocks.

$$\begin{bmatrix} F_x \\ F_y \end{bmatrix} = b \begin{bmatrix} 0 & d \\ -d & 0 \end{bmatrix} \begin{bmatrix} v_x \\ v_y \end{bmatrix} \quad (4.1)$$

x and y correspond to the lateral and sagittal directions. F_x and F_y describe the forces applied to the hand, v_x and v_y describe the velocity of the hand, b is the field constant, and d corresponds to the direction (d = 1 for a clockwise force field (CWFF), -1 for a counterclockwise force field (CCWFF) or 0 for a null field (NF)).

All participants completed five blocks of 96 trials. Each block consisted of 12 reaches to each of the 8 targets presented in random order. The five blocks occurred in the following order: NFa (null field), FF1a (CWFF), NFb (null field), FF1b (CWFF), FF2 (CCWFF). Trials 6, 24, 35, 50, 71, and 91 of each block were "catch trials", during which reaches

occurred in a null field. When a force field is suddenly removed in catch trials, adaptation causes errors to occur in the opposite direction of the force field. A reduction in reach error during force field trials may reflect either adaptation to the force field, stiffening of the arm, or changes in feedback corrections. The magnitude of errors opposite the force field in catch trials is thought to better capture adaptation of feedforward control. Similar to catch trials, we expected after-effects at the beginning of NFa in the form of counterclockwise reach errors after the sudden removal of the clockwise force field in FF1a.

4.2.3 Data analysis

Robot handle positional data were low-pass filtered with a 40 Hz cutoff frequency and differentiated to yield instantaneous velocity and acceleration. On each trial, movement onset and end of movement were defined according to a velocity threshold set at 5% of the maximum tangential velocity of the robot endpoint. Our behavioral measure of interest was the lateral deviation of the hand at the time of peak tangential velocity. Perpendicular deviation (PD) was calculated relative to a line drawn from the position of movement onset in the direction of the target angle (either 0°, 45°, 90°, 135°, 180°, 225°, 270°, or 315°). PD was calculated for each trial as the perpendicular distance between the position of the hand at peak velocity and this line, with positive PD corresponding to clockwise deviations. For non-catch trials, PD was averaged across trials within 12 bins of 8 trials each. We analyzed effects related to adaptation separately for an early and late period of each block. The early period consisted of the first 5 bins (trials 1-40, catch trials: 6,24,35) and the late period consisted of the remaining 7 bins (trials 41-96, catch trials: 50,71,91). Baseline PD was computed as the average PD in the late period of NFa. We computed metrics for adaptation, savings, after-effects, and learning with interference separately for the early and late periods, and separately for catch trials and non-catch trials. All metrics were computed so that positive values corresponded to the effects of interest, and values of zero correspond to no effect. We tested for adaptation, savings, after-effects, and learning with interference using 1-sample t-tests against zero. We tested for differences between the placebo and levodopa groups using paired t-tests.

Non-catch trials: Adaptation metrics were computed to capture reductions in error during FF1a relative to the initial errors caused by the onset of the force field. Our measure of early adaptation was the average PD in the first bin of FF1a minus the average PD across subsequent bins within the early period of FF1a (bins 2-5). Our measure of late adaptation was the average PD in the first bin of FF1a minus the average PD across bins in the late period of FF1a (bins 6-12). Savings metrics were computed to measure reductions in errors during the second exposure to FF1 compared to the first. Savings was measured as the difference in PD between FF1a and FF1b (FF1a – FF1b), separately for PD averaged across bins within the early and late periods. Adaptation to FF1a caused after-effects in the form of errors upon its sudden removal at the onset of NFb. After-effects were measured as the difference between baseline PD and the PD in NFb (baseline – NFb), separately for PD averaged across bins in the early and late periods of NFb. We expected large initial errors at the onset of FF2 due to a combination of after-effects from the removal of FF1b and the introduction of a novel force field. Previous adaptation to FF1b was also expected to cause anterograde interference during adaptation to FF2 as the force fields were opposite. Metrics for adaptation with interference were computed to capture reductions in errors during FF2 relative to the initial errors caused by the onset of the force field. Early adaptation with interference was measured by subtracting the average PD from the first bin of FF2 from the average PD across subsequent bins within the early period of FF2 (bins 2-5). Late adaptation with interference was measured by subtracting the average PD in the first bin of FF2 from the average PD across subsequent bins in the late period of FF2 (bins 6-12).

Catch trials: When a force field is suddenly removed during catch trials, adaptation to the force field is reflected in errors opposite the direction of the force field. Adaptation effects were computed as the baseline PD minus the PD in FF1a averaged across catch trials, separately for catch trials in the early and late period. Improved adaptation due to savings was expected to cause larger errors in catch trial during FF1b compared to FF1a. Savings was computed as the PD in FF1a minus the PD in FF1b, averaged across catch trials separately for the early and late periods. Learning effects with interference were

computed using data from FF2. There was no suitable baseline PD to analyze learning in this block. Instead, the PD of the first catch trial was subtracted from the PD of each of the later catch trials, separately for catch trials in the early and late periods. This captures changes in catch trial PD opposite the direction of FF2 due to adaptation.

4.2.4 Statistics

Statistical tests were implemented using JASP v0.14.1. We compared sample means using 1 sample T-Tests and independent sample T-Tests. These comparisons allowed us to compute one-tailed Bayes factors representing $p(data|H_+) \,/\, p(data|H_0)$, where H_0 represents the null hypothesis corresponding to the standard *t*-distribution for an effect size of 0, and H_+ represents the alternative hypothesis corresponding to a *t*-distribution constructed using a one-tailed prior distribution of effect sizes. The use of 1-tailed priors is recommended in the case of directional hypotheses to provide "a fairer balance between the ability to provide evidence for *H*0 and H_+" (Keysers et al., 2020). We used the default effect size priors implemented in JASP (Cauchy scale 0.707). These priors are generally appropriate for effect sizes typical of neuroscience research, and the use of default priors is recommended for standardized and objective analysis (Keysers et al., 2020; Rouder et al., 2012; Wetzels et al., 2011). Bayesian estimates of effect size are reported as median posterior Cohen's δ with 95% credibility interval using 2-tailed priors for H_+ to avoid biasing the estimate in the expected direction. We also report T-statistics, p-values, and 95% confidence intervals generated using 2-tailed frequentist T-Tests. Bayes factors >3 and >10 were taken as moderate and strong evidence in favor of the alternative hypothesis, respectively. Bayes factors <1/3 and <1/10 were taken as moderate and strong evidence in favor of the null hypothesis, respectively. Bayes factors between 1/3 and 3 were taken as inconclusive evidence (Keysers et al., 2020).

Directional priors used for alternative hypotheses specified our predictions that learning metrics would be greater than zero (Force field adaptation, savings, after-effects, and adaptation with interference). In comparing placebo and levodopa conditions, our alternative hypotheses specified that learning metrics would be lower in levodopa

conditions than placebo conditions, in accordance with the "dopamine overdose" hypothesis. The only exception was that we predicted adaptation with interference would be increased by levodopa. If anterograde interference is caused by dopaminergic reinforcement learning, then the "dopamine overdose" effect should reduce interference and facilitate adaptation. All other Bayes factors are computed with 2-tailed priors, as they were conducted without directional a priori hypotheses (control measures, etc.).

4.3 Results

4.3.1 Control measures

Control measures: Participants' judgments as to whether they received placebo or drug was near chance level (52.63%) and only 13.16% of participants responded that they thought they had received the drug. The values for heart rate, blood pressure, and alertness are reported in Table 4-1 for both the placebo and levodopa groups at the beginning and end of each experimental session. There were no reliable differences between the levodopa and placebo conditions in the percent change of heart rate ($t(36) = -1.09$, $p=0.282$, 95%CI for difference = [-0.10 0.03], BF = 0.5, posterior δ: median = -0.273, 95%CI = [-0.875 0.284]), diastolic blood pressure ($t(36) = 1.37$, $p=0.18$, 95%CI for difference = [-0.02 0.11], BF = 0.65, posterior δ: median = 0.346, 95%CI = [-0.218 0.960]), systolic blood pressure ($t(36) = 1.37$, $p=0.18$, 95%CI for difference = [-.02 0.09], BF = 0.65, posterior δ: median = 0.346, 95%CI = [-0.218 0.960]), or alertness ($t(36) = -0.88$, $p=0.39$, 95%CI for difference = [-0.95 0.38], BF = 0.43, posterior δ: median = -0.218, 95%CI = [-0.810 0.337]). There was also no reliable difference between peak movement velocity between the levodopa and placebo groups ($t(36) = -0.09$, $p=0.93$, 95%CI for difference = [-0.01 9.94e-3], BF = 0.32, posterior δ: median = -0.021, 95%CI = [-0.585 0.539]).

4.3.2 Force field adaptation results

In each trial, we measured the perpendicular deviation (PD) of the reach trajectory at peak tangential velocity. PD data from throughout each force field and null field block, excluding catch trials, are shown in Figure 4-1. PD data from catch trials are shown in

Figure 4-2. We computed contrasts to test for adaptation, savings, after-effects, and learning with interference in both the early (bins 1-5) and late (bins 6-12) periods following perturbation onset (Figure 4-3). We tested whether these effects are different from zero using 1-sample T-Tests for both the levodopa and placebo groups. We tested for differences between the levodopa and placebo groups using independent sample T-Tests. Detailed statistical results are shown in Table 4-2.

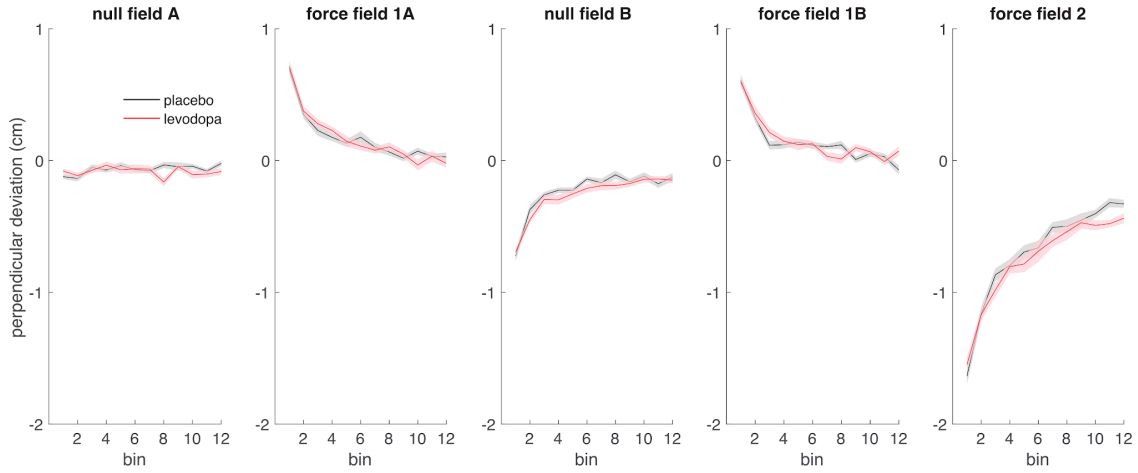

Figure 4-1: Perpendicular deviation of reach trajectory during non-catch trials. Average perpendicular deviation of the hand trajectory within bins consisting of 8 trials each is shown in cm (Shaded region: ± SEM). The placebo condition is shown in black (n=19), and the levodopa condition is shown in red (n=19). Perpendicular deviation was measured on each trial at peak tangential velocity. Trials 6, 24, 35, 50, 71, and 91 of each block were catch trials, and were excluded from the corresponding bins. In *null field A* and *null field B,* the robot did not apply external forces to the hand during reaches. In *force field 1A* and *force field 1B,* participants made reaches in a clockwise force field. In *force field 2* participants made reaches in a counterclockwise force field.

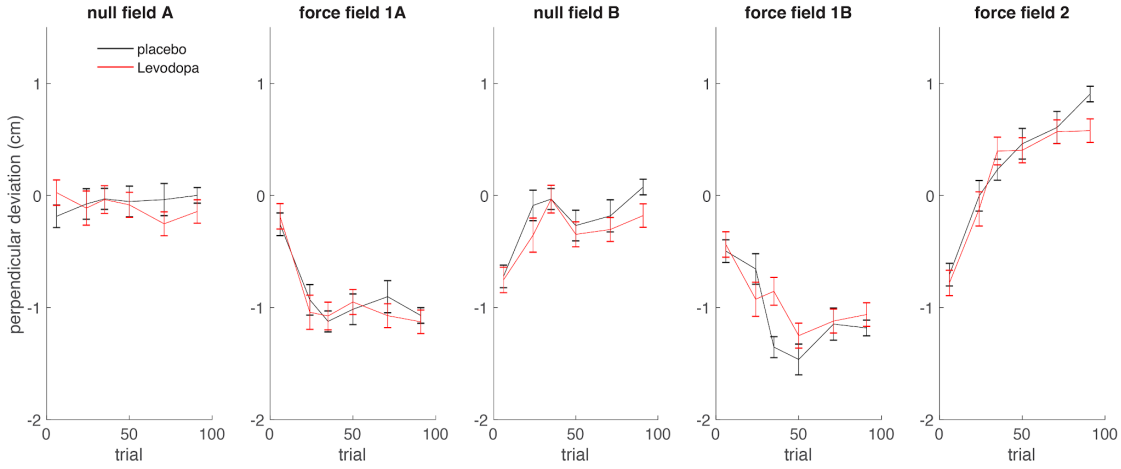

Figure 4-2: Perpendicular deviation of reach trajectory during catch trials. Perpendicular deviation of the hand trajectory, measured at peak tangential velocity, is shown in cm (Error bars: ± SEM). The placebo condition is shown in black (n=19), and the levodopa condition is shown in red (n=19). Catch trials occurred on trials 6, 24, 35, 50, 71, and 91 of each block. In *null field A* and *null field B,* the robot did not apply external forces to the hand during reaches. In *force field 1A* and *force field 1B,* participants made reaches in a clockwise force field. In *force field 2* participants made reaches in a counterclockwise force field.

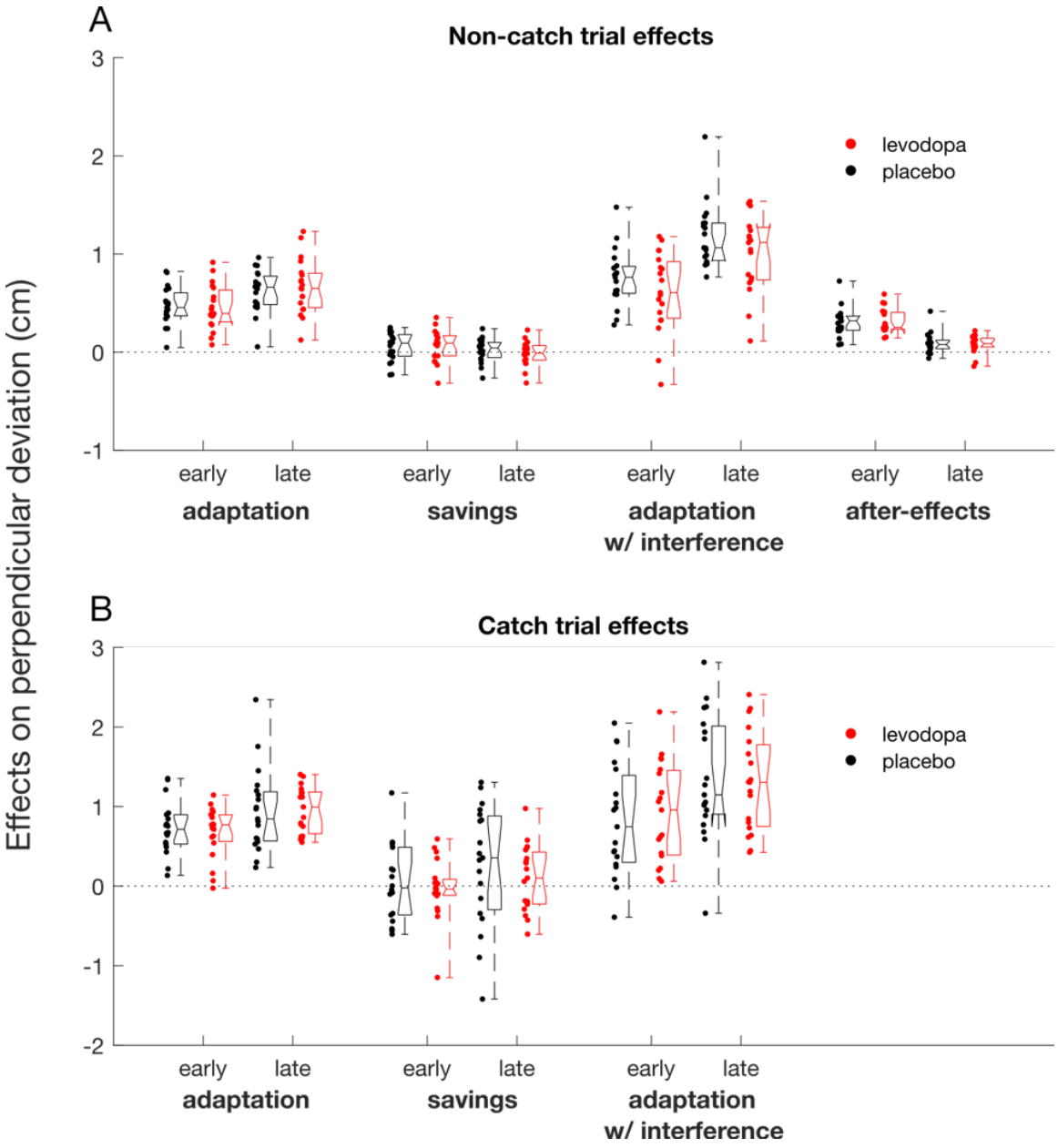

Figure 4-3: Adaptation effects in non-catch trials **A**, and catch trials **B**. Data points show effects for individual participants, box plots show the median, interquartile range, and full range. Effects are contrasts computed using perpendicular deviation (PD) of reach trajectory (cm), such that zero corresponds to no effect. Adaptation: change in PD during FF1a. Savings: difference in PD between FF1a and FF1b. After-effects: difference in PD between NFb and baseline from NFa. Adaptation w/ interference: change in PD during FF2.

One Sample T-Tests

						95% CI for Sample Mean	
Placebo Non-Catch Trials	**t**	**df**	**p**	**BF**	**Sample Mean**	**Lower**	**Upper**
early adaptation	10.60	18	3.62e -9	6.36e +6	0.47	0.38	0.56
late adaptation	12.54	18	2.48e -10	7.72e +7	0.62	0.52	0.73
early savings	1.56	18	0.14	1.23	0.05	-0.02	0.13
late savings	0.70	18	0.50	0.44	0.02	-0.04	0.08
early after-effects	9.06	18	4.00e -8	688519.55	0.31	0.24	0.38
late after-effects	3.78	18	1.37e -3	56.24	0.09	0.04	0.14
early adaptation (interference)	11.46	18	1.06e -9	2.00e +7	0.76	0.62	0.89
late adaptation (interference)	15.70	18	5.98e -12	2.51e +9	1.18	1.02	1.34
Levodopa Non-Catch Trials	**t**	**df**	**p**	**BF**	**Sample Mean**	**Lower**	**Upper**
early adaptation	8.76	18	6.61e -8	432847.50	0.46	0.35	0.57
late adaptation	10.42	18	4.71e -9	4.99e +6	0.67	0.53	0.80
early savings	1.67	18	0.11	1.43	0.06	-0.02	0.14
late savings	-0.40	18	0.70	0.18	-0.01	-0.07	0.05
early after-effects	10.84	18	2.56e -9	8.79e +6	0.31	0.25	0.37
late after-effects	3.94	18	9.66e -4	76.15	0.08	0.04	0.12
early adaptation (interference)	6.59	18	3.42e -6	11657.42	0.61	0.42	0.81
late adaptation (interference)	11.12	18	1.70e -9	1.28e +7	1.02	0.83	1.21
Placebo Catch Trials	**t**	**df**	**p**	**BF**	**Sample Mean**	**Lower**	**Upper**
early adaptation	9.25	17	4.82e -8	574167.17	0.73	0.57	0.90
late adaptation	7.90	18	2.92e -7	110521.90	0.94	0.69	1.19
early savings	0.36	17	0.72	0.33	0.04	-0.20	0.28
late savings	1.54	18	0.14	1.20	0.27	-0.10	0.63
early adaptation (interference)	5.17	18	6.37e -5	837.09	0.82	0.49	1.15
late adaptation (interference)	7.68	18	4.33e -7	77010.30	1.36	0.99	1.74
Levodopa Catch Trials	**t**	**df**	**p**	**BF**	**Sample Mean**	**Lower**	**Upper**
early adaptation	9.12	18	3.62e -8	755029.63	0.68	0.52	0.83
late adaptation	14.40	18	2.54e -11	6.48e +8	0.96	0.82	1.10
early savings	-0.33	18	0.75	0.19	-0.03	-0.21	0.15
late savings	0.99	18	0.33	0.60	0.09	-0.11	0.29
early adaptation (interference)	6.42	18	4.84e -6	8524.02	0.92	0.62	1.22
late adaptation (interference)	9.03	18	4.20e -8	657919.38	1.30	1.00	1.60

Independent Samples T-Tests

						95% CI for Mean Difference	
Placebo vs Levodopa Non-Catch Trials	**t**	**df**	**p**	**BF**	**Mean Diff.**	**Lower**	**Upper**
early adaptation	0.22	36	0.83	0.37	0.01	-0.12	0.15
late adaptation	-0.57	36	0.57	0.22	-0.05	-0.21	0.12
early savings	-0.16	36	0.87	0.28	-7.98e -3	-0.11	0.09
late savings	0.77	36	0.45	0.59	0.03	-0.05	0.11
early after-effects	0.02	36	0.99	0.32	8.12e -4	-0.09	0.09
late after-effects	0.28	36	0.78	0.39	8.67e -3	-0.05	0.07
early adaptation (interference)	1.23	36	0.23	0.16	0.14	-0.09	0.37
late adaptation (interference)	1.38	36	0.18	0.15	0.16	-0.08	0.40
Placebo vs Levodopa Catch Trials	**t**	**df**	**p**	**BF**	**Mean Diff.**	**Lower**	**Upper**
early adaptation	0.51	35	0.61	0.47	0.06	-0.16	0.28
late adaptation	-0.12	36	0.90	0.29	-0.02	-0.29	0.26
early savings	0.49	35	0.63	0.47	0.07	-0.22	0.36
late savings	0.87	36	0.39	0.66	0.17	-0.23	0.57
early adaptation (interference)	-0.46	36	0.65	0.45	-0.10	-0.53	0.33
late adaptation (interference)	0.29	36	0.77	0.26	0.07	-0.40	0.53

Table 4-2: Statistical results. In one-sample T-Tests, the null hypothesis was that the mean was equal to zero. T, T-statistic. DF, degrees of freedom. P, P-value. BF, Bayes factor in favor of the alternative hypothesis. 95% CI, frequentist confidence interval. Mean differences are computed as placebo-levodopa. Bayes factors were computed using one-tailed default priors for the alternative hypothesis. In all one-sample T-Tests, the alternative hypothesis was that the population mean is greater than zero. For independent T-Tests, the alternative hypothesis stated that adaptation with interference would be greater in the levodopa group than the placebo group. For all other independent T-tests, the alternative hypothesis stated that the measure of interest would be smaller in the levodopa group than the placebo group.

4.3.2.1 Adaptation.

Non-catch trials: Early adaptation was greater than zero in both the placebo (p=3.62e-9, BF=6.36e+6) and levodopa conditions (p=6.61e-8, BF=432848). We also observed reliable late adaptation for both the placebo (p=2.48e-10, BF=7.72e+7) and levodopa (p=4.71e-9, BF=4.99e+6) conditions. We did not observe a reliable difference between drug conditions for either early (p=0.83, BF=0.37) or late (p=0.57, BF=0.22) adaptation.

Catch trials: Early adaptation was greater than zero in both the placebo (p=4.82e-8, BF=574167) and levodopa (p=3.62e-8, BF=755029) conditions. We observed reliable late adaptation in both the placebo (p=2.92e-7, BF=110522) and levodopa (p=2.54e-11, BF=6.48e +8) conditions. There was no reliable difference between drug conditions for either early (p=0.61, BF=0.47), or late (p=0.90, BF=0.29) adaptation.

4.3.2.2 Savings

Non-catch trials: Our analyses yielded inconclusive evidence in favor of the hypothesized effect of savings for early adaptation for both the placebo (p=0.14, BF=1.23) and levodopa (p=0.11, BF=1.43) conditions. There was reliable evidence of savings for early adaptation when both groups were combined (p=0.03, BF=3.63). In the late period of adaptation, Non-catch trials provided inconclusive evidence against the hypothesized effect of savings following placebo (p=0.50, BF=0.44), and moderate evidence against the hypothesized

effect of savings following levodopa (p=0.70, BF=0.18). There was moderate evidence against the hypothesis that savings would be reduced by levodopa in early adaptation (p=0.87, BF=0.28), and inconclusive evidence that savings would be reduced in late adaptation (p=0.45, BF=0.59).

Catch trials: There was moderate evidence against the hypothesized effects of savings for early adaptation following both placebo (p=0.72, BF=0.33) and levodopa (p=0.75, BF=0.19). Evidence for savings in late adaptation was inconclusive following both placebo (p=0.14, BF=1.20) and levodopa (p=0.33, BF=0.60). There was inconclusive evidence against the hypothesis that levodopa would reduce savings for both early (p=0.63, BF=0.47) and late (p=0.39, BF=0.66) adaptation.

4.3.2.3 After-Effects

Non-catch trials: We observed reliable after-effects in the early portion of NFb following adaptation in both the placebo (p=4.00e-8, BF=688519.55) and levodopa (p=2.56e-9, BF=8.79e+6) conditions. We also observed reliable after-effects extending to the later period of NFb after both placebo (p=1.37e-3, BF=56.24) and levodopa (p=9.66e -4, BF=76.15). We observed no reliable evidence that levodopa impaired after-effects in either the early (p=0.99, BF=0.32) or late (p=0.78, BF=0.39) periods.

4.3.2.4 Adaptation with interference

Non-catch trials: Early adaptation following exposure to an opposing force field was reliably greater than zero in both the placebo (p=1.06e-9, BF=2.00e+7) and levodopa (p=3.42e -6, BF=11657.42) conditions. We also observed reliable late adaptation in both the placebo (p=5.98e -12, BF=2.51e +9) and levodopa (p=1.70e -9, BF=1.28e +7) conditions. We observed moderate evidence against the hypothesized effect that levodopa would result in improved adaptation with interference in both the early (p=0.23, BF=0.16) and late (p=0.18, BF=0.15) periods.

Catch trials: Early adaptation following exposure to an opposing force field was reliably greater than zero in both the placebo (p=6.37e-5, BF=837.09) and levodopa (p=4.84e-6,

BF=8524.02) conditions. We also observed reliable late adaptation in both the placebo (p=4.33e-7, BF=77010.30) and levodopa (p=4.20e-8, BF=657919.38) conditions. We observed inconclusive evidence against the hypothesis that levodopa would result in improved adaptation with interference in the early period (p=0.65, BF = 0.45), and moderate evidence in the late period (p = 0.77, BF = 0.26).

4.4 Discussion

Previous work suggests that savings and anterograde interference for motor adaptation may result from model-free reinforcement learning processes mediated by dopaminergic inputs to the striatum (Huang et al., 2011; Leow et al., 2012, 2013; Orban de Xivry & Lefèvre, 2015). Here, we tested the hypothesis that savings and anterograde interference for motor adaptation depend on dopamine release. A sample of healthy young volunteers performed a force field adaptation task designed to produce effects of initial adaptation, savings, and interference. Prior to performing the task, participants either ingested placebo or levodopa, a dopamine precursor. Levodopa is thought to impair reinforcement learning processes that depend on the ventral striatum (Cools et al., 2001, 2007; Feigin et al., 2003; Frank et al., 2004; Graef et al., 2010; Hiebert et al., 2014, 2019; Jahanshahi et al., 2010; Kwak et al., 2010; MacDonald et al., 2011; Swainson et al., 2000; Torta et al., 2009; Vo et al., 2016, 2018). We hypothesized that levodopa would diminish savings and interference while initial adaptation would be unaffected.

Participants reliably adapted to the clockwise force field imposed in blocks FF1a and FF1b, and we found no evidence that adaptation was affected by levodopa. This was expected as force field adaptation is thought to rely primarily on sensory error-based learning mechanisms involving the cerebellum (Block & Bastian, 2012; Izawa et al., 2012; Maschke et al., 2004; M. A. Smith & Shadmehr, 2005; Tseng et al., 2007). Previous studies have found that people with Parkinson's disease (PD) show deficient savings and interference despite relatively normal adaptation (Bédard & Sanes, 2011; Leow et al., 2012, 2013; Marinelli et al., 2009). Because many deficits in reinforcement-based learning in PD can be attributed to side effects of dopaminergic medication, we predicted that

levodopa would impair savings and interference in healthy volunteers. However, levodopa did not impact our measures of either savings or adaptation with interference. Impaired savings may therefore be a specific effect of Parkinson's disease as opposed to a side-effect of levodopa. This is consistent with the findings of Marinelli et al. (2009), who observed a lack of savings effects in drug-naive and off-medication PD patients. Early-stage PD is thought to result in dopamine depletion primarily in the dorsal striatum, while levodopa likely impairs dopaminergic learning mechanisms mediated by the ventral striatum (Hiebert et al., 2019). Thus, savings and interference may occur due to plasticity in the dorsal striatum. There is also evidence that savings occurs through plasticity in primary motor cortex (Orban de Xivry et al., 2011, 2012; Orban de Xivry & Lefèvre, 2015; Richardson et al., 2006). PD is associated with structural and functional alterations of primary motor cortex, including abnormalities in plasticity induced by transcranial magnetic stimulation (Bagnato et al., 2006; Burciu & Vaillancourt, 2018; Ueki et al., 2006). Importantly, levodopa ameliorates altered motor cortical plasticity in PD, suggesting that these deficits are not caused by the medication. Future work might test whether deficits in savings and interference due to PD are linked to altered plasticity in primary motor cortex, and whether they are improved by dopaminergic medication.

An important limitation is that our experimental protocol may have been insufficient to produce reliable savings or interference effects even in the control group, as we observed weak evidence of savings overall. Savings and interference have been shown to depend on sufficient repetition of the adapted movements to produce reinforcement of the adapted movements (Huang et al., 2011; Orban de Xivry & Lefèvre, 2015). Because the current study involved a limited number of reaches to 8 different targets, adaptation may not have resulted in sufficient repetition required to elicit model-free reinforcement of the adapted motor outputs.

4.5 References

Bagnato, S., Agostino, R., Modugno, N., Quartarone, A., & Berardelli, A. (2006). Plasticity of the motor cortex in Parkinson's disease patients on and off therapy. *Movement Disorders*, *21*(5), 639–645. https://doi.org/10.1002/mds.20778

Bédard, P., & Sanes, J. N. (2011). Basal ganglia-dependent processes in recalling learned visual-motor adaptations. *Experimental Brain Research*, *209*(3), 385–393. https://doi.org/10.1007/s00221-011-2561-y

Block, H. J., & Bastian, A. J. (2012). Cerebellar involvement in motor but not sensory adaptation. *Neuropsychologia*, *50*(8), 1766–1775. https://doi.org/10.1016/j.neuropsychologia.2012.03.034

Bond, A., & Lader, M. (1974). The use of analogue scales in rating subjective feelings. *British Journal of Medical Psychology*, *47*(3), 211–218. https://doi.org/10.1111/j.2044-8341.1974.tb02285.x

Burciu, R. G., & Vaillancourt, D. E. (2018). Imaging of Motor Cortex Physiology in Parkinson's Disease. *Movement Disorders*, *33*(11), 1688–1699. https://doi.org/10.1002/mds.102

Coltman, S. K., Cashaback, J. G. A., & Gribble, P. L. (2019). Both fast and slow learning processes contribute to savings following sensorimotor adaptation. *Journal of Neurophysiology*, *121*(4), 1575–1583. https://doi.org/10.1152/jn.00794.2018

Cools, R., Barker, R. A., Sahakian, B. J., & Robbins, T. W. (2001). Enhanced or impaired cognitive function in Parkinson's disease as a function of dopaminergic medication and task demands. *Cerebral Cortex (New York, N.Y.: 1991)*, *11*(12), 1136–1143. https://doi.org/10.1093/cercor/11.12.1136

Cools, R., Lewis, S. J. G., Clark, L., Barker, R. A., & Robbins, T. W. (2007). L-DOPA disrupts activity in the nucleus accumbens during reversal learning in Parkinson's disease. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, *32*(1), 180–189. https://doi.org/10.1038/sj.npp.1301153

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, *8*(12), 1704–1711. https://doi.org/10.1038/nn1560

Diedrichsen, J., Hashambhoy, Y., Rane, T., & Shadmehr, R. (2005). Neural Correlates of Reach Errors. *Journal of Neuroscience*, *25*(43), 9919–9931. https://doi.org/10.1523/JNEUROSCI.1874-05.2005

Dolan, R. J., & Dayan, P. (2013). Goals and Habits in the Brain. *Neuron*, *80*(2), 312–325. https://doi.org/10.1016/j.neuron.2013.09.007

Doll, B. B., Bath, K. G., Daw, N. D., & Frank, M. J. (2016). Variability in Dopamine Genes Dissociates Model-Based and Model-Free Reinforcement Learning. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *36*(4), 1211–1222. https://doi.org/10.1523/JNEUROSCI.1901-15.2016

Donchin, O., Francis, J. T., & Shadmehr, R. (2003). Quantifying Generalization from Trial-by-Trial Behavior of Adaptive Systems that Learn with Basis Functions: Theory and Experiments in Human Motor Control. *Journal of Neuroscience*, *23*(27), 9032–9045. https://doi.org/10.1523/JNEUROSCI.23-27-09032.2003

Feigin, A., Ghilardi, M. F., Carbon, M., Edwards, C., Fukuda, M., Dhawan, V., Margouleff, C., Ghez, C., & Eidelberg, D. (2003). Effects of levodopa on motor sequence learning in Parkinson's disease. *Neurology*, *60*(11), 1744–1749. https://doi.org/10.1212/01.wnl.0000072263.03608.42

Frank, M. J., Seeberger, L. C., & O'reilly, R. C. (2004). By carrot or by stick: Cognitive reinforcement learning in parkinsonism. *Science (New York, N.Y.)*, *306*(5703), 1940–1943. https://doi.org/10.1126/science.1102941

Gläscher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). States versus rewards: Dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*, *66*(4), 585–595. https://doi.org/10.1016/j.neuron.2010.04.016

Glimcher, P. W. (2011). Understanding dopamine and reinforcement learning: The dopamine reward prediction error hypothesis. *Proceedings of the National Academy of Sciences*, *108*(Supplement 3), 15647–15654. https://doi.org/10.1073/pnas.1014269108

Graef, S., Biele, G., Krugel, L. K., Marzinzik, F., Wahl, M., Wotka, J., Klostermann, F., & Heekeren, H. R. (2010). Differential influence of levodopa on reward-based learning in Parkinson's disease. *Frontiers in Human Neuroscience*, *4*, 169. https://doi.org/10.3389/fnhum.2010.00169

Herzfeld, D. J., Vaswani, P. A., Marko, M. K., & Shadmehr, R. (2014). A memory of errors in sensorimotor learning. *Science*, *345*(6202), 1349–1353. https://doi.org/10.1126/science.1253138

Hiebert, N. M., Owen, A. M., Ganjavi, H., Mendonça, D., Jenkins, M. E., Seergobin, K. N., & MacDonald, P. A. (2019). Dorsal striatum does not mediate feedback-based, stimulus-response learning: An event-related fMRI study in patients with Parkinson's disease tested on and off dopaminergic therapy. *NeuroImage*, *185*, 455–470. https://doi.org/10.1016/j.neuroimage.2018.10.045

Hiebert, N. M., Seergobin, K. N., Vo, A., Ganjavi, H., & MacDonald, P. A. (2014). Dopaminergic therapy affects learning and impulsivity in Parkinson's disease. *Annals of Clinical and Translational Neurology*, *1*(10), 833–843. https://doi.org/10.1002/acn3.128

Holland, P., Codol, O., & Galea, J. M. (2018). Contribution of explicit processes to reinforcement-based motor learning. *Journal of Neurophysiology*, *119*(6), 2241–2255. https://doi.org/10.1152/jn.00901.2017

Holland, P., Codol, O., Oxley, E., Taylor, M., Hamshere, E., Joseph, S., Huffer, L., & Galea, J. M. (2019). Domain-Specific Working Memory, But Not Dopamine-Related Genetic Variability, Shapes Reward-Based Motor Learning. *The Journal of Neuroscience*, *39*(47), 9383–9396. https://doi.org/10.1523/JNEUROSCI.0583-19.2019

Huang, V. S., Haith, A., Mazzoni, P., & Krakauer, J. W. (2011). Rethinking Motor Learning and Savings in Adaptation Paradigms: Model-Free Memory for Successful Actions Combines with Internal Models. *Neuron*, *70*(4), 787–801. https://doi.org/10.1016/j.neuron.2011.04.012

Izawa, J., Criscimagna-Hemminger, S. E., & Shadmehr, R. (2012). Cerebellar Contributions to Reach Adaptation and Learning Sensory Consequences of Action. *Journal of Neuroscience*, *32*(12), 4230–4239. https://doi.org/10.1523/JNEUROSCI.6353-11.2012

Jahanshahi, M., Wilkinson, L., Gahir, H., Dharmaindra, A., Dharmarinda, A., Dharminda, A., & Lagnado, D. A. (2010). Medication impairs probabilistic classification learning in Parkinson's disease. *Neuropsychologia*, *48*(4), 1096–1103. https://doi.org/10.1016/j.neuropsychologia.2009.12.010

Kawato, M. (1999). Internal models for motor control and trajectory planning. *Current Opinion in Neurobiology*, *9*(6), 718–727. https://doi.org/10.1016/S0959-4388(99)00028-8

Keysers, C., Gazzola, V., & Wagenmakers, E.-J. (2020). Using Bayes factor hypothesis testing in neuroscience to establish evidence of absence. *Nature Neuroscience*, *23*(7), 788–799. https://doi.org/10.1038/s41593-020-0660-4

Krakauer, J. W. (2009). Motor Learning and Consolidation: The Case of Visuomotor Rotation. In D. Sternad (Ed.), *Progress in Motor Control: A Multidisciplinary Perspective* (pp. 405–421). Springer US. https://doi.org/10.1007/978-0-387-77064-2_21

Kwak, Y., Müller, M. L. T. M., Bohnen, N. I., Dayalu, P., & Seidler, R. D. (2010). Effect of dopaminergic medications on the time course of explicit motor sequence learning in Parkinson's disease. *Journal of Neurophysiology*, *103*(2), 942–949. https://doi.org/10.1152/jn.00197.2009

Leow, L.-A., de Rugy, A., Loftus, A. M., & Hammond, G. (2013). Different mechanisms contributing to savings and anterograde interference are impaired in Parkinson's disease. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00055

Leow, L.-A., Loftus, A. M., & Hammond, G. R. (2012). Impaired savings despite intact initial learning of motor adaptation in Parkinson's disease. *Experimental Brain Research*, *218*(2), 295–304. https://doi.org/10.1007/s00221-012-3060-5

Li, C.-S. R., Padoa-Schioppa, C., & Bizzi, E. (2001). Neuronal Correlates of Motor Performance and Motor Learning in the Primary Motor Cortex of Monkeys Adapting to an External Force Field. *Neuron*, *30*(2), 593–607. https://doi.org/10.1016/S0896-6273(01)00301-4

MacDonald, P. A., MacDonald, A. A., Seergobin, K. N., Tamjeedi, R., Ganjavi, H., Provost, J.-S., & Monchi, O. (2011). The effect of dopamine therapy on ventral and dorsal striatum-mediated cognition in Parkinson's disease: Support from functional MRI. *Brain: A Journal of Neurology*, *134*(Pt 5), 1447–1463. https://doi.org/10.1093/brain/awr075

Marinelli, L., Crupi, D., Di Rocco, A., Bove, M., Eidelberg, D., Abbruzzese, G., & Felice Ghilardi, M. (2009). Learning and consolidation of visuo-motor adaptation in Parkinson's disease. *Parkinsonism & Related Disorders*, *15*(1), 6–11. https://doi.org/10.1016/j.parkreldis.2008.02.012

Maschke, M., Gomez, C. M., Ebner, T. J., & Konczak, J. (2004). Hereditary cerebellar ataxia progressively impairs force adaptation during goal-directed arm movements. *Journal of Neurophysiology*, *91*(1), 230–238. https://doi.org/10.1152/jn.00557.2003

McNamee, D., & Wolpert, D. M. (2019). Internal Models in Biological Control. *Annual Review of Control, Robotics, and Autonomous Systems*, *2*(1), 339–364. https://doi.org/10.1146/annurev-control-060117-105206

Miall, R. C., Jenkinson, N., & Kulkarni, K. (2004). Adaptation to rotated visual feedback: A re-examination of motor interference. *Experimental Brain Research*, *154*(2), 201–210. https://doi.org/10.1007/s00221-003-1630-2

Orban de Xivry, J.-J., Ahmadi-Pajouh, M. A., Harran, M. D., Salimpour, Y., & Shadmehr, R. (2012). Changes in corticospinal excitability during reach adaptation in force fields. *Journal of Neurophysiology*, *109*(1), 124–136. https://doi.org/10.1152/jn.00785.2012

Orban de Xivry, J.-J., Criscimagna-Hemminger, S. E., & Shadmehr, R. (2011). Contributions of the Motor Cortex to Adaptive Control of Reaching Depend on the Perturbation Schedule. *Cerebral Cortex*, *21*(7), 1475–1484. https://doi.org/10.1093/cercor/bhq192

Orban de Xivry, J.-J., & Lefèvre, P. (2015). Formation of model-free motor memories during motor adaptation depends on perturbation schedule. *Journal of Neurophysiology*, *113*(7), 2733–2741. https://doi.org/10.1152/jn.00673.2014

Padoa-Schioppa, C., Li, C.-S. R., & Bizzi, E. (2004). Neuronal Activity in the Supplementary Motor Area of Monkeys Adapting to a New Dynamic Environment. *Journal of Neurophysiology*, *91*(1), 449–473. https://doi.org/10.1152/jn.00876.2002

Perich, M. G., Gallego, J. A., & Miller, L. E. (2018). A Neural Population Mechanism for Rapid Learning. *Neuron*, *100*(4), 964-976.e7. https://doi.org/10.1016/j.neuron.2018.09.030

Richardson, A. G., Overduin, S. A., Valero-Cabré, A., Padoa-Schioppa, C., Pascual-Leone, A., Bizzi, E., & Press, D. Z. (2006). Disruption of Primary Motor Cortex before Learning Impairs Memory of Movement Dynamics. *Journal of Neuroscience*, *26*(48), 12466–12470. https://doi.org/10.1523/JNEUROSCI.1139-06.2006

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, *56*(5), 356–374. https://doi.org/10.1016/j.jmp.2012.08.001

Shadmehr, R., & Brashers-Krug, T. (1997). Functional stages in the formation of human long-term motor memory. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *17*(1), 409–419.

Shadmehr, R., & Mussa-Ivaldi, F. A. (1994). Adaptive representation of dynamics during learning of a motor task. *Journal of Neuroscience*, *14*(5), 3208–3224. https://doi.org/10.1523/JNEUROSCI.14-05-03208.1994

Shadmehr, R., Smith, M. A., & Krakauer, J. W. (2010). Error correction, sensory prediction, and adaptation in motor control. *Annual Review of Neuroscience*, *33*, 89–108. https://doi.org/10.1146/annurev-neuro-060909-153135

Sidarta, A., van Vugt, F. T., & Ostry, D. J. (2018). Somatosensory working memory in human reinforcement-based motor learning. *Journal of Neurophysiology*, *120*(6), 3275–3286. https://doi.org/10.1152/jn.00442.2018

Smith, M. A., Ghazizadeh, A., & Shadmehr, R. (2006). Interacting Adaptive Processes with Different Timescales Underlie Short-Term Motor Learning. *PLOS Biology*, *4*(6), e179. https://doi.org/10.1371/journal.pbio.0040179

Smith, M. A., & Shadmehr, R. (2005). Intact Ability to Learn Internal Models of Arm Dynamics in Huntington's Disease But Not Cerebellar Degeneration. *Journal of Neurophysiology*, *93*(5), 2809–2821. https://doi.org/10.1152/jn.00943.2004

Steinberg, E. E., Keiflin, R., Boivin, J. R., Witten, I. B., Deisseroth, K., & Janak, P. H. (2013). A causal link between prediction errors, dopamine neurons and learning. *Nature Neuroscience*, *16*(7), 966–973. https://doi.org/10.1038/nn.3413

Swainson, R., Rogers, R. D., Sahakian, B. J., Summers, B. A., Polkey, C. E., & Robbins, T. W. (2000). Probabilistic learning and reversal deficits in patients with Parkinson's disease or frontal or temporal lobe lesions: Possible adverse effects of dopaminergic medication. *Neuropsychologia*, *38*(5), 596–612. https://doi.org/10.1016/s0028-3932(99)00103-7

Thoroughman, K. A., & Shadmehr, R. (2000). Learning of action through adaptive combination of motor primitives. *Nature*, *407*(6805), 742–747. https://doi.org/10.1038/35037588

Torta, D. M. E., Castelli, L., Zibetti, M., Lopiano, L., & Geminiani, G. (2009). On the role of dopamine replacement therapy in decision-making, working memory, and reward in Parkinson's disease: Does the therapy-dose matter? *Brain and Cognition*, *71*(2), 84–91. https://doi.org/10.1016/j.bandc.2009.04.003

Tseng, Y., Diedrichsen, J., Krakauer, J. W., Shadmehr, R., & Bastian, A. J. (2007). Sensory Prediction Errors Drive Cerebellum-Dependent Adaptation of Reaching. *Journal of Neurophysiology*, *98*(1), 54–62. https://doi.org/10.1152/jn.00266.2007

Ueki, Y., Mima, T., Kotb, M. A., Sawada, H., Saiki, H., Ikeda, A., Begum, T., Reza, F., Nagamine, T., & Fukuyama, H. (2006). Altered plasticity of the human motor cortex in Parkinson's disease. *Annals of Neurology*, *59*(1), 60–71. https://doi.org/10.1002/ana.20692

Vo, A., Seergobin, K. N., & MacDonald, P. A. (2018). Independent effects of age and levodopa on reversal learning in healthy volunteers. *Neurobiology of Aging*, *69*, 129–139. https://doi.org/10.1016/j.neurobiolaging.2018.05.014

Vo, A., Seergobin, K. N., Morrow, S. A., & MacDonald, P. A. (2016). Levodopa impairs probabilistic reversal learning in healthy young adults. *Psychopharmacology*, *233*(14), 2753–2763. https://doi.org/10.1007/s00213-016-4322-x

Wetzels, R., Matzke, D., Lee, M. D., Rouder, J. N., Iverson, G. J., & Wagenmakers, E.-J. (2011). Statistical Evidence in Experimental Psychology: An Empirical Comparison Using 855 t Tests. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *6*(3), 291–298. https://doi.org/10.1177/1745691611406923

Xiao, J., Padoa-Schioppa, C., & Bizzi, E. (2006). Neuronal correlates of movement dynamics in the dorsal and ventral premotor area in the monkey. *Experimental Brain Research*, *168*(1), 106–119. https://doi.org/10.1007/s00221-005-0074-2

Zarahn, E., Weston, G. D., Liang, J., Mazzoni, P., & Krakauer, J. W. (2008). Explaining Savings for Visuomotor Adaptation: Linear Time-Invariant State-Space Models Are Not Sufficient. *Journal of Neurophysiology*, *100*(5), 2537–2548. https://doi.org/10.1152/jn.90529.2008

Chapter 5

EEG correlates of physical effort and reward processing during reinforcement learning

5.1 Introduction

Humans and other animals tend to make decisions that lead to more rewarding and less physically effortful outcomes (Hartmann et al., 2013; Kennerley et al., 2009; Morel et al., 2017; Rangel & Hare, 2010; Walton et al., 2006). Movement parameters may be planned and adapted not only to increase reward, but also to reduce effort as a cost (Cos et al., 2011, 2014; Selinger et al., 2015, 2019). For example, an increase of movement velocity may allow reward to be attained at a higher rate, but may also require more physical effort. A recent theory suggests that movement parameters such as speed are determined to maximize subjective utility, a function of reward value discounted by effort cost (Shadmehr et al., 2016a, 2016b; Summerside et al., 2018; Yoon et al., 2020). Chapters 2 and 3 describe findings suggesting that an EEG signal called the feedback-related negativity/reward positivity (FRN/RP) reflects neural reward processing during reward-based motor adaptation. Here, we test the hypothesis that the FRN/RP reflects subjective utility for which physical effort functions as a cost.

During decision making, dopaminergic, prefrontal, and striatal structures are implicated in motivating effortful behavior to obtain reward and in integrating reward and effort costs to make value-based choices (Denk et al., 2005; Hosking et al., 2015; Kurniawan et al., 2010, 2011; Rudebeck et al., 2008; Salamone et al., 2003, 2007; Schweimer et al., 2005; Walton et al., 2003). After decisions are made and actions are produced, midbrain dopaminergic neurons signal the difference between expected and obtained reward to the ventral striatum and prefrontal regions (Bayer & Glimcher, 2005; Gläscher et al., 2010; Graybiel, 2008; Holroyd & Coles, 2002; Pessiglione et al., 2006; Puig & Miller, 2012; Schultz, 2006). This reward prediction error signal is thought to drive reinforcement learning by updating reward expectations, allowing for adaptive behavior in uncertain or changing environments. Although many reward-processing areas have been shown to also process effort costs during decision making before action selection, we know relatively little about whether motor costs associated with effort also modulate reward processing after an action is completed, in response to feedback about success or failure.

During decision making, the anterior cingulate cortex (ACC) is known to encode prospective reward and effort cost and to integrate both into a unitary subjective utility signal characterized by effort-discounted reward (Croxson et al., 2009; Kennerley et al., 2011; Klein-Flügge et al., 2016; Porter et al., 2019; Prévost et al., 2010; Rudebeck et al., 2006). During outcome evaluation the ACC encodes reward prediction error and supports reinforcement learning (Amiez et al., 2005; S. Ito et al., 2003; Kennerley et al., 2011; H. Seo & Lee, 2007; Walsh & Anderson, 2012; Williams et al., 2004), but it remains to be shown whether reward learning signals in the ACC also integrate motor effort costs. Contrary to this idea, fMRI studies have argued that separate neural systems underlie reward and effort learning, with ACC activity reflecting prediction errors for effort but not reward (Hauser et al., 2017; Skvortsova et al., 2014). However, an event-related potential (ERP) measured by EEG called the feedback-related negativity, or alternatively the reward positivity (FRN/RP), is a reliable neural correlate of reward prediction error and is consistently localized to the ACC (Becker et al., 2014; Cohen & Ranganath, 2007; Emeric et al., 2008; Gehring & Willoughby, 2002; Hauser, Iannaccone, Stämpfli, et al., 2014; Holroyd & Coles, 2002; Mathewson et al., 2008; Miltner et al., 1997; Vezoli & Procyk, 2009; Walsh & Anderson, 2012; Warren et al., 2015). We sought to test whether the FRN/RP not only acts as a learning signal for reward outcomes but also physical effort requirements.

Economic theories assert that effort is a cost that devalues reward and thus predict diminished neural responses to reinforcement for more costly rewards (Botvinick et al., 2009; Hartmann et al., 2013; Hauser et al., 2017; Shadmehr et al., 2016a). If the FRN/RP encodes reward value discounted by effort cost, then it could reflect a neural mechanism for adapting behavior not only to increase reward but to increase the overall economy of actions. Because other ERP components have also been implicated in outcome processing, we tested for effects in a broad temporal range (Glazer et al., 2018). In particular, the P300 is affected by various properties of motivationally relevant feedback including valence, magnitude, likelihood, and other high-level variables (Ma et al., 2014;

San Martín, 2012; Sato et al., 2005; J. Wang et al., 2014; Y. Wu & Zhou, 2009; Yeung & Sanfey, 2004).

Paradoxically, it has been found in humans and animals that effort can enhance the reinforcing quality of rewards (Clement et al., 2000; Inzlicht et al., 2018; Lydall et al., 2010; Zentall, 2010). It may be that prospective effort devalues reward, while retrospective effort amplifies reinforcement. For example, when given a choice between responses requiring high and low effort, participants choose to produce less effort in the immediate future. However, when given a choice between conditioned reinforcers that follow either low or high effort, humans and other animals tend to prefer the reinforcer that followed greater effort in the past (Alessandri et al., 2008; Clement et al., 2000; Hernandez Lallement et al., 2014; Zentall, 2010). Like many real-world situations, uncertain reward was obtained only after effort expenditure in the present study.

Previous EEG experiments have investigated interactions between effort and reward processing (Gheza et al., 2018). In tasks that provide reinforcement feedback after performance of cognitive tasks with varying attentional or mental demands, increased effort has been shown to enhance the FRN/RP (Ma et al., 2014; Schevernels et al., 2014; L. Wang et al., 2017). This is consistent with the notion that preceding effort enhances reward signals. However, increased effort in cognitive tasks is almost invariably associated with higher difficulty and thus lower probability of success. In this case, it is difficult to determine whether enhanced reward signals are due to increased effort or lower reward expectations, which would result in larger reward prediction error.

In the present study, physical effort is manipulated by changing the magnitude of muscle contractions required to complete the task, but the probabilities of reward and success are equated across the different effort conditions. This allows us to assess the effects of effort in terms of motor cost without confounds related to reward expectation. Furthermore, unlike previous EEG experiments that varied cognitive effort randomly, physical effort requirements in the present study were affected by participants' choice behavior. It has been shown that the FRN/RP and other neural correlates of outcome processing are

typically more sensitive when outcomes are attributable to agents' actions (Hassall et al., 2019; Martin & Potts, 2011; Sambrook & Goslin, 2015; Walsh & Anderson, 2012; Yeung et al., 2005; Zink et al., 2004).

Participants first made binary choices, and then they received feedback about the resulting effort requirements, which were probabilistic and uncertain. Subsequently, they performed an effortful electromyographic (EMG) production task for which they received variable reward that was dependent on precisely producing a target level of EMG activity. This trial sequence allowed us to test the hypothesis that effort information is maintained during the course of an action and that this information is integrated retrospectively with reward feedback. According to this hypothesis, feedback indicating effort requirements in the present study would not elicit neural reinforcement signals such as the FRN/RP, whereas the neural response to reward feedback at the end of each trial would be modulated by both reinforcement outcome and the preceding effort. Alternatively, if effort is treated simply as an aversive stimulus or an economic loss by a standard temporal difference learning process, then feedback that predicts the upcoming effort but not the reward outcome should reflect effort as a cost (Mulligan & Hajcak, 2018).

5.2 Materials and methods

5.2.1 Participants

Data from a total of $n = 18$ healthy participants were analyzed and reported (mean age: 22.12 yr, SD: 3.66; 9 men, 9 women). Four additional participants underwent the experimental procedure but were excluded because of excessive EEG artifacts caused by sweat or movement associated with the task. Participants provided written informed consent to experimental procedures approved by the Research Ethics Board at The University of Western Ontario.

5.2.2 Experimental setup

To allow for isometric contractions of the quadriceps muscles, participants were restrained to a chair by straps on their shoulders and waists. Participants' ankles were

strapped to a rack fixed at the base of the chair, with the knees bent at ~90°. Participants were seated in front of a CRT monitor with their hands resting on a table positioned to make button presses on a response box.

5.2.3 EMG and EEG recording

Unreferenced EEG activity was recorded at 512 Hz with a 64-channel Biosemi ActiveTwo system. Electrodes were mounted in an elastic cap and distributed according to the extended 10-20 system with electrode Cz placed over the vertex. Instead of the typical ground electrode, Biosemi forms a feedback loop between an active Common Mode Sense electrode and a passive Driven Right Leg electrode. The Common Mode Sense electrode was located in the center of the area between P1, Pz, PO3, and POz. The Driven Right Leg electrode was located in the center of the area between Pz, P2, PO3, and PO4. Electrooculogram (EOG) was recorded with electrodes placed above and below each eye and the outer canthus of each eye. Additional electrodes were placed on each mastoid.

EMG activity was recorded at 2,400 Hz bilaterally from the vastus lateralis muscles of the quadriceps with an active electrode system and amplifier (g.USBamp; g.tec Medical Engineering). Two electrodes were placed on each muscle belly for bipolar recordings, and a ground electrode was placed on the left shin. EMG signals were filtered at the time of recording with a 5- to 500-Hz band-pass filter and a 60-Hz notch filter.

5.2.4 Visual feedback of EMG

The EMG signal used to provide online visual feedback of quadriceps muscle activity was first rectified, low-pass filtered with a 10-Hz cutoff frequency, and then downsampled to 120 Hz. At the beginning of each block, participants performed isometric knee extensions with maximum effort continuously for 4 s. All samples greater than the median value recorded during maximum effort were averaged to determine the value of maximum voluntary contraction (MVC) used throughout the block. Subsequently, participants were cued to remain completely still and keep their legs relaxed for 4 s. The mean EMG signal during this period was used as a baseline value throughout the block.

During each trial, an animation of a thermometer was displayed to participants. The fluid level of the thermometer increased in real time (monitor refresh rate: 60 Hz) as a linear function of the processed EMG signal. In "hard" effort trials, the top of the thermometer corresponded to 85% of MVC and the bottom of the thermometer corresponded to the baseline measure. In "easy" effort trials, the top of the thermometer corresponded to 15% of MVC. Because easy trials required only a small amount of muscle activity to reach the target, the gain of the visual feedback relative to the EMG was high. To reduce the gain and provide smooth feedback, the baseline measure (resting EMG) was made to correspond to the point halfway up the thermometer for the easy condition, so that the temperature moved half the distance on the display. The fluid level was calculated separately for each leg based on their respective MVC and baseline measures, and the average was used to display feedback. A running average of fluid level for the previous 60 samples was drawn to the screen to provide smooth feedback. During each trial, the maximum fluid level for that trial was continuously displayed such that the fluid level only increased, and if the participant relaxed their quadriceps muscles the feedback would remain at the same level. This allowed for smooth, ballistic isometric contractions. It also made it so that participants were not required to hold the fluid level constant without visual feedback, which often resulted in the fluid level fluctuating or drifting away from the target during pilot experiments.

5.2.5 Experimental task

Participants first performed a block of 28 practice trials (see below). Participants then performed four blocks of 74 trials with self-paced rest periods between blocks. Each block consisted of 12 control condition trials, followed by 50 experimental condition trials and finally 12 additional control condition trials.

Experimental condition: During each trial, participants made a binary choice that probabilistically determined whether the trial would require easy or hard physical effort. The effort contingencies had to be learned through experience. Participants then performed isometric knee extensions to control visual EMG feedback on a screen.

Participants were instructed to exceed a minimum level of muscle activation indicated by a visual target while remaining as close as possible to the target. Binary reinforcement feedback was provided at the end of each trial to indicate success or failure, which corresponded to a small monetary reward.

Visual stimuli are shown in figure 5-1. An animated thermometer was drawn on the screen throughout the task. A cross was drawn at the top of the thermometer to serve as a target for EMG feedback. Letters "A" and "B" drawn to the left and right of the thermometer represented the options for binary choices made in each trial. Participants initiated each trial by pressing either a left or right button on a response box with their left or right index finger, respectively. Immediately on each button press, the choice was indicated by a box appearing around the letter "A" or "B" for the left and right response buttons, respectively. The box remained throughout the trial.

Effort feedback: One second after the button press participants received feedback indicating the effort that they would be required to exert on the present trial. The word "easy" or "hard" replaced the target cross for 700 ms to indicate upcoming required effort. The effort condition was determined probabilistically by the participants' response, and the effort contingencies had to be learned through experience. One of the responses led to a hard effort trial with a probability of 0.8 and an easy effort trial with a probability of 0.2. The other response led to a hard effort trial with a probability of 0.2 and an easy effort trial with a probability of 0.8. Unannounced to participants, the effort contingencies periodically reversed. Reversals occurred after the response more likely to produce easy effort was chosen a cumulative number of times, which was randomly selected to be between 5 and 9 for each reversal. Participants were instructed that their responses would affect the effort requirements in some way but were not informed of the specific nature of the task. Participants were not instructed to respond in any particular way other than to sample both choices.

After the effort feedback was removed from the display, the target cross reappeared for 800 ms. Subsequently, the effort production phase of the trial began. During this phase,

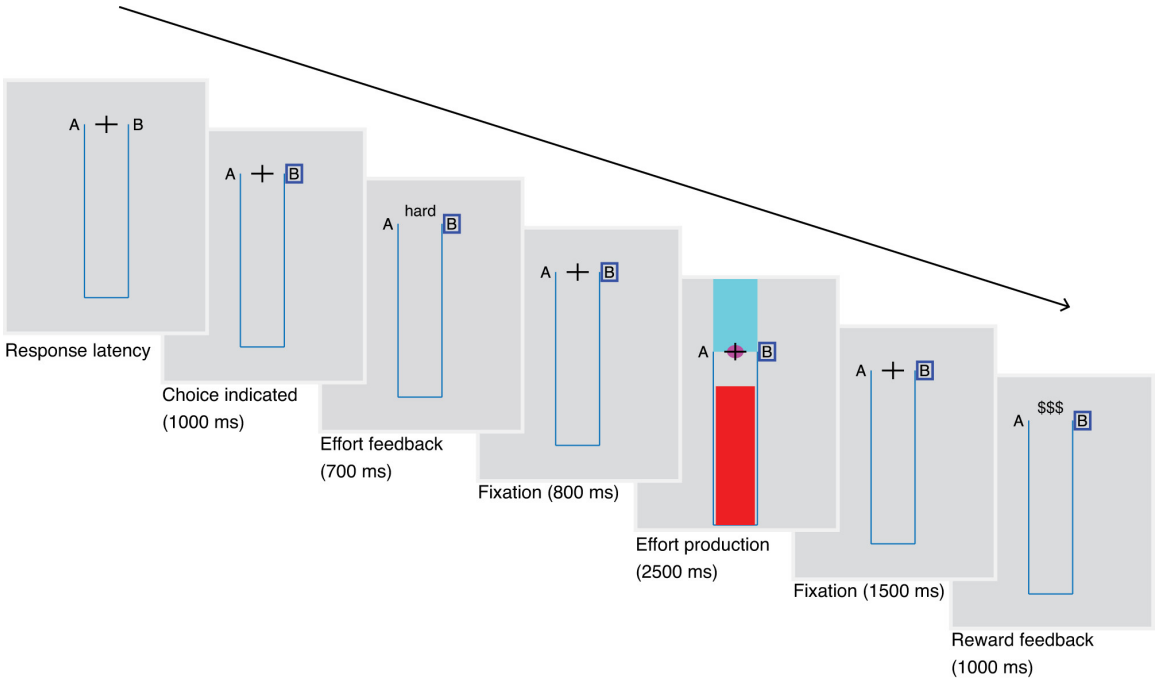

Figure 5-1: Stimuli. Participants initiated each trial by indicating a binary decision through button press. A box immediately appeared around either the letter "A" or "B," corresponding to the choice options. After 1,000 ms, feedback appeared to inform participants that their choice resulted in either easy or hard physical effort requirements for the upcoming electromyographic (EMG) production task. A purple circle appeared over the target cross to cue the onset of the EMG production phase, during which participants performed isometric knee extension and the fluid level of a thermometer indicated quadriceps muscle activation. The circle shrank continuously, disappearing in 2.5 s to cue the end of the EMG production phase. Participants attempted to bring the fluid level above the target, represented by a cross, while remaining as close as possible to the target. However, a mask drawn above the target prevented participants from seeing the extent of errors that they made in overshooting the target. Instead, binary reinforcement feedback was provided 1.5 s after the EMG production phase ended, indicating whether or not participants had successfully exceeded the target while remaining sufficiently close.

the fluid level of the thermometer was drawn continuously to provide EMG feedback (see *Visual Feedback of EMG*). The fluid level increased with increasing EMG signal but represented the maximum signal for the trial, and thus never decreased. A purple circle was drawn under the target to cue the beginning of the effort production phase, and participants were instructed to keep their legs relaxed until they saw this cue. The circle shrank continuously during the course of the trial, disappearing in 2,500 ms to signal the end of the effort production phase, at which point EMG feedback disappeared. Participants were instructed that to complete the task successfully the final fluid level must exceed the target represented by the center of the cross. The target corresponded to 15% and 85% of MVC in the easy effort and hard effort conditions, respectively. Furthermore, participants were instructed to keep the fluid level as close as possible to the target; thus their goal was to always overshoot the target but to minimize the extent of overshoot. Participants were instructed to relax their legs as soon as possible after reaching the target, as the fluid level did not decrease during a trial. EMG feedback was withheld above the target by a mask drawn on the top of the thermometer. This prevented participants from seeing the extent of their overshoot errors.

Reinforcement feedback: Feedback about performance was provided at the end of the trial with binary reinforcement. At the end of the effort production phase, the EMG feedback and the mask disappeared. After 1,500 ms of fixation, the target cross was replaced with either "$$$" or "XXX" to indicate a rewarded or failed trial, with a reward being indicated if the fluid level exceeded the target while remaining sufficiently close to it. Participants were instructed that they could earn up to an additional 10 CAD throughout the task according to the number of trials in which they received feedback indicating success. The error threshold for overshoot was adjusted with a 1-up-1-down adaptive staircase separately for the two effort conditions to ensure a 50% reinforcement rate overall for both conditions.

Control condition: Each block began and ended with 12 control trials, during which the task was the same as the experimental condition except no reinforcement feedback was provided and the effort condition was deterministic and independent of participants'

responses. Both runs of 12 control trials consisted of 6 easy effort trials and 6 hard effort trials, with the trials of each effort condition occurring consecutively. The text "easy effort" or "hard effort" was displayed at the top of the screen continuously to cue the effort condition for all control trials. Participants were instructed to make a button press to initiate each trial but were instructed that the choice was arbitrary and that the effort condition would always correspond to the cue at the top of the screen. In the first 12 control trials of each block, there was no mask drawn on the top of the thermometer, so participants could see their overshoot errors in order to practice the task more effectively. In the final 12 control trials of each block, the mask was drawn for each trial as in the experimental condition. The orders of easy and hard condition runs during the control trials were randomized and balanced across the four blocks for each participant.

Practice trials: Participants first performed a practice block to learn how to control the EMG feedback. As in the control trials, no reinforcement feedback was provided and the effort condition was cued to participants before each trial and independent of participants' responses. The practice block began with seven easy effort trials followed by seven hard effort trials without the mask drawn at the top of the thermometer. Participants then performed seven easy effort trials followed by seven hard effort trials with the mask.

5.2.6 Behavioral analysis

The effect of effort and reinforcement outcomes on behavioral choice was analyzed with logistic regression performed with the *Glmnet* package in R. The dependent variable was whether the participants' choice on trial n corresponded to staying or switching from the choice on trial $n - 1$, coded as 0 or 1. The independent variables were determined by the effort and reinforcement outcomes on trial $n - 1$:

Effort: −1 for easy effort, 1 for hard effort

Reward: −1 for nonreward, 1 for reward

Effort/Reward interaction: Effort × Reward

Switch: 1 for all trials

Logistic regression was calculated separately for each participant. Regularization was applied with an L2-norm penalty. The penalty constant, λ, was selected by leave-out cross-validation. A value of 0.04297 was chosen as it is the largest value that minimizes the cross-validated misclassification error, averaged across subjects. The coefficients for Effort, Reward, and the interaction term were each submitted to one-sample *t* tests against zero.

5.2.7 EEG preprocessing and denoising

EEG data were preprocessed with the EEGLAB toolbox (see Delorme & Makeig, 2004, for details), except for filtering, which was performed with the MATLAB *filtfilt* function. Data, initially referenced to linked mastoids, were band-pass filtered with a second-order Butterworth filter with a passband of 0.1–45 Hz. Channels with poor recording quality or excessive artifacts were identified with visual inspection and interpolated with spherical interpolation. EEG data were then rereferenced to the average scalp potential, and interpolated electrodes were subsequently removed from the data before independent component analysis (ICA). Two epochs were extracted for each trial corresponding to effort condition feedback following the button press response and reinforcement feedback following the effort production phase. Continuous data were segmented into 2.5-s epochs time-locked to stimulus onset at 0 ms (time range: −1,000 to +1,500 ms). Data epochs containing artifacts other than blinks were removed by visual inspection. Subsequently, extended infomax ICA was performed on each participant's data (Delorme & Makeig, 2004). Components reflecting eye movements and blink artifacts were identified by visual inspection and subtracted by projection of the remaining components back to the voltage time series.

5.2.8 Event related potential analysis

Trial averaging: We computed event-related potentials (ERPs) on an individual participant basis by trial-averaging EEG time series epochs recorded from electrode FCz after artifact removal. ERPs were analyzed after time-locking signals to two points in time: effort feedback and reinforcement (performance) feedback. The FRN/RP is typically

maximal at electrode FCz, and this selection is consistent with previous work including our own (Holroyd & Krigolson, 2007; Miltner et al., 1997; Palidis et al., 2019; Pfabigan et al., 2011). We selected trials corresponding to various feedback conditions in each task.

For ERPs time locked to reinforcement feedback, we computed ERPs corresponding to "easy nonreward" (45.4 ± 6.9 trials), "easy reward" (46.3 ± 6.2 trials), "hard nonreward" (37.9 ± 8.4 trials), and "hard reward" (45.6 ± 7.7 trials) conditions. In the control condition, participants performed the effort production task but did not receive any reinforcement feedback. We computed ERPs for the "control easy" (41.5 ± 5.9 trials) and "control hard" (38.5 ± 5.5 trials) conditions time locked to the moment when reinforcement feedback would have been delivered in the experimental condition. For ERPs corresponding to reinforcement feedback and the control condition, we excluded all trials in which the visual EMG feedback did not reach the target, as in this case a nonreward outcome was evident before the reinforcement feedback was delivered.

We also extracted ERPs time locked to the effort condition feedback, which indicated the upcoming effort requirements after each button press but before the participant performed the EMG production task ("easy feedback" 94.9 ± 10.2 trials and "hard feedback" 92.4 ± 13.5 trials). All ERPs were baseline corrected by subtracting the average voltage in the 100-ms period immediately before stimulus onset. Finally, ERPs were low-pass filtered with a cutoff frequency of 30 Hz.

Statistical analysis: We performed statistical tests on each sample between 100 and 600 ms after feedback onset. We selected this time window as it is wide enough to capture effects outside of the FRN/RP yet constrained to a range during which ERPs are likely to be affected by feedback processing (Glazer et al., 2018). We corrected significance values for multiple comparisons across time with the Benjamini–Hochberg procedure for estimating the false discovery rate (FDR), implemented by the MATLAB *mafdr* function.

To analyze the neural response to reinforcement feedback, we performed 2 × 2 repeated-measures ANOVA with the MATLAB *ranova* function. The factors were reward outcome (levels: nonreward, reward) and effort condition (levels: easy, hard). We used one-

sample t-tests against zero on the difference waves computed between easy feedback and hard feedback ERPs, aligned to feedback indicating effort condition after each button press but before the EMG production phase. To test for artifacts related to the isometric leg extension, we used one-sample t-tests against zero on the difference waves computed between easy control and hard control ERPs. These ERPs were aligned to the moment when reinforcement feedback would have been delivered in the experimental condition, but instead the target cross simply disappeared briefly. Participants were told that they would not receive feedback in this condition and thus did not expect a possible reward.

Scalp distributions: Scalp distributions were plotted with the EEGLAB *topoplot* function using the mean amplitude of difference waves within specified time windows, averaged across subjects.

Source separation: Because of volume conduction, potential differences between any particular scalp electrodes contain mixed contributions from nearly all active neural sources and artifacts. Measurements of any event-related potential (ERP) component using scalp regions of interest, such as the measurements of the FRN/RP described above, are thus prone to contamination by other ERP components with distinct neural sources. Independent component analysis (ICA) can be used to produce spatial filters that isolate activity measured from separate cortical sources (Onton & Makeig, 2006). Each component returned by ICA is a linear weighting of all electrodes, computed to produce signals with maximal temporal independence (Delorme & Makeig, 2004).

Because ICA decomposition of neural activity sources can be particularly sensitive to signal properties and noise, we preprocessed the data using a modified procedure to produce ICA weightings for source separation. Except for the differences described below, we followed the same preprocessing used for the ERP analysis, including rejection of the same channels and epochs for artifact removal, as described in *EEG Preprocessing and Denoising*. EEG data were downsampled to 256 Hz and high-pass filtered with a second-order Butterworth filter with a cutoff of 0.25 Hz instead of 0.1 Hz. Aggressive high-pass cutoff frequencies of 1–2 Hz have been shown to improve ICA decomposition

(Winkler et al. 2015). However, high-pass filtering at or above 0.3 Hz has been shown to attenuate and distort long-latency ERP components, and 0.1-Hz cutoff is generally recommended for ERP analysis (Acunzo et al., 2012; Holinger et al., 2000; Tanner et al., 2015). We chose to compute ICA weights with data high-pass filtered with a 0.25-Hz cutoff as a suitable compromise. Sixty-hertz power line noise was removed with the *CleanLine* EEGLAB plugin. Data were selected with the time range −100 to +600 ms centered around effort feedback and reinforcement feedback, instead of −1,000 to +1500, so that the ICA would primarily account for variance in the time window of interest. The same epochs previously identified to contain artifacts, as described in *EEG Preprocessing and Denoising*, were rejected before extended infomax ICA was performed on each participant's data. Subsequently, the time series of the independent components' activities were visually inspected, and additional epochs containing artifacts were flagged and removed before recomputing the ICA weights. According to the tutorial wiki maintained by the developers of EEGLAB, ICA can "concentrate" artifacts for easier rejection, and recomputing ICA after such a rejection "may improve the quality of the ICA decomposition, revealing more independent components accounting for neural, as opposed to mixed artifactual activity" (see https://sccn-ucsd-edu.proxy1.lib.uwo.ca/wiki/Chapter_01:_Rejecting_Artifacts).

To analyze the contribution of individual independent components (ICs) to the ERPs (IC-ERPs), the ICA weights computed for each participants' data were then applied to the data originally preprocessed for ERP analysis as described in *EEG Preprocessing and Denoising*. Thus, identical preprocessed data were used for traditional ERP analysis and the IC-ERP analysis, along with identical procedures described above in *Trial averaging* and *Statistical analysis*. The only difference was that the traditional ERP analysis used mixed data recorded from electrode FCz, whereas the IC-ERP analysis used activity from selected ICs back-projected onto electrode FCz. ICs corresponding to brain activity as opposed to artifacts were identified by stereotyped properties including scalp topographies resembling dipolar projections and spectral peaks at frequencies typical of EEG activity. The traditional ERP analysis revealed multiple ERP components peaking at

different latencies. ICs corresponding to particular ERP components were identified by the presence of maximal peaks in the IC-ERPs in corresponding time windows (see *IC-ERP Results*).

5.3 Results

5.3.1 Behavioral results

Participants made binary decisions that probabilistically determined the effort requirements for each trial. Participants underwent the hard effort condition in 49.6% (SD: 4.8%) of trials. Reward was delivered if EMG feedback exceeded a target level while staying sufficiently close to the target. Participants received reward in 49.4% (SD: 0.01%) of trials. We performed logistic regression for each subject to predict switching of responses between trials $n - 1$ and n, with the effort condition and reward outcome on trial $n - 1$ as the predictors. Figure 5-2B shows the coefficients estimated for each subject, and Figure 5-2A shows the proportion of trials after which participants switched responses for the different reward and effort outcomes. We found that the coefficients for the effect of effort on switching were significantly greater than zero [1-sample t test; $t(17) = 2.263$, $P = 0.037$]. The coefficients for the effect of reward were not reliably different from zero [$t(17) = -0.871$, $P = 0.3959$], nor were the coefficients for the interaction term [$t(17) = 0.252$, $P = 0.8043$].

5.3.2 ERP results

Figure 5-3A shows the ERPs elicited by reinforcement feedback. We analyzed the neural response to reinforcement feedback by performing 2 × 2 repeated-measures ANOVA for each individual time point 100–600 ms after feedback onset. P-values are corrected for multiple comparisons across time points with FDR. In response to reinforcement feedback, that ERP amplitude was larger in response to reward compared with nonreward between 184ms and 336ms after feedback onset (Figure 5-3B; reward main effect, ranges for significant time points: $F = [7.99\ 52.99]$, $P = [0.045\ 0.0001]$, uncorrected $P = [0.012\ <0.0001]$). We also found that voltage was lower in the hard effort condition than in the easy effort condition 238–254 ms after feedback onset (Figure 5-3C; main effect effort,

ranges for significant time points: F = [13.74 14.45], P = [0.050 0.050], uncorrected P = [0.0018 0.0011]). We found effort/reward interaction effects starting 250 ms after feedback onset and up to 600 ms, the end of our time window for statistical testing (Figure 5-3D; ranges for significant time points: F = [5.73 38.94], P = [0.048 0.0008], uncorrected P = [0.028 <0.0001]).

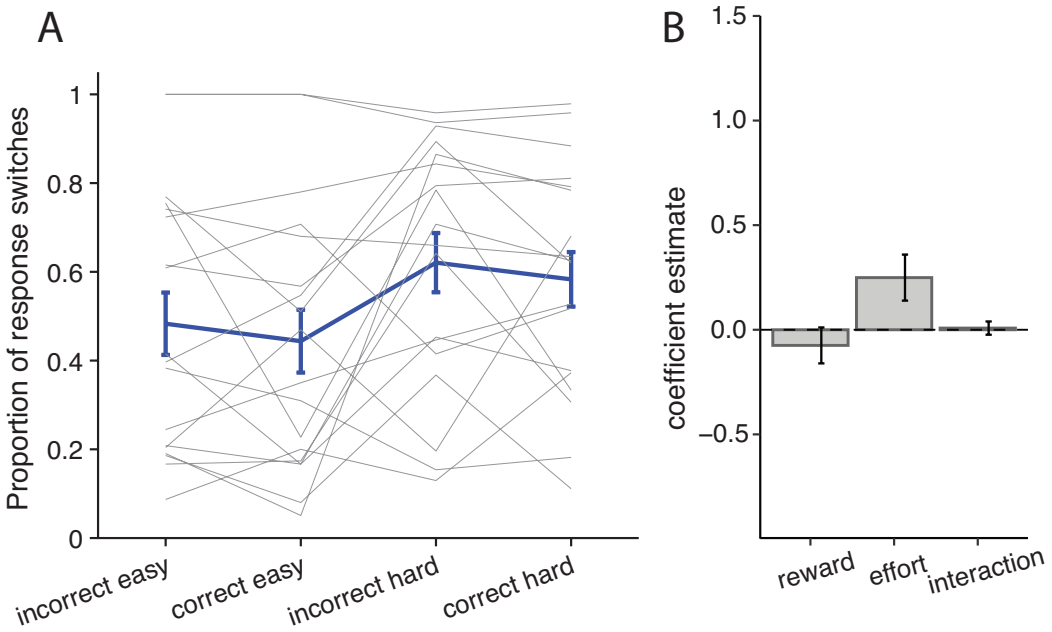

Figure 5-2: Adaptive responses to effort and reward. Participants switch responses more frequently after hard physical effort than easy effort. ***A***: the proportion of trials on which participants switched responses between trial $n - 1$ and trial n for the different reward outcomes and effort conditions on trial $n - 1$ (error bars: ±SE). ***B***: coefficients estimated by using logistic regression to predict response switching for each participant (bars indicate mean ± 1 SE). Predictors were reward outcome, effort condition, and reward/effort interaction. The effort term was significantly greater than 0 (P = 0.037), indicating that participants were more likely to switch responses after hard effort than easy effort.

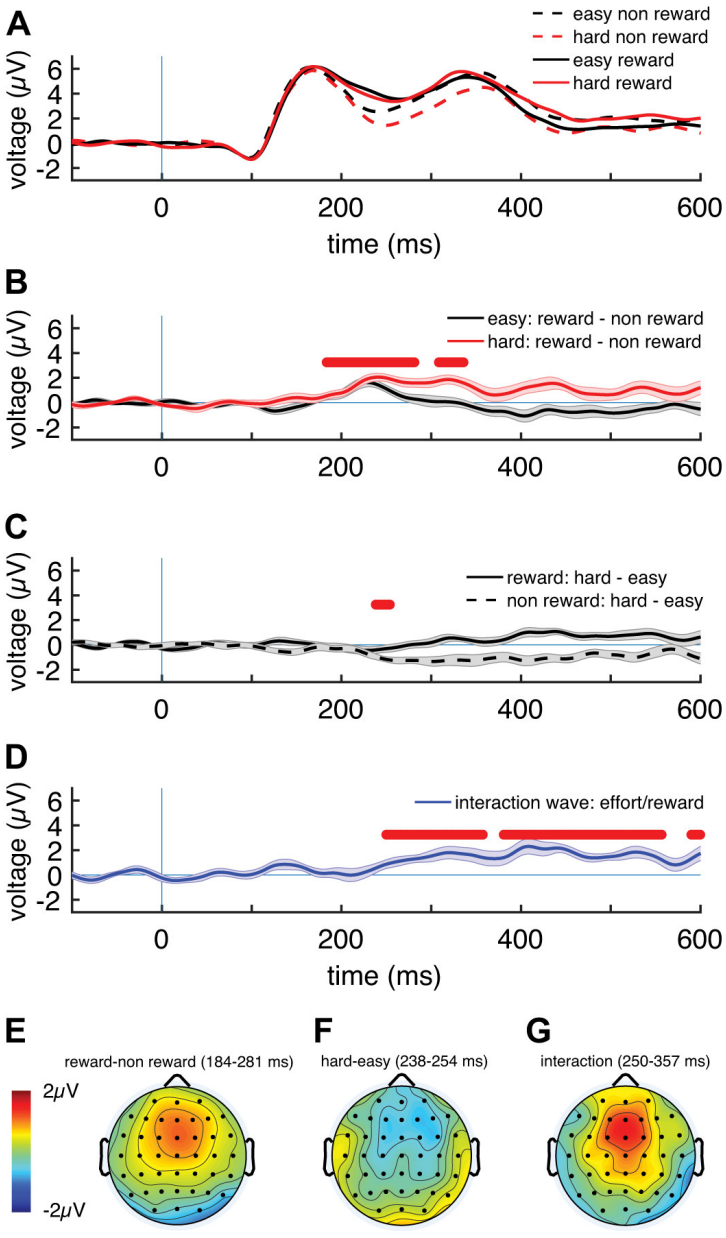

Figure 5-3: Event-related potential elicited by reinforcement feedback. ***A***: trial-averaged ERPs recorded from electrode FCz aligned to reinforcement feedback presentation (0 ms: vertical blue line), selected for reinforcement outcome (reward or nonreward) and the physical effort requirement on that trial (easy or hard). ***B***: mean difference waves computed as reward ERP − nonreward ERP

separately for the easy and hard effort conditions (shaded region: ±SE). Red markers indicate time points between 100 and 600 ms with significant main effect of reward outcome [$P < 0.05$, false discovery rate (FDR) corrected]. ***C***: mean difference waves computed as hard effort ERP − easy effort ERP separately for the reward and nonreward effort conditions (shaded region: ±SE). Red marker indicates time points between 100 and 600 ms with significant main effect of effort condition ($P < 0.05$, FDR corrected). ***D***: mean interaction wave computed as (hard reward ERP − hard nonreward ERP) − (easy reward ERP − easy nonreward ERP). Shaded region: ±SE. Red markers indicate time points between 100 and 600 ms with significant interaction effect of reward outcome ($P < 0.05$, FDR corrected). ***E***: scalp distribution of reward − nonreward ERPs, irrespective of effort condition, between 184 and 281 ms (1st cluster of significant time points shown in *B*). ***F***: scalp distribution of hard effort − easy effort ERPs, irrespective of reward outcome, between 238 and 254 ms (cluster of significant time points shown in *C*). ***G***: scalp distribution of interaction wave between 250 and 357 ms (1st cluster of significant time points shown in *D*).

In the control condition, participants performed the EMG production task, but the effort condition was predetermined and no reinforcement feedback was provided. We found no reliable differences between the control easy and control hard ERPs (ranges for all time points between 100 and 600 ms: $t = [-1.56\ 1.42]$, uncorrected $P = [1.00\ 0.139]$).

After participants produced binary decisions by button press, feedback was provided to indicate the resulting effort condition for the current trial. Figure 5-4A shows the ERPs elicited on each trial by feedback indicating the effort condition. We found significant differences between the easy feedback and hard feedback ERPs between 373 and 522 ms after effort feedback (Figure 5-4*B*; ranges for significant time points: $t = [-4.31\ -2.81]$, $P = [0.049\ 0.015]$, uncorrected $P = [0.012\ 0.0005]$). Although we only performed statistical testing up to 600 ms to avoid sacrificing statistical power, we observed that the easy feedback − hard feedback difference wave remains at least 1 standard error below zero until 1,018 ms after feedback.

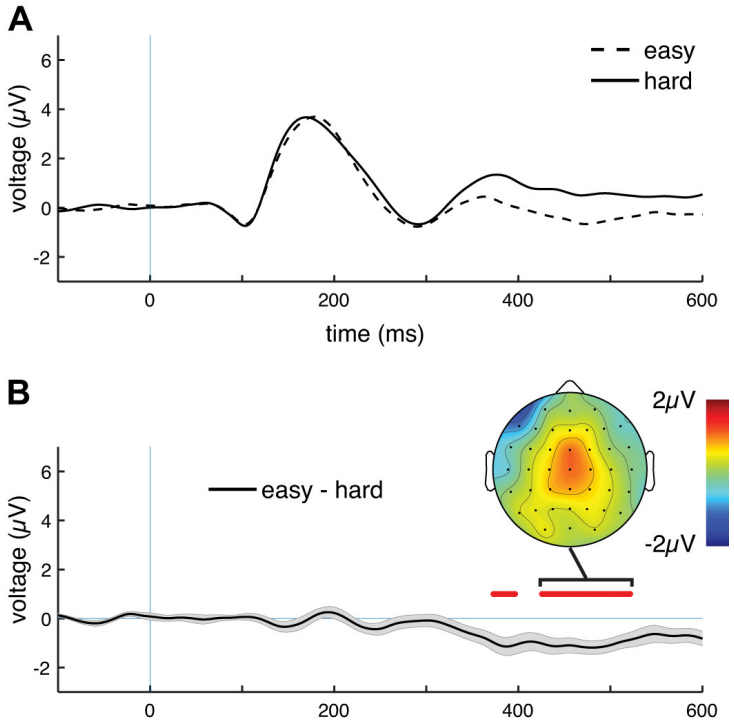

Figure 5-4: Event-related potentials elicited by effort condition feedback. ***A***: trial-averaged ERPs recorded from electrode FCz aligned to effort condition feedback presentation (0 ms: vertical blue line), selected for physical effort condition (easy or hard). ***B***: mean difference waves computed as easy feedback ERP − hard feedback ERP (shaded region: ±SE). Red markers indicate time points between 100 and 600 ms significantly different from 0 [$P < 0.05$, false discovery rate (FDR) corrected]. *Inset*: scalp distribution of difference wave between 426 and 521 ms.

5.3.3 IC-ERP results

The ERP responses to reinforcement feedback at electrode FCz clearly contained multiple components (Figure 5-3), with the most obvious contributions being from a relatively early positive deflection peaking between 164 and 172 ms after feedback onset depending on the feedback condition and a later positive deflection peaking between 334 and 359 ms. Independent component analysis applied to individual participants' data consistently outputted ICs corresponding to these ERP components. We identified ICs that resembled neural EEG activity and computed the average ERPs of IC activity projected onto electrode FCz. For each participant, we identified the early ERP component by selecting the IC with the largest maximal positive peak occurring before 200 ms. We identified the later component by selecting the IC with the largest maximal positive peak after 200 ms. We identified neural ICs containing maximal peaks in the early time window for 16/18 participants and in the late time window for 17/18 participants.

Early ERP component: Figure 5-5A shows the IC-ERPs elicited by reinforcement feedback for the early ERP components. We performed 2 × 2 repeated-measures ANOVA on IC-ERP amplitude at each individual time point 100–600 ms after feedback onset. *P* values are corrected for multiple comparisons across time points with FDR. We found reliable main effects of reward outcome between 164 and 342 ms after feedback, with larger IC-ERP amplitude for reward compared with nonreward (Figure 5-5*C*; ranges for significant time points: F = [8.10 34.86], P = [0.047 0.0011], uncorrected P = [0.012 <0.0001]). We found no reliable main effects of effort on the response to reinforcement feedback (Figure 5-5*E*; ranges for all time points 100–600 ms: F = [0.00 4.97], uncorrected P = [1.00 0.0416]). We also found no reliable interaction effects between effort and reward (Figure 5-5*G*; ranges for all time points 100–600 ms: F = [0.00 5.58], uncorrected P = [0.99 0.032]). Figure 5-6*A* shows the IC-ERPs elicited by effort feedback for the early ERP components. We observed no difference between the response to easy effort and hard effort feedback (Figure 5-6*C*; ranges for all time points between 100 and 600 ms: t = [−0.94 1.67], uncorrected P = [1.00 0.12]). We observed no differences between the

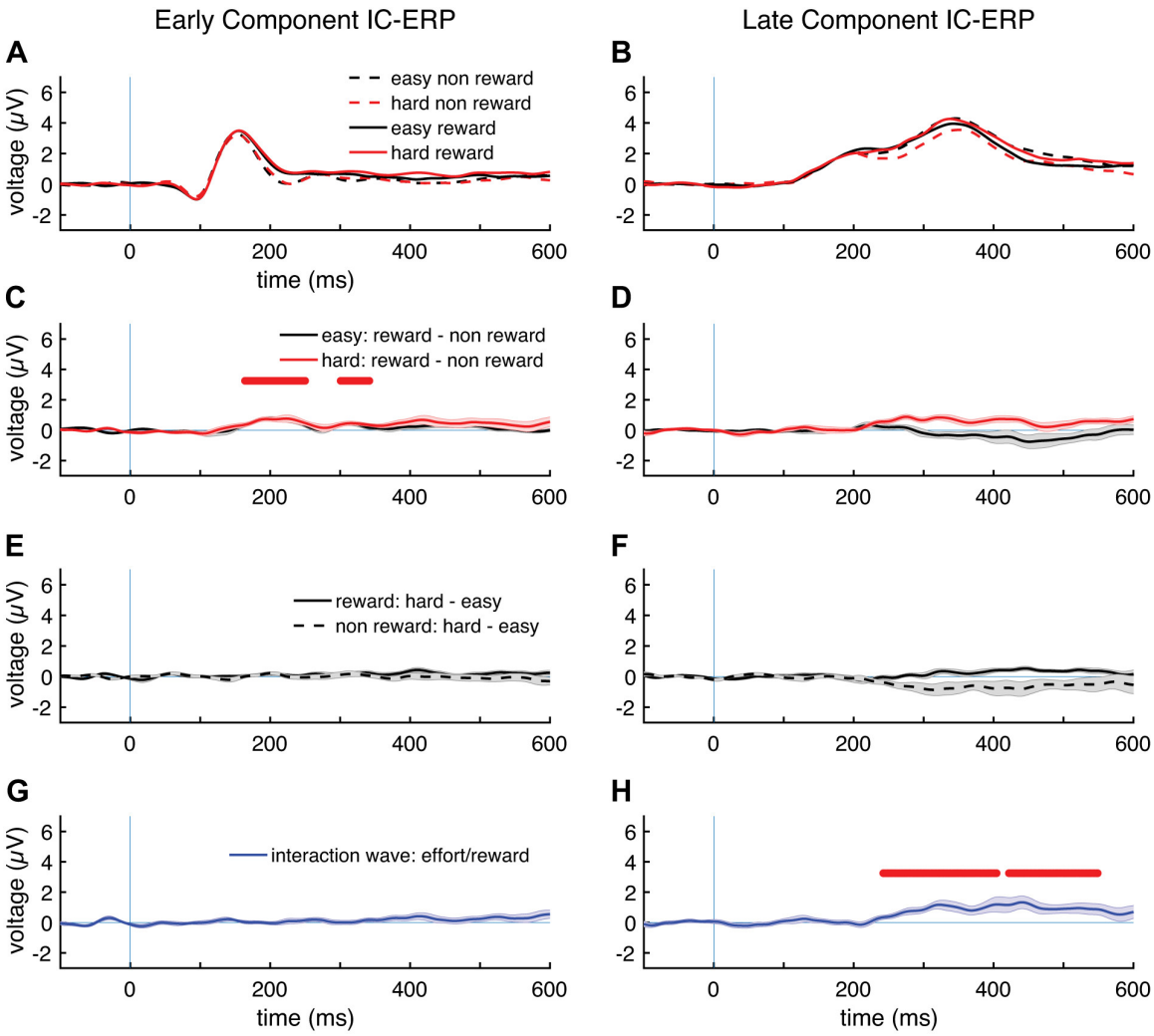

Figure 5-5: Independent components of event-related potentials (IC-ERPs) elicited by reinforcement feedback. ***A*** and ***B***: trial-averaged IC-ERPs for the early and late IC-ERP components back-projected to electrode FCz, aligned to reinforcement feedback presentation (0 ms: vertical blue line), selected for reinforcement outcome (reward or nonreward) and the physical effort requirement on that trial (easy or hard). ***C*** and ***D***: mean difference waves computed as reward IC-ERP − nonreward IC-ERP separately for the easy and hard effort conditions (shaded region: ±SE). Red markers indicate time points between 100 and 600 ms with significant main effect of reward outcome [$P < 0.05$, false discovery rate (FDR) corrected]. ***E*** and ***F***: mean difference waves computed as hard effort IC-ERP − easy effort IC-ERP separately for the reward and nonreward effort conditions (shaded region: ±SE). ***G*** and ***H***: mean interaction wave computed as (hard reward IC-ERP − hard nonreward IC-ERP) − (easy reward IC-ERP − easy nonreward IC-ERP). Shaded region: ±SE. Red markers indicate time points between 100 and 600 ms with significant interaction effect of reward outcome ($P < 0.05$, FDR corrected).

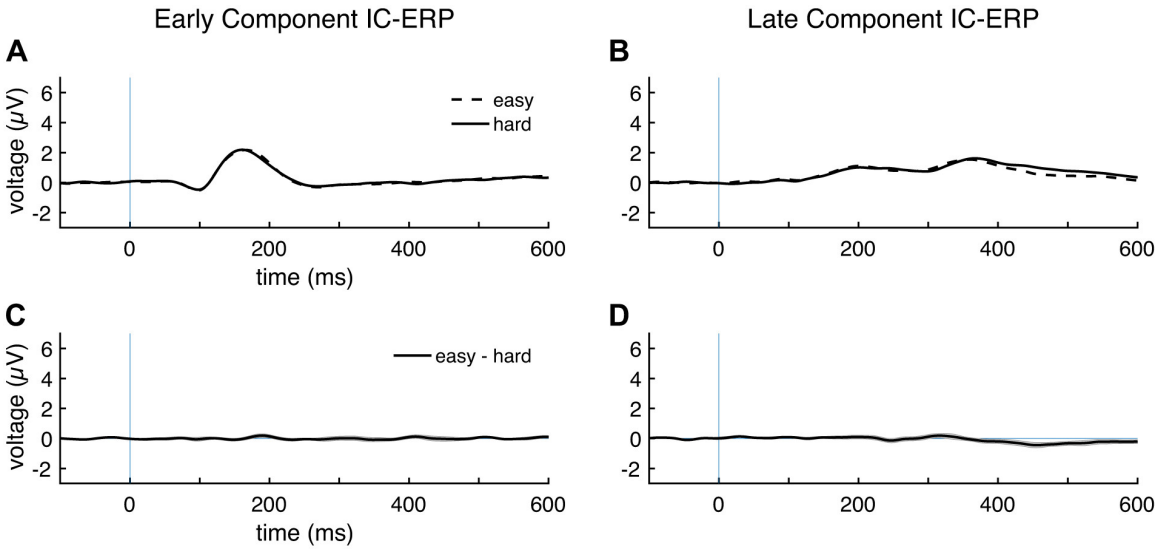

Figure 5-6: Independent components of event-related potentials (IC-ERPs) elicited by effort condition feedback. A and **B**: trial-averaged IC-ERPs for the early and late IC-ERP components back-projected to electrode FCz, aligned to effort condition feedback presentation (0 ms: vertical blue line), selected for physical effort condition (easy or hard). **C** and **D**: mean difference waves computed as easy feedback IC-ERP – hard feedback IC-ERP (shaded region: ±SE).

control easy and control hard IC-ERPs (ranges for all time points between 100 and 600 ms: t = [−1.38 2.46], uncorrected P = [0.027 0.99]).

Late ERP component: Figure 5-5*B* shows the IC-ERPs elicited by reinforcement feedback for the late ERP components. We found no reliable main effects of reward outcome on the response to reinforcement feedback (Figure 5-5*D*; ranges for all time points 100–600 ms: F = [0.00 16.40], uncorrected P = [1.00 0.001]). We also found no reliable main effects of effort on the response to reinforcement feedback (Figure 5-5*F*; ranges for all time points 100–600 ms: F = [0.00 4.70], uncorrected P = [0.99 0.046]). We did observe an

interaction effect between effort and reward, with significant time points between 242 and 549 ms (Figure 5-5*H*; ranges for significant time points: F = [5.78 24.06], P = [0.0492 0.010], uncorrected P = [0.0287 0.0002]). Figure 5-6*B* shows the IC-ERPs elicited by effort feedback for the late ERP components. We observed no difference between the response to easy effort and hard effort feedback (Figure 5-6*D*; ranges for all time points between 100 and 600 ms: t = [−2.89 1.29], uncorrected P = [1.00 0.011]). We observed no differences between the control easy and control hard IC-ERPs (ranges for all time points between 100 and 600 ms: t = [−0.27 2.01], uncorrected P = [0.97 0.06]).

To directly compare the IC-ERP responses of the early and late components, we also analyzed both components with a single statistical model. We performed three-way repeated-measures ANOVA at each time point between 100 and 600 ms, with factors reward outcome, effort, and IC (early vs. late ERP component). We included the 15 participants for whom both early and late components were identified. P values are corrected for multiple comparisons across time points with FDR. We found reliable main effects of reward outcome between 170 and 342 ms (ranges for significant time points: F = [7.39 29.87], P = [0.0451 0.0030], uncorrected P = [0.0167 0.0001]). We found reliable interaction effects between effort and reward between 238 and 535 ms (ranges for significant time points: F = [6.071 14.13], P = [0.0494 0.0425], uncorrected P = [0.0273 0.0021]). We found reliable three-way interaction effects between reward, effort, and IC between 299 and 325 ms (ranges for significant time points: F = [13.75 22.24], P = [0.0429 0.0192], uncorrected P = [0.0023 0.0003]). The three-way interaction reflects a larger interaction between reward and effort in the late component IC-ERP than the early component IC-ERP. Separate analyses revealed a reliable effect of reward in the early component but not the late component. This was reflected in a trend toward interaction effects between IC and reward outcome, although these effects were not reliable after correction for multiple comparisons across time (ranges for significant time points uncorrected: time = [172 221 ms], F = [4.65 10.65], uncorrected P = [0.0489 0.0057]).

5.4 Discussion

Participants were more likely to switch responses after choices that led to hard effort than easy effort, suggesting that they adapted behavior to reduce physical effort in response to uncertain outcomes. At the end of each trial, binary reinforcement feedback indicated whether participants achieved a monetary reward, which depended on precisely producing a target level of EMG activity. Unsurprisingly, reinforcement feedback elicited a robust feedback related negativity/reward positivity (FRN/RP) response, measured as a relative positivity in the ERPs elicited by reward compared to those elicited by nonreward feedback over the medial frontal scalp.

Samplewise analysis revealed interesting temporal dynamics of effort and reward processing in the midfrontal EEG during outcome evaluation. After reinforcement feedback was delivered, an effect of reward outcome first emerged with a latency of 184 ms, which remained significant while an additional main effect of preceding effort emerged at 238 ms. Finally, a sustained interaction between reward and effort first occurred around 250 ms after feedback onset. These dynamics suggest a process whereby upon receiving reward feedback the brain first encodes the immediate reward outcome and subsequently integrates signals related to the preceding effort. This process culminates in an interaction whereby the effect of reward outcome depends on the preceding effort.

The main effect of reward outcome, which is generally the definitive feature of the FRN/RP, occurred with a typical spatial and temporal distribution. The interaction effect showed similar medial frontal scalp topography and substantially overlapping temporal properties, suggesting that it may originate in the same neural generator. However, the interaction effect persisted upwards of 600 ms. Although the FRN/RP is not typically measured beyond 400 ms, meta-analysis has shown sensitivity in medial frontal ERPs to reward prediction error upwards of 500 ms (Sambrook & Goslin, 2015). These long-latency effects may be attributed to the P300, which is modulated by various features of reinforcement processing and can overlap with the FRN/RP (Glazer et al., 2018). We used

independent component analysis to separate the contributions of various neural sources to the scalp ERP. We found that the main effect of reward and the interaction between effort and reward reliably load onto separate sets of independent components. We take this as evidence that these effects arise in distinct neural sources, as a single source cannot produce multiple effects with separable scalp projections. However, we do not claim that the early and late IC-ERPs necessarily represent purely isolated single ERP components, as our ICA procedure could fail to separate distinct sources. Although ICA should ameliorate issues due to component overlap, they cannot be verifiably ruled out. Nonetheless, the reward × effort interaction effect showed latencies and IC loadings consistent with P300 effects.

Although the FRN/RP is reliably sensitive to outcome valence and likelihood aspects of reward prediction error, reward magnitude seems to be coded independently in the P300 (Sambrook & Goslin, 2015; Sato et al., 2005; Walsh & Anderson, 2012; Y. Wu & Zhou, 2009; Yeung & Sanfey, 2004). Thus, physical effort may modulate the subjective magnitude of reinforcement outcomes. The P300 elicited by outcome processing has also been shown to be affected by various high-level properties such as motivational salience, temporal waiting cost, and cognitive effort (Glazer et al., 2018; Ma et al., 2014; San Martín, 2012; J. Wang et al., 2014). Multiple variants of the P300 have been reported with medial frontal or posterior scalp distributions, and widespread cortical association networks are implicated including parietal, temporal, and prefrontal regions (Polich, 2007; Soltani & Knight, 2000).

Although these results show a clear effect of the physical effort associated with an action on the neural processing of subsequent reinforcement outcomes, specific interpretation of the effects depends on the theoretical understanding of the underlying components. The prominent interaction effect of effort and reward observed in the present study shows that effort increases the differential neural response to reinforcement outcomes. This could be explained either by effort increasing a positive ERP deflection to reward or a negative ERP deflection to nonreward. Due to the possibility of ERP component overlap, simple effects cannot necessarily be interpreted in the case of an interaction

effect on ERP amplitude (Sambrook & Goslin, 2015, 2016). An fMRI study by Hernandez Lallement et al., (2014) found that cognitive effort increased neural sensitivity to both reward and loss, with reward sensitivity being modulated in the anterior cingulate and nucleus accumbens and loss sensitivity being increased in the anterior insula.

An increased neural response to nonreward due to effort is consistent with economic theories whereby a motor cost would further devalue a nonreward outcome. An increased neural response to reward may be consistent with other paradoxical findings reported in the literature. Although normative economic models of behavior predict that effort costs should devalue reward, it has often been reported in humans and other species that rewards produce stronger reinforcement when they require more effort to obtain (Clement et al., 2000; Inzlicht et al., 2018; Lydall et al., 2010; Zentall, 2010). Unfortunately, we were not able to assess such a behavioral interaction in the present study as there was no effect of reward outcome on the binary decisions that participants made on each trial. This was not surprising, as reward outcome was not determined by these decisions but rather by performance on the EMG production task. The binary decisions only determined the effort required, and the reinforcement threshold was controlled to produce approximately equal reward rate in both effort conditions.

After participants produced binary responses, feedback indicated the resulting physical effort condition for the subsequent EMG production portion of the trial. In line with theoretical accounts of the FRN/RP as a temporal difference learning prediction error signal, stimuli that predict aversive outcomes or economic loss typically elicit FRN/RP responses (Mulligan & Hajcak, 2018). Thus, we predicted that effort feedback might elicit an FRN/RP component as a learning signal for effort minimization. However, we observed no FRN/RP modulation when comparing the ERP responses to feedback indicating easy or hard effort trials. Rather, effort modulated the response to reinforcement feedback at the end of the trial. This suggests that physical effort is not immediately treated by the reinforcement learning system as a loss or a punishing stimulus. Rather, effort information can be maintained during the course of an action and incorporated with reward information at the time of outcome evaluation.

We often undertake protracted tasks for which the effort requirements and ultimate payoffs are uncertain. It may not be efficient to punish the value representation of a task every time an unexpected effort is encountered, as the eventual payoff may be well worth the effort. Instead, it may be more efficient to integrate effort over the entire course of an undertaking and evaluate the cost and benefit simultaneously when the final outcome is observed. This process can also support interactions in which the effect of effort depends on reward that is only delivered later. Alternatively, some work suggests that we learn about effort requirements and reward separately and integrate them at the time of decision making (Hauser et al., 2017; Skvortsova et al., 2014). It is likely that economic decision making and learning involves distributed hierarchical computations and that it is possible to observe a distribution of signals with varying dependencies on effort, reward, and integrated utility throughout the brain (Hunt & Hayden, 2017).

Some limitations of this study should be noted. Participants adapted their behavior to reduce physical effort, but the behavioral effect of effort was variable and relatively weak. Participants were more likely to switch responses after choices that led to hard effort than easy effort. However, participants often switched responses after easy trials or stayed with responses that produced hard effort: on average, participants switched responses after 46.5% of easy trials and 60% of hard trials. Furthermore, negative coefficients for the effect of effort on switching were estimated for several participants. The relatively weak and highly variable effects of effort are consistent with the notion that although effort is generally treated as a cost that is minimized, in many cases people are undeterred by effort or even purposefully select more effortful options (Eisenberger, 1992; Inzlicht et al., 2018). These effects are often attributed to state-dependent learning in which reinforcement outcomes are evaluated relative to the value of the current state. In the present study, variable reinforcement outcomes were only evaluated after effort production and thus may have been more valuable when received after a costly, high-effort action. Other details of the task may have affected effort-related choice. Unlike some previous studies of effort minimization, participants were not instructed to avoid effort. Furthermore, success in the task was not dependent on exerting effort that

exceeded an unknown criterion. These features may enhance effort minimization, but they could also conflate effort prediction errors with errors relative to the goals of the task at hand, which are also strongly represented in the ACC (Fu et al., 2019; Holroyd & Krigolson, 2007; Swick & Turken, 2002; Ullsperger et al., 2014).

Although the excellent temporal resolution offered by EEG proved instrumental in uncovering the dynamics of effort and reward processing in the brain, it invariably measures a mixture of signals from neurons with different response properties. Kennerley et al., (2011) identified diverse tuning to economic value across ACC, orbitofrontal cortex, and lateral prefrontal cortex, such that many neurons that are selective to value with opposite tunings will cancel out at the population level measured by EEG of fMRI. We report effects with midfrontal scalp topographies. However, EEG measured at the scalp is difficult to localize and can represent mixtures of activity from entirely separate brain regions. Although the FRN/RP is a well-characterized response and convergent lines of evidence suggest a source in the ACC, we did not attempt any source localization. The spatial localization of the measured signals remains speculative.

5.5 References

Acunzo, D. J., Mackenzie, G., & van Rossum, M. C. W. (2012). Systematic biases in early ERP and ERF components as a result of high-pass filtering. *Journal of Neuroscience Methods*, *209*(1), 212–218. https://doi.org/10.1016/j.jneumeth.2012.06.011

Alessandri, J., Darcheville, J.-C., Delevoye-Turrell, Y., & Zentall, T. R. (2008). Preference for rewards that follow greater effort and greater delay. *Learning & Behavior*, *36*(4), 352–358. https://doi.org/10.3758/LB.36.4.352

Amiez, C., Joseph, J.-P., & Procyk, E. (2005). Anterior cingulate error-related activity is modulated by predicted reward. *The European Journal of Neuroscience*, *21*(12), 3447–3452. https://doi.org/10.1111/j.1460-9568.2005.04170.x

Bayer, H. M., & Glimcher, P. W. (2005). Midbrain Dopamine Neurons Encode a Quantitative Reward Prediction Error Signal. *Neuron*, *47*(1), 129–141. https://doi.org/10.1016/j.neuron.2005.05.020

Becker, M. P. I., Nitsch, A. M., Miltner, W. H. R., & Straube, T. (2014). A single-trial estimation of the feedback-related negativity and its relation to BOLD responses in a time-estimation task. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *34*(8), 3005–3012. https://doi.org/10.1523/JNEUROSCI.3684-13.2014

Botvinick, M. M., Huffstetler, S., & McGuire, J. T. (2009). Effort discounting in human nucleus accumbens. *Cognitive, Affective & Behavioral Neuroscience*, *9*(1), 16–27. https://doi.org/10.3758/CABN.9.1.16

Clement, T. S., Feltus, J. R., Kaiser, D. H., & Zentall, T. R. (2000). "work ethic" in pigeons: Reward value is directly related to the effort or time required to obtain the reward. *Psychonomic Bulletin & Review*, *7*(1), 100–106. https://doi.org/10.3758/BF03210727

Cohen, M. X., & Ranganath, C. (2007). Reinforcement Learning Signals Predict Future Decisions. *Journal of Neuroscience*, *27*(2), 371–378. https://doi.org/10.1523/JNEUROSCI.4421-06.2007

Cos, I., Bélanger, N., & Cisek, P. (2011). The influence of predicted arm biomechanics on decision making. *Journal of Neurophysiology*, *105*(6), 3022–3033. https://doi.org/10.1152/jn.00975.2010

Cos, I., Duque, J., & Cisek, P. (2014). Rapid prediction of biomechanical costs during action decisions. *Journal of Neurophysiology*, *112*(6), 1256–1266. https://doi.org/10.1152/jn.00147.2014

Croxson, P. L., Walton, M. E., O'Reilly, J. X., Behrens, T. E. J., & Rushworth, M. F. S. (2009). Effort-Based Cost–Benefit Valuation and the Human Brain. *Journal of Neuroscience*, *29*(14), 4531–4541. https://doi.org/10.1523/JNEUROSCI.4515-08.2009

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Denk, F., Walton, M. E., Jennings, K. A., Sharp, T., Rushworth, M. F. S., & Bannerman, D. M. (2005). Differential involvement of serotonin and dopamine systems in cost-benefit decisions about delay or effort. *Psychopharmacology*, *179*(3), 587–596. https://doi.org/10.1007/s00213-004-2059-4

Eisenberger, R. (1992). Learned industriousness. *Psychological Review*, *99*(2), 248–267. https://doi.org/10.1037/0033-295x.99.2.248

Emeric, E. E., Brown, J. W., Leslie, M., Pouget, P., Stuphorn, V., & Schall, J. D. (2008). Performance monitoring local field potentials in the medial frontal cortex of primates: Anterior cingulate cortex. *Journal of Neurophysiology*, *99*(2), 759–772. https://doi.org/10.1152/jn.00896.2006

Fu, Z., Wu, D.-A. J., Ross, I., Chung, J. M., Mamelak, A. N., Adolphs, R., & Rutishauser, U. (2019). Single-Neuron Correlates of Error Monitoring and Post-Error Adjustments in Human Medial Frontal Cortex. *Neuron*, *101*(1), 165-177.e5. https://doi.org/10.1016/j.neuron.2018.11.016

Gehring, W. J., & Willoughby, A. R. (2002). The Medial Frontal Cortex and the Rapid Processing of Monetary Gains and Losses. *Science*, *295*(5563), 2279–2282. https://doi.org/10.1126/science.1066893

Gheza, D., De Raedt, R., Baeken, C., & Pourtois, G. (2018). Integration of reward with cost anticipation during performance monitoring revealed by ERPs and EEG spectral perturbations. *NeuroImage*, *173*, 153–164. https://doi.org/10.1016/j.neuroimage.2018.02.049

Gläscher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). States versus rewards: Dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*, *66*(4), 585–595. https://doi.org/10.1016/j.neuron.2010.04.016

Glazer, J. E., Kelley, N. J., Pornpattananangkul, N., Mittal, V. A., & Nusslock, R. (2018). Beyond the FRN: Broadening the time-course of EEG and ERP components implicated in reward processing. *International Journal of Psychophysiology*, *132*, 184–202. https://doi.org/10.1016/j.ijpsycho.2018.02.002

Graybiel, A. M. (2008). Habits, rituals, and the evaluative brain. *Annual Review of Neuroscience*, *31*, 359–387. https://doi.org/10.1146/annurev.neuro.29.051605.112851

Hartmann, M. N., Hager, O. M., Tobler, P. N., & Kaiser, S. (2013). Parabolic discounting of monetary rewards by physical effort. *Behavioural Processes*, *100*, 192–196. https://doi.org/10.1016/j.beproc.2013.09.014

Hassall, C. D., Hajcak, G., & Krigolson, O. E. (2019). The importance of agency in human reward processing. *Cognitive, Affective & Behavioral Neuroscience*, *19*(6), 1458–1466. https://doi.org/10.3758/s13415-019-00730-2

Hauser, T. U., Eldar, E., & Dolan, R. J. (2017). Separate mesocortical and mesolimbic pathways encode effort and reward learning signals. *Proceedings of the National Academy of Sciences*, *114*(35), E7395–E7404. https://doi.org/10.1073/pnas.1705643114

Hauser, T. U., Iannaccone, R., Stämpfli, P., Drechsler, R., Brandeis, D., Walitza, S., & Brem, S. (2014). The feedback-related negativity (FRN) revisited: New insights into the localization, meaning and network organization. *NeuroImage*, *84*, 159–168. https://doi.org/10.1016/j.neuroimage.2013.08.028

Hernandez Lallement, J., Kuss, K., Trautner, P., Weber, B., Falk, A., & Fliessbach, K. (2014). Effort increases sensitivity to reward and loss magnitude in the human brain. *Social Cognitive and Affective Neuroscience*, *9*(3), 342–349. https://doi.org/10.1093/scan/nss147

Holinger, D. P., Hill, S. A., Martin, D. L., Faux, S. F., Ives, J. R., & Schomer, D. L. (2000). Reappraisal of filter effects on P300 voltage and latency. *Journal of Clinical Neurophysiology: Official Publication of the American Electroencephalographic Society*, *17*(3), 331–335. https://doi.org/10.1097/00004691-200005000-00011

Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*(4), 679–709. https://doi.org/10.1037/0033-295X.109.4.679

Holroyd, C. B., & Krigolson, O. E. (2007). Reward prediction error signals associated with a modified time estimation task. *Psychophysiology*, *44*(6), 913–917. https://doi.org/10.1111/j.1469-8986.2007.00561.x

Hosking, J. G., Floresco, S. B., & Winstanley, C. A. (2015). Dopamine Antagonism Decreases Willingness to Expend Physical, But Not Cognitive, Effort: A Comparison of Two Rodent Cost/Benefit Decision-Making Tasks. *Neuropsychopharmacology*, *40*(4), 1005–1015. https://doi.org/10.1038/npp.2014.285

Hunt, L. T., & Hayden, B. Y. (2017). A distributed, hierarchical and recurrent framework for reward-based choice. *Nature Reviews. Neuroscience*, *18*(3), 172–182. https://doi.org/10.1038/nrn.2017.7

Inzlicht, M., Shenhav, A., & Olivola, C. Y. (2018). The Effort Paradox: Effort Is Both Costly and Valued. *Trends in Cognitive Sciences*, *22*(4), 337–349. https://doi.org/10.1016/j.tics.2018.01.007

Ito, S., Stuphorn, V., Brown, J. W., & Schall, J. D. (2003). Performance monitoring by the anterior cingulate cortex during saccade countermanding. *Science (New York, N.Y.)*, *302*(5642), 120–122. https://doi.org/10.1126/science.1087847

Kennerley, S. W., Behrens, T. E. J., & Wallis, J. D. (2011). Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. *Nature Neuroscience*, *14*(12), 1581–1589. https://doi.org/10.1038/nn.2961

Kennerley, S. W., Dahmubed, A. F., Lara, A. H., & Wallis, J. D. (2009). Neurons in the Frontal Lobe Encode the Value of Multiple Decision Variables. *Journal of Cognitive Neuroscience*, *21*(6), 1162–1178. https://doi.org/10.1162/jocn.2009.21100

Klein-Flügge, M. C., Kennerley, S. W., Friston, K., & Bestmann, S. (2016). Neural Signatures of Value Comparison in Human Cingulate Cortex during Decisions Requiring an Effort-Reward Trade-off. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *36*(39), 10002–10015. https://doi.org/10.1523/JNEUROSCI.0292-16.2016

Kurniawan, I. T., Guitart-Masip, M., & Dolan, R. J. (2011). Dopamine and Effort-Based Decision Making. *Frontiers in Neuroscience*, *5*. https://doi.org/10.3389/fnins.2011.00081

Kurniawan, I. T., Seymour, B., Talmi, D., Yoshida, W., Chater, N., & Dolan, R. J. (2010). Choosing to Make an Effort: The Role of Striatum in Signaling Physical Effort of a Chosen Action. *Journal of Neurophysiology*, *104*(1), 313–321. https://doi.org/10.1152/jn.00027.2010

Lydall, E. S., Gilmour, G., & Dwyer, D. M. (2010). Rats place greater value on rewards produced by high effort: An animal analogue of the "effort justification" effect. *Journal of Experimental Social Psychology*, *46*(6), 1134–1137. https://doi.org/10.1016/j.jesp.2010.05.011

Ma, Q., Meng, L., Wang, L., & Shen, Q. (2014). I endeavor to make it: Effort increases valuation of subsequent monetary reward. *Behavioural Brain Research*, *261*, 1–7. https://doi.org/10.1016/j.bbr.2013.11.045

Martin, L. E., & Potts, G. F. (2011). Medial frontal event-related potentials and reward prediction: Do responses matter? *Brain and Cognition*, *77*(1), 128–134. https://doi.org/10.1016/j.bandc.2011.04.001

Mathewson, K. J., Dywan, J., Snyder, P. J., Tays, W. J., & Segalowitz, S. J. (2008). Aging and electrocortical response to error feedback during a spatial learning task. *Psychophysiology*, *45*(6), 936–948. https://doi.org/10.1111/j.1469-8986.2008.00699.x

Miltner, W. H., Braun, C. H., & Coles, M. G. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: Evidence for a "generic" neural system for error detection. *Journal of Cognitive Neuroscience*, *9*(6), 788–798. https://doi.org/10.1162/jocn.1997.9.6.788

Morel, P., Ulbrich, P., & Gail, A. (2017). What makes a reach movement effortful? Physical effort discounting supports common minimization principles in decision making and motor control. *PLOS Biology*, *15*(6), e2001323. https://doi.org/10.1371/journal.pbio.2001323

Mulligan, E. M., & Hajcak, G. (2018). The electrocortical response to rewarding and aversive feedback: The reward positivity does not reflect salience in simple gambling tasks. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *132*(Pt B), 262–267. https://doi.org/10.1016/j.ijpsycho.2017.11.015

Onton, J., & Makeig, S. (2006). Information-based modeling of event-related brain dynamics. *Progress in Brain Research*, *159*, 99–120. https://doi.org/10.1016/S0079-6123(06)59007-7

Palidis, D. J., Cashaback, J. G. A., & Gribble, P. L. (2019). Neural signatures of reward and sensory error feedback processing in motor learning. *Journal of Neurophysiology*, *121*(4), 1561–1574. https://doi.org/10.1152/jn.00792.2018

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*, *442*(7106), 1042–1045. https://doi.org/10.1038/nature05051

Pfabigan, D. M., Alexopoulos, J., Bauer, H., & Sailer, U. (2011). Manipulation of feedback expectancy and valence induces negative and positive reward prediction error signals manifest in event-related brain potentials. *Psychophysiology*, *48*(5), 656–664. https://doi.org/10.1111/j.1469-8986.2010.01136.x

Polich, J. (2007). Updating P300: An Integrative Theory of P3a and P3b. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, *118*(10), 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019

Porter, B. S., Hillman, K. L., & Bilkey, D. K. (2019). Anterior cingulate cortex encoding of effortful behavior. *Journal of Neurophysiology*, *121*(2), 701–714. https://doi.org/10.1152/jn.00654.2018

Prévost, C., Pessiglione, M., Météreau, E., Cléry-Melin, M.-L., & Dreher, J.-C. (2010). Separate Valuation Subsystems for Delay and Effort Decision Costs. *Journal of Neuroscience*, *30*(42), 14080–14090. https://doi.org/10.1523/JNEUROSCI.2752-10.2010

Puig, M. V., & Miller, E. K. (2012). The role of prefrontal dopamine D1 receptors in the neural mechanisms of associative learning. *Neuron*, *74*(5), 874–886. https://doi.org/10.1016/j.neuron.2012.04.018

Rangel, A., & Hare, T. (2010). Neural computations associated with goal-directed choice. *Current Opinion in Neurobiology*, *20*(2), 262–270. https://doi.org/10.1016/j.conb.2010.03.001

Rudebeck, P. H., Behrens, T. E., Kennerley, S. W., Baxter, M. G., Buckley, M. J., Walton, M. E., & Rushworth, M. F. S. (2008). Frontal cortex subregions play distinct roles in choices between actions and stimuli. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *28*(51), 13775–13785. https://doi.org/10.1523/JNEUROSCI.3541-08.2008

Rudebeck, P. H., Walton, M. E., Smyth, A. N., Bannerman, D. M., & Rushworth, M. F. S. (2006). Separate neural pathways process different decision costs. *Nature Neuroscience*, *9*(9), 1161–1168. https://doi.org/10.1038/nn1756

Salamone, J. D., Correa, M., Farrar, A., & Mingote, S. M. (2007). Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits. *Psychopharmacology*, *191*(3), 461–482. https://doi.org/10.1007/s00213-006-0668-9

Salamone, J. D., Correa, M., Mingote, S., & Weber, S. M. (2003). Nucleus accumbens dopamine and the regulation of effort in food-seeking behavior: Implications for studies of natural motivation, psychiatry, and drug abuse. *The Journal of Pharmacology and Experimental Therapeutics*, *305*(1), 1–8. https://doi.org/10.1124/jpet.102.035063

Sambrook, T. D., & Goslin, J. (2015). A neural reward prediction error revealed by a meta-analysis of ERPs using great grand averages. *Psychological Bulletin*, *141*(1), 213–235. http://dx.doi.org.proxy1.lib.uwo.ca/10.1037/bul0000006

Sambrook, T. D., & Goslin, J. (2016). Principal components analysis of reward prediction errors in a reinforcement learning task. *NeuroImage*, *124*, 276–286. https://doi.org/10.1016/j.neuroimage.2015.07.032

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

Sato, A., Yasuda, A., Ohira, H., Miyawaki, K., Nishikawa, M., Kumano, H., & Kuboki, T. (2005). Effects of value and reward magnitude on feedback negativity and P300. *NeuroReport*, *16*(4), 407–411.

Schevernels, H., Krebs, R. M., Santens, P., Woldorff, M. G., & Boehler, C. N. (2014). Task preparation processes related to reward prediction precede those related to task-difficulty expectation. *NeuroImage*, *84*, 639–647. https://doi.org/10.1016/j.neuroimage.2013.09.039

Schultz, W. (2006). Behavioral theories and the neurophysiology of reward. *Annual Review of Psychology*, *57*, 87–115. https://doi.org/10.1146/annurev.psych.56.091103.070229

Schweimer, J., Saft, S., & Hauber, W. (2005). Involvement of catecholamine neurotransmission in the rat anterior cingulate in effort-related decision making. *Behavioral Neuroscience*, *119*(6), 1687–1692. https://doi.org/10.1037/0735-7044.119.6.1687

Selinger, J. C., O'Connor, S. M., Wong, J. D., & Donelan, J. M. (2015). Humans Can Continuously Optimize Energetic Cost during Walking. *Current Biology*, *25*(18), 2452–2456. https://doi.org/10.1016/j.cub.2015.08.016

Selinger, J. C., Wong, J. D., Simha, S. N., & Donelan, J. M. (2019). How humans initiate energy optimization and converge on their optimal gaits. *The Journal of Experimental Biology*, *222*(Pt 19). https://doi.org/10.1242/jeb.198234

Seo, H., & Lee, D. (2007). Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *27*(31), 8366–8377. https://doi.org/10.1523/JNEUROSCI.2369-07.2007

Shadmehr, R., Huang, H. J., & Ahmed, A. A. (2016a). A Representation of Effort in Decision-Making and Motor Control. *Current Biology*, *26*(14), 1929–1934. https://doi.org/10.1016/j.cub.2016.05.065

Shadmehr, R., Huang, H. J., & Ahmed, A. A. (2016b). Effort, reward, and vigor in decision-making and motor control. *Current Biology : CB*, *26*(14), 1929–1934. https://doi.org/10.1016/j.cub.2016.05.065

Skvortsova, V., Palminteri, S., & Pessiglione, M. (2014). Learning to minimize efforts versus maximizing rewards: Computational principles and neural correlates. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *34*(47), 15621–15630. https://doi.org/10.1523/JNEUROSCI.1350-14.2014

Soltani, M., & Knight, R. T. (2000). Neural origins of the P300. *Critical Reviews in Neurobiology*, *14*(3–4), 199–224.

Summerside, E. M., Shadmehr, R., & Ahmed, A. A. (2018). Vigor of reaching movements: Reward discounts the cost of effort. *Journal of Neurophysiology*, *119*(6), 2347–2357. https://doi.org/10.1152/jn.00872.2017

Swick, D., & Turken, A. U. (2002). Dissociation between conflict detection and error monitoring in the human anterior cingulate cortex. *Proceedings of the National Academy of Sciences*, *99*(25), 16354–16359. https://doi.org/10.1073/pnas.252521499

Tanner, D., Morgan-Short, K., & Luck, S. J. (2015). How inappropriate high-pass filters can produce artifactual effects and incorrect conclusions in ERP studies of language and cognition. *Psychophysiology*, *52*(8), 997–1009. https://doi.org/10.1111/psyp.12437

Ullsperger, M., Danielmeier, C., & Jocham, G. (2014). Neurophysiology of performance monitoring and adaptive behavior. *Physiological Reviews*, *94*(1), 35–79. https://doi.org/10.1152/physrev.00041.2012

Vezoli, J., & Procyk, E. (2009). Frontal feedback-related potentials in nonhuman primates: Modulation during learning and under haloperidol. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *29*(50), 15675–15683. https://doi.org/10.1523/JNEUROSCI.4943-09.2009

Walsh, M. M., & Anderson, J. R. (2012). Learning from experience: Event-related potential correlates of reward processing, neural adaptation, and behavioral choice. *Neuroscience and Biobehavioral Reviews*, *36*(8), 1870–1884. https://doi.org/10.1016/j.neubiorev.2012.05.008

Walton, M. E., Bannerman, D. M., Alterescu, K., & Rushworth, M. F. S. (2003). Functional Specialization within Medial Frontal Cortex of the Anterior Cingulate for Evaluating Effort-Related Decisions. *Journal of Neuroscience*, *23*(16), 6475–6479. https://doi.org/10.1523/JNEUROSCI.23-16-06475.2003

Walton, M. E., Kennerley, S. W., Bannerman, D. M., Phillips, P. E. M., & Rushworth, M. F. S. (2006). Weighing up the benefits of work: Behavioral and neural analyses of effort-related decision making. *Neural Networks: The Official Journal of the International Neural Network Society*, *19*(8), 1302–1314. https://doi.org/10.1016/j.neunet.2006.03.005

Wang, J., Chen, J., Lei, Y., & Li, P. (2014). P300, not feedback error-related negativity, manifests the waiting cost of receiving reward information. *NeuroReport*, *25*(13), 1044–1048. https://doi.org/10.1097/WNR.0000000000000226

Wang, L., Zheng, J., & Meng, L. (2017). Effort provides its own reward: Endeavors reinforce subjective expectation and evaluation of task performance. *Experimental Brain Research*, *235*(4), 1107–1118. https://doi.org/10.1007/s00221-017-4873-z

Warren, C. M., Hyman, J. M., Seamans, J. K., & Holroyd, C. B. (2015). Feedback-related negativity observed in rodent anterior cingulate cortex. *Journal of Physiology-Paris*, *109*(1), 87–94. https://doi.org/10.1016/j.jphysparis.2014.08.008

Williams, Z. M., Bush, G., Rauch, S. L., Cosgrove, G. R., & Eskandar, E. N. (2004). Human anterior cingulate neurons and the integration of monetary reward with motor responses. *Nature Neuroscience*, *7*(12), 1370–1375. https://doi.org/10.1038/nn1354

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

Yeung, N., Holroyd, C. B., & Cohen, J. D. (2005). ERP Correlates of Feedback and Reward Processing in the Presence and Absence of Response Choice. *Cerebral Cortex*, *15*(5), 535–544. https://doi.org/10.1093/cercor/bhh153

Yeung, N., & Sanfey, A. G. (2004). Independent Coding of Reward Magnitude and Valence in the Human Brain. *Journal of Neuroscience*, *24*(28), 6258–6264. https://doi.org/10.1523/JNEUROSCI.4537-03.2004

Yoon, T., Jaleel, A., Ahmed, A. A., & Shadmehr, R. (2020). Saccade vigor and the subjective economic value of visual stimuli. *Journal of Neurophysiology*, *123*(6), 2161–2172. https://doi.org/10.1152/jn.00700.2019

Zentall, T. R. (2010). Justification of Effort by Humans and Pigeons: Cognitive Dissonance or Contrast? *Current Directions in Psychological Science*, *19*(5), 296–300. https://doi.org/10.1177/0963721410383381

Zink, C. F., Pagnoni, G., Martin-Skurski, M. E., Chappelow, J. C., & Berns, G. S. (2004). Human striatal responses to monetary reward depend on saliency. *Neuron*, *42*(3), 509–517. https://doi.org/10.1016/s0896-6273(04)00183-7

Chapter 6

General Discussion

This thesis presents a series of experiments testing whether fronto-striatal and dopaminergic neural circuits contribute to motor adaptation according to principles of biological reinforcement learning. Experimental tasks were designed to isolate neural and behavioral adaptive responses to feedback conveying information about reward, physical effort, and sensory error. Neural event-related potentials (ERPs) elicited by feedback were measured using scalp electroencephalography (EEG). Of particular interest was the feedback-related negativity/ reward positivity (FRN/RP), a midfrontal component of the scalp ERP that occurs approximately 200-350ms following task-related feedback. The FRN/RP is theorized to index activity in fronto-striatal circuits in response to dopaminergic signaling of reward-prediction error, a canonical computation in biological reinforcement learning (Glimcher, 2011; Holroyd & Coles, 2002; Sambrook & Goslin, 2015; Schultz, 2016; Walsh & Anderson, 2012). Indeed, measures of the FRN/RP correlate with neural responses to reward in the ventral striatum and medial frontal cortex, as well as with gray matter volume of the dopaminergic midbrain (Becker et al., 2014; Carlson et al., 2011, 2015).

Levodopa, a dopamine precursor, was administered to healthy volunteers to test whether dopamine release mediates neural feedback processing and behavioral adaptation. Levodopa has been shown to disrupt reward-based learning in healthy volunteers, and it is thought to do so by overstimulating dopamine release in the striatum (Cools et al., 2001, 2006; Frank et al., 2004; Hiebert et al., 2019; MacDonald et al., 2011; Vo et al., 2016). The following discussion first summarizes the primary findings reported throughout this thesis. Second, limitations of the current results are acknowledged. Third, the findings are interpreted in the context of the literature. Important questions and avenues for future research are discussed.

6.1 Summary of results

In the experiment presented in chapter 2, we sought to identify neural correlates of reward-prediction error during reward-based motor adaptation, and to dissociate them from neural processes related to sensory-error based learning. In a reward-based learning

task, adaptation of reach angle was induced through binary reinforcement feedback that indicated whether movements were successful or not. Cursor feedback indicating the position of the hand was withheld to preclude sensory error-based learning (Izawa & Shadmehr, 2011). In a separate visuomotor rotation task, cursor feedback of hand position was perturbed to induce sensory-error based learning. To preclude reward prediction error, the task was designed such that the rotations did not disrupt task performance or reward attainment. Binary reinforcement feedback elicited an FRN/RP ERP component, a relatively positive ERP deflection at the mid-frontal scalp in response to reward relative to the response elicited by non-reward. The FRN/RP component was larger for infrequent reinforcement outcomes, which is consistent with the notion that it encodes reward prediction error (Krigolson, 2018; Sambrook & Goslin, 2015; Walsh & Anderson, 2012). Visuomotor rotations of cursor feedback did not elicit an FRN/RP component. These findings suggest that the FRN/RP is a specific index of reward prediction error during motor adaptation, and that the process generating the FRN/RP does not occur in response to sensory error.

A centro-parietal ERP component called the P300 was elicited by both reinforcement feedback and sensory error feedback. In response to sensory error feedback, P300 amplitude was increased by larger visuomotor perturbations but not smaller ones. In response to reinforcement feedback, P300 amplitude was larger for reward compared to non-reward, and for infrequent outcomes regardless of valence. These results suggest that the process generating the P300 is generally sensitive to any feature of feedback that signals a requirement for adaptive changes in motor control.

In the experiment presents in chapter 3, we tested the hypothesis that reward-based motor adaptation is mediated by dopaminergic signaling of reinforcement outcomes. Participants ingested either levodopa or placebo in a repeated measures, double-blinded design before performing reward- and sensory error-based learning tasks similar to those described in chapter 2. We predicted that levodopa would impair learning in the reward-based motor adaptation task, as it has been shown to do in various cognitive reward-

based learning tasks (Cools et al., 2001, 2006; Frank et al., 2004; Hiebert et al., 2019; MacDonald et al., 2011; Vo et al., 2016). However, levodopa did not affect adaptation of reach angle in either the reward- or sensory error-based learning tasks used in the current study. Nor did levodopa affect neural processing of reinforcement outcomes in the medial-frontal cortex, as indexed by the FRN/RP. These findings suggest that reward-based motor adaptation and neural reward processing measured by the FRN/RP are not mediated by the same dopaminergic mechanisms that have previously been shown to be impaired by levodopa. Previous work has shown that people with Parkinson's disease have a reduced tendency to explore variable movements in response to decreased reward frequency (Pekny et al., 2015). We found no effect of levodopa on reward-dependent modulation of trial-by-trial movement variability, suggesting that deficits in motor exploration are a specific effect of PD, not a side effect of levodopa.

In the experiment presented in chapter 4, we tested the hypothesis that levodopa would impair savings and anterograde interference for force-field adaptation. Previous behavioral findings suggest that savings and interference may depend on reinforcement learning processes (Huang et al., 2011). People with PD show reduced savings and interference despite intact initial adaptation, suggesting a potential role for dopaminergic input to fronto-striatal circuits (Bédard & Sanes, 2011; Leow et al., 2012, 2013; Marinelli et al., 2009). However, we found no effects of levodopa on either adaptation, savings, or anterograde interference. These findings suggest that savings and interference depend on different mechanisms than those shown to be impaired by levodopa in cognitive reward-based learning tasks. The deficits observed in people with PD are therefore likely to be specific effects of the disease process as opposed to side effects of dopaminergic medication.

In the experiment presented in chapter 5, we tested the hypothesis that reinforcement signaling in the medial frontal cortex, indexed by the FRN/RP, is sensitive not only to reward outcomes, but also to effort as a cost. The neural generator of the FRN/RP has been consistently localized to the anterior cingulate cortex (Becker et al., 2014; Gehring &

Willoughby, 2002; Hauser, Iannaccone, Ball, et al., 2014; Hauser, Iannaccone, Stämpfli, et al., 2014; Santesso et al., 2008; E. H. Smith et al., 2015; Warren et al., 2015). During decision making, the anterior cingulate cortex encodes effort-discounted reward, with activity being positively related to reward and negatively related to effort (Croxson et al., 2009; Hauser et al., 2017; Kennerley et al., 2009; Klein-Flügge et al., 2016; Porter et al., 2019). However, we did not find evidence that the FRN/RP was modulated by effort as a cost or a loss during outcome evaluation. Participants were required to produce precise levels of muscle activity in order to receive rewards, and the level of physical effort required to obtain reward was determined probabilistically according to choices under uncertainty. After choices, feedback indicating the resulting physical effort requirement did not elicit an FRN/RP component, as would be expected if effort carried negative value or loss. Following completion of effort, binary reinforcement feedback indicated reward outcomes, and reward depended on accurate performance. The FRN/RP was not diminished by effort, as would be expected for a signal encoding the value of reward discounted by effort. Rather, the differential response to reward vs nonreward in midfrontal ERPs was increased by effort. Source decomposition suggests that the FRN/RP was only affected by reward, while a later sustained signal resembling the P300 was sensitive to reward only when effort was high.

6.2 Limitations

Chapter 2 limitations:

We found that the FRN/RP was elicited specifically by reinforcement outcomes but not sensory error induced by visuomotor rotations (VMR). Small VMR perturbations were used to avoid disrupting task performance which could result in reward prediction error. One potential criticism is that 1.5 degree VMRs may simply not have been sufficiently large to elicit an FRN/RP. However, sensory error feedback did elicit a reliable P300 component that was modulated in amplitude by VMR perturbations, demonstrating neural sensitivity to sensory error despite the small rotations. Both tasks resulted in highly reliable behavioral adaptation of comparable magnitude. Furthermore, we

replicated these findings in the results presented in chapter 4 using larger VMR perturbations, up to 4 degrees.

The current study adds to a growing number of experiments characterizing the properties of the FRN/RP in response to various experimental manipulations. However, the physiological underpinnings of the FRN/RP are still not well understood, and we did not demonstrate a direct relationship between the FRN/RP and behavior. Some avenues for further work are discussed in the next section. (see "The role of midfrontal reinforcement signals in motor adaptation: *future directions")*

Chapter 3 limitations:

Contrary to our hypothesis, we found no effect of levodopa on reward-based motor adaptation. Levodopa has been repeatedly shown to impair reward-based learning in cognitive tasks. Thus, we interpret this result as indicating a dissociation between the neural mechanisms for reward-based motor adaptation and reward-based learning involving choices between discrete stimuli. However these results would be more convincing if we had demonstrated behavioral effects of levodopa on a positive control task within the same sample.

Another possibility is that some component of reward-based motor learning was impaired but that other adaptive processes compensated for this deficit. For example, when implicit sensory-error based learning is impaired by cerebellar damage, an explicit aiming strategy can produce compensatory adaptation to VMR perturbations (Taylor et al., 2010). Adaptation in the current task is likely also achievable through a simple and verbalizable strategy (Holland et al., 2018). The current reward-based learning task was designed to be comparable with standard VMR paradigms, but it may not be particularly relevant to real-world motor control or learning. An important challenge for future work will be to design more demanding and ecologically valid paradigms that depend on reward-based learning mechanisms.

Chapter 4 limitations:

Contrary to our hypothesis, we found no effects of levodopa on savings or anterograde interference in force field adaptation. There was reliable evidence that levodopa did not affect measures of savings or learning with interference. However, even when combining the placebo and levodopa groups the evidence that savings occurred at all was fairly weak and the effect size was very small. Thus, the experimental paradigm may not have produced sufficient savings effects to detect effects of levodopa. However, we did observe robust adaptation effects and clear evidence against effects of levodopa on initial force field learning.

Chapter 5 limitations:

We found that effort increased the neural sensitivity to reward outcomes. However, reward outcomes did not affect choice behavior in the task. Thus, we were unable to determine whether the effect of effort on neural reward processing corresponded to any effects of effort on behavioral responses to reward. The task was designed primarily to elicit effort-based learning, as choices only affected the effort level, and reward probability was equated across the effort conditions. Future work may attempt to replicate these findings in a task where effort and reward outcomes both produce behavioral adaptation.

These findings are also limited by the poor spatial resolution of EEG. Source decomposition through ICA suggests that effort did not affect the FRN/RP responses to reward, but rather effort affected a distinct later ERP component, most likely the P300. Numerous lines of evidence point to the dorsal anterior cingulate as a generator for the FRN/RP, however we can only speculate on the neural sources of the later component at this time. Furthermore, EEG is only sensitive to local field potentials caused by synchronous responses of large neuronal populations. It is possible that neurons in the ACC have mixed tuning to effort, with effort increasing the firing rate of some neurons and decreasing the firing rate of others (Kennerley et al., 2009). If the direction of tuning

is balanced across the entire neural population, then scalp ERPs would not detect effects of effort even if it was strongly encoded in the neural source of the FNR/RP.

6.3 The role of midfrontal reinforcement signals in motor adaptation

The FRN/RP is a specific index of reward-based learning processes

Studies of motor adaptation have largely focused on sensory error-based learning. However, motor output can also be adapted and refined through a reward-based learning process (Cashaback et al., 2017, 2019; Izawa & Shadmehr, 2011; Therrien et al., 2016). In reward-based motor adaptation, variable movements are produced during repeated practice, and movements that result in success or reward are reinforced. In the results presented in chapters 2 and 3, we found that the FRN/RP was elicited specifically by reinforcement outcomes that produced reward-based motor adaptation, and not by sensory error. These findings seem to be at odds with some previous reports that the FRN/RP is elicited by perturbations in sensory error-based learning tasks including VMR and force field paradigms (Anguera et al., 2009; Reuter et al., 2018; Savoie et al., 2018; Torrecillos et al., 2014). To the extent that task success is desired and expected, VMR and force field perturbations may disrupt task performance and produce reward prediction error. We argue that our experiments dissociated sensory error and reward prediction error by using visuomotor rotations that did not displace the cursor outside of the reach target. Importantly, participants were explicitly instructed that reaches within the target would be considered successful and would result in reward attainment. Our results are consistent with findings that errors caused by sudden shifts in target location or by prism goggles only elicit an FRN/RP component when they result in failure to reach a visual target (Aziz et al., 2020; Krigolson et al., 2008; Krigolson & Holroyd, 2007).

It is also important to consider that some discrepancies in the literature may be caused by entirely different ERP components being conflated as the FRN/RP. Motor adaptation studies have typically identified midfrontal ERP negativities as FRN/RP components, as

these responses resemble the classic FRN elicited by feedback indicating non-reward or error (Anguera et al., 2009; Reuter et al., 2018; Savoie et al., 2018; Torrecillos et al., 2014). However, almost any salient stimulus causes an N200 ERP component that is indistinguishable from the classic FRN (Holroyd et al., 2008; Sokhadze et al., 2017). Recent findings suggest that rewarding and unfavorable outcomes both elicit an N200 component (Baker & Holroyd, 2011; Holroyd et al., 2008). There is compelling evidence to suggest that the variance measured by the FRN/RP is primarily attributable to a positive ERP deflection in response to rewarding outcomes that cancels the N200 by superposition (Baker & Holroyd, 2011; Becker et al., 2014; Carlson et al., 2011; Holroyd et al., 2008; Proudfit, 2015; Sambrook & Goslin, 2016). The N200 is typically larger in response to stimuli that elicit increased cognitive control or response conflict (Baker & Holroyd, 2011; Cavanagh & Frank, 2014; Folstein & Petten, 2008; Iannaccone et al., 2015; Ridderinkhof et al., 2004; Yeung et al., 2004). Perturbations during adaptation tasks likely result in an immediate allocation of neural resources to restore control. Furthermore, perturbations during ongoing movement may induce response conflict as the current motor plan may be suppressed to allow for a corrective movement. Thus, it is important to note that motor error processing may elicit an N200 response related to regulation of control but not motor adaptation per se, which may commonly be conflated with the FNR/RP.

Functional significance of the FRN/RP in reward-based motor adaptation

The vast majority of studies measuring the FRN/RP have used cognitive learning tasks such as time-estimation and gambling style tasks (Cohen et al., 2007; Cohen & Ranganath, 2007; Gehring & Willoughby, 2002; Holroyd & Krigolson, 2007; Sambrook & Goslin, 2015; Glazer et al., 2018). In these contexts, the FRN/RP is thought to measure neural activity in the dorsal anterior cingulate cortex (dACC) indexing reward-prediction error, a canonical teaching signal for reinforcement learning (Holroyd & Coles, 2002; Sambrook & Goslin, 2015; Walsh & Anderson, 2012). Reward-based motor adaptation has also been modeled using reinforcement learning algorithms that compute reward-

prediction error (Dhawale et al., 2019; Izawa & Shadmehr, 2011). A natural question, which this thesis begins to address, is whether biological reinforcement learning mechanisms that produce learning for discrete choices also underlie the adaptation and refinement of motor control. It may be that analogous reinforcement learning processes occur in distinct, hierarchically organized circuits for decision making and motor control. For example, the selection and execution of actions may be adapted in parallel fronto-striatal loop circuits involving prefrontal and motor cortices, respectively. It is also possible that the selection and execution of actions both emerge from the same neural representation of movement (Cisek, 2006; Cisek & Kalaska, 2005, 2010). Reinforcement learning for decision making and motor adaptation might both occur on a unified representation of movement without structural distinction.

The notion that the FRN/RP measures reinforcement learning signals in the dACC for decision making is consistent with data from single-cell recordings. Neuronal firing rates in the dACC encode reward prediction errors for the outcomes of decision (Amiez et al., 2005; Hill et al., 2016; Kennerley et al., 2009; Matsumoto et al., 2007; H. Seo & Lee, 2007; Warren et al., 2015), and lesions or inactivation of the ACC impair the ability to adapt choices according to reward outcomes (Camille et al., 2011; Kennerley et al., 2006; Shima & Tanji, 1998; Walton et al., 2007; Williams et al., 2004). The dACC may receive rich information about reinforcement outcomes through extensive connections to other prefrontal subregions, the striatum, and the dopaminergic midbrain (Kolling et al., 2016). The dACC also forms dense connections to motor cortex as well as direct corticospinal outputs (Paus, 2001; Strick et al., 1998). Stimulation of motor fields within the dACC can elicit movements of the upper limb (Luppino et al., 1991; Paus, 2001; Showers, 1959). Thus, this region seems well-positioned to coordinate reward-based motor learning through extensive connections to reward circuits and functional proximity to motor output.

The FRN/RP does not reflect physical effort

During value-based decision making, neural activity in the ACC tends to encode effort in opposition to reward (Croxson et al., 2009; Hauser et al., 2017; Kennerley et al., 2009; Klein-Flügge et al., 2016; Porter et al., 2019; Walton et al., 2006). It is thought that this activity signals subjective value or utility, a common neural currency for choice that accounts for both the costs and benefits of actions. In almost all studies of decision making involving effort and reward, the result of each potential choice is cued explicitly before decisions. In the experiment presented in chapter 5, participants instead made choices under uncertainty that resulted in varying physical effort requirements and reward outcomes. We asked whether the FRN/RP was sensitive to physical effort as a cost during outcome processing.

Contrary to our expectations, feedback indicating the effort requirements resulting from participants' choices did not elicit FRN/RP signals. Nor did we find evidence that effort diminished the FRN/RP responses to reward feedback, as would be expected if effort acted as a cost that devalues reward. These results are inconsistent with two previous fMRI studies that reported effort-prediction error signals in the ACC while participants adapted their behavior to minimize effort (Hauser et al., 2017; Skvortsova et al., 2014). However, in Skvortsova et al., (2014) participants were specifically instructed to avoid unnecessary effort. This instruction confounds effort feedback with error or reward signals related to meeting the goals of the task, which are likely to be observed in the ACC. In Hauser et al., (2017), participants squeezed a handgrip and received varying rewards if their grip force met an unknown and changing threshold. Effort-prediction error was supposed to occur in response to binary feedback indicating whether or not the unknown force threshold was met on a given trial. In this case, effort-prediction error could also be interpreted as reward prediction error related to meeting the goals of the task, or as motor error. In the present study participants were not instructed to avoid effort, and effort was orthogonal to task success and reward attainment. Our results suggest that outcome processing in the ACC does not reflect effort under these

conditions. It is certainly plausible that these discrepancies are due to EEG and fMRI measuring different aspects of neural activity. However EEG measures of the FRN/RP have been shown to correlate with BOLD responses to reinforcement in the ACC (Becker et al., 2014; Carlson et al., 2011).

Future directions:

Overall, our results suggest that the FRN/RP is specific index of reinforcement processing during motor adaptation, and that it is dissociable from sensory error- and effort-based learning processes. However, it is still not clear whether the process generating the FRN/RP is a necessary component of reward-based motor learning. It would be useful for future work to identify causal manipulations of neurophysiology that result in specific effects on behavior and the FRN/RP. Chapter 3 describes our attempt to do so using levodopa, which however did not produce any effects on the FRN/RP or behavior. Other pharmacological techniques, studies of humans with focal lesions of the dACC, and the use of animal models might prove more successful. Furthermore, larger datasets may be able to identify correlational relationships between the FRN/RP and behavior, either through variation in single trial measures within participants, or through individual differences between participants.

Although the FRN/RP was not directly sensitive to effort as a cost, it may be a measure of reward processing that influences the motivation to produce effort. Some theories posit that reward signals in the dACC do not guide the selection or implementation of individual actions, but rather the regulation of effortful control applied to the task at hand according to its expected utility (Botvinick & Braver, 2015; Holroyd & Umemoto, 2016; Shenhav et al., 2016). Reward contingencies can improve the control of movements through upregulation of feedback gains or limb stiffness (Carroll et al., 2019; Codol et al., 2020; Manohar et al., 2019). Reward can also increase the speed and vigor of movements, which is thought to reflect a tradeoff between the effort and incentives associated with moving quickly (Shadmehr et al., 2016a; Summerside et al., 2018; Yoon et al., 2020). It would be interesting to test whether the FRN/RP reflects tracking of reward outcomes for

the purpose of allocating effortful control according to the demands and utility of the task at hand. For example, increased stiffness may improve feedforward control and stability in the face of external perturbations, while increased feedback gains may facilitate movements towards unpredictable moving targets. In foraging contexts, normative theories suggest that reward rate should modulate vigor (Yoon et al., 2018).

Midfrontal theta oscillations generated by the dACC have been extensively implicated in cognitive control (Cavanagh & Frank, 2014; Holroyd & Umemoto, 2016). Recent work suggests midfrontal theta-band activity may also regulate visuomotor feedback control (Watanabe et al., 2021). Above, we speculated that midfrontal ERP negativities in response to motor perturbations may reflect the N200. The N200 is thought to measure phase-locked, event-related theta synchronization that facilitates neural control (Cavanagh & Frank, 2014), and indeed the midfrontal ERP response to visuomotor perturbation reflects theta-band activity (Savoie et al., 2018). Midfrontal regions may regulate control through long-range connections to other task-relevant brain regions including motor areas, evidenced by interregional phase synchronization in theta-band activity (Asanowicz et al., 2021). Beta-band desynchronization in sensory-motor regions is associated with improvements in motor performance due to reward cues and error feedback, and may be moderated by midfrontal theta-band activity (Savoie et al., 2019; Watanabe et al., 2021). Future work might test whether reward processing indexed by the FRN/RP is associated with the regulation of control effort mediated by midfrontal theta, particularly in response to motor errors.

6.4 The P300 ERP component is modulated by error, reinforcement, and effort

P300 effects during motor adaptation

Both reinforcement feedback and sensory error feedback elicited P300 ERP component responses, centro-parietal positive ERP deflections that peaking between 300-500ms after feedback. The amplitude of P300 responses to sensory error feedback was increased by

visuomotor rotation, and this effect depended on the size of the perturbations. This finding is consistent with reports that visuomotor perturbations caused by prism goggles increase P300 amplitude (Aziz et al., 2020; MacLean et al., 2015). During reward-based motor adaptation, the amplitude of P300 responses to reinforcement feedback was larger for infrequent outcomes relative to frequent outcomes regardless of valence. Similar effects of outcome frequency on P300 amplitude in cognitive reward based learning tasks are common in the literature (Hajcak et al., 2005, 2007; Leng & Zhou, 2010; Pfabigan et al., 2011; Y. Wu & Zhou, 2009; Zhou et al., 2010). If participants learned to expect the frequent outcomes, infrequent feedback should elicit larger prediction errors and drive a larger adaptive response. We found that P300 amplitude was larger in response to reward outcomes than non-reward outcomes, however this effect may have been due to spatiotemporal overlap with the FRN/RP. The effects of error and surprise suggest that the P300 response is increased by features of feedback that elicit adaptation across both reward- and sensory error-based learning processes.

The P300 is thought to arise from diffuse cortical sources including parietal, frontal, and motor areas (Polich, 2007; Soltani & Knight, 2000). A prominent theory states that the P300 is generated by cortical responses to phasic norepinephrine release from the locus coeruleus (Aston-Jones & Cohen, 2005; de Rover et al., 2015; De Taeye et al., 2014; Murphy et al., 2011; Nieuwenhuis et al., 2005; Polich, 2007; Swick et al., 1994; Vazey et al., 2018). Phasic norepinephrine release in response to task-related outcomes is theorized to facilitate behavioral adaptation, possibly by interrupting the current neural activity state to allow rapid reorganization (Aston-Jones & Cohen, 2005; Bouret & Sara, 2005; Dayan & Yu, 2006). Norepinephrine release in response to salient or unexpected events may regulate the rate of learning according to uncertainty about the environment (Sales et al., 2019). Norepinephrine has been shown to modulate synaptic long term potentiation and long term depression through actions on beta-adrenoreceptors (Gibbs & Summers, 2002; Hagena et al., 2016; Lemon et al., 2009; Pettigrew, 1982; Pettigrew & Kasamatsu, 1978; Salgado et al., 2012; Stanton & Sarvey, 1985). The locus coeruleus innervates nearly all cortical areas, and so it is possible that phasic norepinephrine release serves as a highly

general mechanism that facilitates behavioral adaptation across different tasks (Poe et al., 2020). This is consistent with a role in both sensory error- and reward-based motor adaptation. Locus coeruleus neurons display highly plastic responses to reward and reward-predictive stimuli during learning (G. Aston-Jones et al., 1997; Bouret & Sara, 2004; Sara & Segal, 1991). Recent evidence suggests that norepinephrine release in response to force field perturbations modulates the rate of motor adaptation according to surprise and uncertainty (Yokoi & Weiler, 2021).

Effects of physical effort on P300 responses to reward feedback

In the experiment presented in chapter 5, we found that P300 amplitude was larger in response to feedback indicating reward compared to non-reward only when the preceding effort was high. This finding suggests that physical effort may increase the salience of subsequent reward outcomes, possibly through increased norepinephrine release caused by effort-related arousal. We speculate that this effect may underlie paradoxical findings in which humans and other animals tend to overvalue rewards and conditioned reinforcers that are obtained under conditions of high effort (Clement et al., 2000; Inzlicht et al., 2018; Lydall et al., 2010; Zentall, 2010). Norepinephrine is also implicated in motivating and energizing effortful behavior (Borderies et al., 2020; Bouret et al., 2012; Varazzani et al., 2015; Zénon et al., 2014). In a high-effort context, reward-related norepinephrine release may be important to motivate continued performance of the effortful behavior.

Future directions:

Future work might test whether P300 responses to reward and sensory error feedback correlate with increases in pupil area, an index of norepinephrine release. Adrenoreceptor signaling could also be manipulated pharmacologically using propranolol or clonidine, and both of these drugs have been shown to affect P300 amplitude in some tasks (de Rover et al., 2015; Nieuwenhuis et al., 2005). It would be useful to test whether these

drugs cause effects on P300 responses to feedback during motor adaptation that correspond with changes of learning rate.

We found that effort increased the sensitivity of P300 responses to reward feedback. Future work might test whether this effect corresponds to positive effects of effort on the rate of reward-based learning. This could be accomplished using a reward-based learning task in which physical effort is manipulated independent of reward outcomes. Future experiments might also test whether effort-related increases in reward responsivity relate to the motivation to persist in effortful behavior.

6.5 The role of dopamine in motor adaptation

Contrary to our expectations, we observed no effects of levodopa on reward-based motor adaptation, reward-dependent modulation of movement variability, savings, or anterograde interference. Levodopa has been shown repeatedly to impair learning in tasks that involve associating visual stimuli with discrete responses through reinforcement feedback (Cools et al., 2001, 2007; Feigin et al., 2003; Frank et al., 2004; Graef et al., 2010; Hiebert et al., 2014; Jahanshahi et al., 2010; Kwak et al., 2010; MacDonald et al., 2011; Swainson et al., 2000; Torta et al., 2009; Vo et al., 2016, 2018). Our findings suggest a dissociation between dopaminergic mechanisms for motor adaptation and stimulus-response learning. This dissociation is not trivial as dopaminergic signaling of reward-prediction error is thought to be a ubiquitous and canonical mechanism for reward-based learning (Bayer & Glimcher, 2005; Glimcher, 2011; Schultz, 2016; Schultz et al., 1997).

In people with Parkinson's disease (PD), neuroimaging studies suggest that levodopa impairs stimulus-response learning by disrupting neural responses to reward feedback in the ventral striatum (Cools et al., 2007; Hiebert et al., 2019; MacDonald et al., 2011). Early-stage PD primarily causes dopamine depletion in the dorsal striatum while the ventral striatum is relatively spared. According to the "dopamine overdose" hypothesis, the dosages of levodopa necessary to restore motor functions mediated by the dorsal striatum result in overstimulation of dopamine release in the ventral striatum.

"Dopamine overdose" in the ventral striatum is thought to also explain levodopa-induced learning impairments in healthy volunteers (Vo et al., 2016, 2017, 2018).

Monkeys with ventral striatum lesions are impaired in learning to choose stimuli associated with reward outcomes, but are still able to learn action-reward associations (Rothenhoefer et al., 2017). Action-based reinforcement learning, such as in the current task, may instead rely on the dorsal striatum (E. Lee et al., 2015; N. F. Parker et al., 2016; M. Seo et al., 2012). Sidarta et al. (2016) found that reward-based motor learning was associated with increased functional connectivity between the putamen and ventromedial prefrontal cortex using a task similar to that of the current study. Currently, there is no evidence that levodopa causes "dopamine overdose" effects for behaviors mediated by the dorsal striatum. During tasks that measure response selection for stimulus-response associations that have already been learned, levodopa remediates behavioral performance and associated neural activity in the dorsal striatum in people with PD (Hiebert et al., 2019). However, levodopa does not impair response selection in healthy young participants despite impairing learning, suggesting that the dorsal striatum may not be vulnerable to "dopamine overdose" (Vo et al., 2017).

It is also possible that reward-based motor adaptation is mediated by dopaminergic projections directly to motor cortex, which may not be susceptible to "dopamine overdose". It has been shown that dopaminergic innervation of the motor cortex is essential for some forms of motor learning in rodents (Beeler et al., 2010; Hosp et al., 2011; Luft & Schwarz, 2009; Molina-Luna et al., 2009). Finally, reward-based motor learning may simply not depend on dopamine at all. Recent studies have found that effects of reward on motor adaptation are not influenced by either dopaminergic genetic polymorphisms or haloperidol, a dopamine receptor antagonist (Holland et al., 2019; Quattrocchi et al., 2018). However, early-stage PD has been shown to impair trial-by-trial responses to feedback in reward-based motor adaptation as well as savings and anterograde interference (Bédard & Sanes, 2011; Leow et al., 2012, 2013; Marinelli et al., 2009; Pekny et al., 2015). Because our results rule out dopamine overdose effects caused

by levodopa, it is likely that the effects of PD on reward-based motor adaptation, savings, and interference are caused by dysfunction of learning mechanisms mediated by the dorsal striatum.

Future directions:

Future studies might image neural activity using fMRI during reward-based motor learning. The experimental paradigms reported in chapters 2 and 3 could be adapted to use a joystick in a scanner. One hypothesis is that reward-prediction errors elicit specific activations of the dorsal and ventral striatum during motor learning and stimulus-response learning, respectively. A causal role of dopamine function in the dorsal striatum could be assessed by scanning people with PD on and off levodopa medication in a repeated measures design.

Another hypothesis is that dopaminergic reward prediction errors drive motor learning by improving movement representations in the motor cortex. This hypothesis could be tested by having participants learn to produce a set of distinct movements through binary reinforcement feedback. Reward prediction error could be experimentally induced through probabilistic reinforcement outcomes. The fidelity of movement representations could be assessed through multivariate distances between the cortical activity patterns associated with different movements. This hypothesis predicts that reward prediction error signals in the dorsal striatum would be positively related to subsequent increases in pattern distance between cortical representations of distinct movements during motor preparation and execution.

FMRI could also be used to test whether reinforcement learning processes underlie savings and anterograde interference. It is hypothesized that when sensory error-based learning processes restore accurate performance, the successful movements are reinforced (Huang et al., 2011). When a subsequent perturbation is encountered, the previously reinforced movements are thought to be recalled, resulting in either savings or interference. This hypothesis leads to the prediction that fronto-striatal reward circuits

would be activated upon attainment of successful performance in the task in the late stages of adaptation. Furthermore, this hypothesis implies that the expression of savings and interference should be associated with activity in fronto-striatal circuits. It predicts that previous adaptation would cause an increase in fronto-striatal activation immediately upon the onset of a second perturbation, and that this activity would be correlated to the expression of savings or interference.

6.6 References

Amiez, C., Joseph, J.-P., & Procyk, E. (2005). Anterior cingulate error-related activity is modulated by predicted reward. *The European Journal of Neuroscience*, *21*(12), 3447–3452. https://doi.org/10.1111/j.1460-9568.2005.04170.x

Anguera, J. A., Seidler, R. D., & Gehring, W. J. (2009). Changes in Performance Monitoring During Sensorimotor Adaptation. *Journal of Neurophysiology*, *102*(3), 1868–1879. https://doi.org/10.1152/jn.00063.2009

Asanowicz, D., Panek, B., & Kotlewska, I. (2021). Selection for Action: The Medial Frontal Cortex Is an Executive Hub for Stimulus and Response Selection. *Journal of Cognitive Neuroscience*, 1-28.

Aston-Jones, G., Rajkowski, J., & Kubiak, P. (1997). Conditioned responses of monkey locus coeruleus neurons anticipate acquisition of discriminative behavior in a vigilance task. *Neuroscience*, *80*(3), 697–715. https://doi.org/10.1016/s0306-4522(97)00060-2

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annual Review of Neuroscience*, *28*, 403–450. https://doi.org/10.1146/annurev.neuro.28.061604.135709

Aziz, J. R., MacLean, S. J., Krigolson, O. E., & Eskes, G. A. (2020). Visual Feedback Modulates Aftereffects and Electrophysiological Markers of Prism Adaptation. *Frontiers in Human Neuroscience*, *14*. https://doi.org/10.3389/fnhum.2020.00138

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

Bayer, H. M., & Glimcher, P. W. (2005). Midbrain Dopamine Neurons Encode a Quantitative Reward Prediction Error Signal. *Neuron*, *47*(1), 129–141. https://doi.org/10.1016/j.neuron.2005.05.020

Becker, M. P. I., Nitsch, A. M., Miltner, W. H. R., & Straube, T. (2014). A single-trial estimation of the feedback-related negativity and its relation to BOLD responses in a time-estimation task. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *34*(8), 3005–3012. https://doi.org/10.1523/JNEUROSCI.3684-13.2014

Bédard, P., & Sanes, J. N. (2011). Basal ganglia-dependent processes in recalling learned visual-motor adaptations. *Experimental Brain Research*, *209*(3), 385–393. https://doi.org/10.1007/s00221-011-2561-y

Beeler, J. A., Cao, Z. F. H., Kheirbek, M. A., Ding, Y., Koranda, J., Murakami, M., Kang, U. J., & Zhuang, X. (2010). Dopamine-dependent motor learning: Insight into levodopa's long-duration response. *Annals of Neurology*, *67*(5), 639–647. https://doi.org/10.1002/ana.21947

Borderies, N., Bornert, P., Gilardeau, S., & Bouret, S. (2020). Pharmacological evidence for the implication of noradrenaline in effort. *PLOS Biology*, *18*(10), e3000793. https://doi.org/10.1371/journal.pbio.3000793

Botvinick, M. M., & Braver, T. (2015). Motivation and cognitive control: From behavior to neural mechanism. *Annual Review of Psychology*, *66*, 83–113. https://doi.org/10.1146/annurev-psych-010814-015044

Bouret, S., Richmond, B. J., & Ravel, S. (2012). Complementary neural correlates of motivation in dopaminergic and noradrenergic neurons of monkeys. *Frontiers in Behavioral Neuroscience*, *6*. https://doi.org/10.3389/fnbeh.2012.00040

Bouret, S., & Sara, S. J. (2004). Reward expectation, orientation of attention and locus coeruleus-medial frontal cortex interplay during learning. *The European Journal of Neuroscience*, *20*(3), 791–802. https://doi.org/10.1111/j.1460-9568.2004.03526.x

Bouret, S., & Sara, S. J. (2005). Network reset: A simplified overarching theory of locus coeruleus noradrenaline function. *Trends in Neurosciences*, *28*(11), 574–582. https://doi.org/10.1016/j.tins.2005.09.002

Camille, N., Tsuchida, A., & Fellows, L. K. (2011). Double Dissociation of Stimulus-Value and Action-Value Learning in Humans with Orbitofrontal or Anterior Cingulate Cortex Damage. *Journal of Neuroscience*, *31*(42), 15048–15052. https://doi.org/10.1523/JNEUROSCI.3164-11.2011

Carlson, J. M., Foti, D., Harmon-Jones, E., & Proudfit, G. H. (2015). Midbrain volume predicts fMRI and ERP measures of reward reactivity. *Brain Structure and Function*, *220*(3), 1861–1866. https://doi.org/10.1007/s00429-014-0725-9

Carlson, J. M., Foti, D., Mujica-Parodi, L. R., Harmon-Jones, E., & Hajcak, G. (2011). Ventral striatal and medial prefrontal BOLD activation is correlated with reward-related electrocortical activity: A combined ERP and fMRI study. *NeuroImage*, *57*(4), 1608–1616. https://doi.org/10.1016/j.neuroimage.2011.05.037

Carroll, T. J., McNamee, D., Ingram, J. N., & Wolpert, D. M. (2019). Rapid Visuomotor Responses Reflect Value-Based Decisions. *Journal of Neuroscience*, *39*(20), 3906–3920. https://doi.org/10.1523/JNEUROSCI.1934-18.2019

Cashaback, J. G. A., Lao, C. K., Palidis, D. J., Coltman, S. K., McGregor, H. R., & Gribble, P. L. (2019). The gradient of the reinforcement landscape influences sensorimotor learning. *PLOS Computational Biology*, *15*(3), e1006839. https://doi.org/10.1371/journal.pcbi.1006839

Cashaback, J. G. A., McGregor, H. R., Mohatarem, A., & Gribble, P. L. (2017). Dissociating error-based and reinforcement-based loss functions during sensorimotor learning. *PLoS Computational Biology*, *13*(7), e1005623. https://doi.org/10.1371/journal.pcbi.1005623

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

Cisek, P. (2006). Integrated Neural Processes for Defining Potential Actions and Deciding between Them: A Computational Model. *Journal of Neuroscience*, *26*(38), 9761–9770. https://doi.org/10.1523/JNEUROSCI.5605-05.2006

Cisek, P., & Kalaska, J. F. (2005). Neural Correlates of Reaching Decisions in Dorsal Premotor Cortex: Specification of Multiple Direction Choices and Final Selection of Action. *Neuron*, *45*(5), 801–814. https://doi.org/10.1016/j.neuron.2005.01.027

Cisek, P., & Kalaska, J. F. (2010). Neural Mechanisms for Interacting with a World Full of Action Choices. *Annual Review of Neuroscience*, *33*(1), 269–298. https://doi.org/10.1146/annurev.neuro.051508.135409

Clement, T. S., Feltus, J. R., Kaiser, D. H., & Zentall, T. R. (2000). "work ethic" in pigeons: Reward value is directly related to the effort or time required to obtain the reward. *Psychonomic Bulletin & Review*, *7*(1), 100–106. https://doi.org/10.3758/BF03210727

Codol, O., Holland, P. J., Manohar, S. G., & Galea, J. M. (2020). Reward-Based Improvements in Motor Control Are Driven by Multiple Error-Reducing Mechanisms. *Journal of Neuroscience*, *40*(18), 3604–3620. https://doi.org/10.1523/JNEUROSCI.2646-19.2020

Cohen, M. X., Elger, C. E., & Ranganath, C. (2007). Reward expectation modulates feedback-related negativity and EEG spectra. *NeuroImage*, *35*(2), 968–978. https://doi.org/10.1016/j.neuroimage.2006.11.056

Cohen, M. X., & Ranganath, C. (2007). Reinforcement Learning Signals Predict Future Decisions. *Journal of Neuroscience*, *27*(2), 371–378. https://doi.org/10.1523/JNEUROSCI.4421-06.2007

Cools, R., Altamirano, L., & D'Esposito, M. (2006). Reversal learning in Parkinson's disease depends on medication status and outcome valence. *Neuropsychologia*, *44*(10), 1663–1673. https://doi.org/10.1016/j.neuropsychologia.2006.03.030

Cools, R., Barker, R. A., Sahakian, B. J., & Robbins, T. W. (2001). Enhanced or impaired cognitive function in Parkinson's disease as a function of dopaminergic medication and task demands. *Cerebral Cortex (New York, N.Y.: 1991)*, *11*(12), 1136–1143. https://doi.org/10.1093/cercor/11.12.1136

Cools, R., Lewis, S. J. G., Clark, L., Barker, R. A., & Robbins, T. W. (2007). L-DOPA disrupts activity in the nucleus accumbens during reversal learning in Parkinson's disease. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, *32*(1), 180–189. https://doi.org/10.1038/sj.npp.1301153

Croxson, P. L., Walton, M. E., O'Reilly, J. X., Behrens, T. E. J., & Rushworth, M. F. S. (2009). Effort-Based Cost–Benefit Valuation and the Human Brain. *Journal of Neuroscience*, *29*(14), 4531–4541. https://doi.org/10.1523/JNEUROSCI.4515-08.2009

Dayan, P., & Yu, A. J. (2006). Phasic norepinephrine: A neural interrupt signal for unexpected events. *Network: Computation in Neural Systems*, *17*(4), 335–350. https://doi.org/10.1080/09548980601004024

de Rover, M., Brown, S. B. R. E., Band, G. P., Giltay, E. J., van Noorden, M. S., van der Wee, N. J. A., & Nieuwenhuis, S. (2015). Beta receptor-mediated modulation of the oddball P3 but not error-related ERP components in humans. *Psychopharmacology*, *232*(17), 3161–3172. https://doi.org/10.1007/s00213-015-3966-2

De Taeye, L., Vonck, K., van Bochove, M., Boon, P., Van Roost, D., Mollet, L., Meurs, A., De Herdt, V., Carrette, E., Dauwe, I., Gadeyne, S., van Mierlo, P., Verguts, T., & Raedt, R. (2014). The P3 Event-Related Potential is a Biomarker for the Efficacy of Vagus Nerve Stimulation in Patients with Epilepsy. *Neurotherapeutics*, *11*(3), 612–622. https://doi.org/10.1007/s13311-014-0272-3

Dhawale, A. K., Miyamoto, Y. R., Smith, M. A., & Ölveczky, B. P. (2019). Adaptive Regulation of Motor Variability. *Current Biology*, *29*(21), 3551-3562.e7. https://doi.org/10.1016/j.cub.2019.08.052

Feigin, A., Ghilardi, M. F., Carbon, M., Edwards, C., Fukuda, M., Dhawan, V., Margouleff, C., Ghez, C., & Eidelberg, D. (2003). Effects of levodopa on motor sequence learning in Parkinson's disease. *Neurology*, *60*(11), 1744–1749. https://doi.org/10.1212/01.wnl.0000072263.03608.42

Folstein, J. R., & Petten, C. V. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology*, *45*(1), 152–170. https://doi.org/10.1111/j.1469-8986.2007.00602.x

Frank, M. J., Seeberger, L. C., & O'reilly, R. C. (2004). By carrot or by stick: Cognitive reinforcement learning in parkinsonism. *Science (New York, N.Y.)*, *306*(5703), 1940–1943. https://doi.org/10.1126/science.1102941

Gehring, W. J., & Willoughby, A. R. (2002). The Medial Frontal Cortex and the Rapid Processing of Monetary Gains and Losses. *Science*, *295*(5563), 2279–2282. https://doi.org/10.1126/science.1066893

Gibbs, M. E., & Summers, R. J. (2002). Role of adrenoceptor subtypes in memory consolidation. *Progress in Neurobiology*, *67*(5), 345–391. https://doi.org/10.1016/S0301-0082(02)00023-0

Glazer, J. E., Kelley, N. J., Pornpattananangkul, N., Mittal, V. A., & Nusslock, R. (2018). Beyond the FRN: Broadening the time-course of EEG and ERP components implicated in reward processing. *International Journal of Psychophysiology*, *132*, 184–202. https://doi.org/10.1016/j.ijpsycho.2018.02.002

Glimcher, P. W. (2011). Understanding dopamine and reinforcement learning: The dopamine reward prediction error hypothesis. *Proceedings of the National Academy of Sciences*, *108*(Supplement 3), 15647–15654. https://doi.org/10.1073/pnas.1014269108

Graef, S., Biele, G., Krugel, L. K., Marzinzik, F., Wahl, M., Wotka, J., Klostermann, F., & Heekeren, H. R. (2010). Differential influence of levodopa on reward-based learning in Parkinson's disease. *Frontiers in Human Neuroscience*, *4*, 169. https://doi.org/10.3389/fnhum.2010.00169

Hagena, H., Hansen, N., & Manahan-Vaughan, D. (2016). β-Adrenergic Control of Hippocampal Function: Subserving the Choreography of Synaptic Information Storage and Memory. *Cerebral Cortex (New York, N.Y.: 1991)*, *26*(4), 1349–1364. https://doi.org/10.1093/cercor/bhv330

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

Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2007). It's worse than you thought: The feedback negativity and violations of reward prediction in gambling tasks. *Psychophysiology*, *44*(6), 905–912. https://doi.org/10.1111/j.1469-8986.2007.00567.x

Hauser, T. U., Eldar, E., & Dolan, R. J. (2017). Separate mesocortical and mesolimbic pathways encode effort and reward learning signals. *Proceedings of the National Academy of Sciences*, *114*(35), E7395–E7404. https://doi.org/10.1073/pnas.1705643114

Hauser, T. U., Iannaccone, R., Ball, J., Mathys, C., Brandeis, D., Walitza, S., & Brem, S. (2014). Role of the medial prefrontal cortex in impaired decision making in juvenile attention-deficit/hyperactivity disorder. *JAMA Psychiatry*, *71*(10), 1165–1173. https://doi.org/10.1001/jamapsychiatry.2014.1093

Hauser, T. U., Iannaccone, R., Stämpfli, P., Drechsler, R., Brandeis, D., Walitza, S., & Brem, S. (2014). The feedback-related negativity (FRN) revisited: New insights into the localization, meaning and network organization. *NeuroImage*, *84*, 159–168. https://doi.org/10.1016/j.neuroimage.2013.08.028

Hiebert, N. M., Owen, A. M., Ganjavi, H., Mendonça, D., Jenkins, M. E., Seergobin, K. N., & MacDonald, P. A. (2019). Dorsal striatum does not mediate feedback-based, stimulus-response learning: An event-related fMRI study in patients with Parkinson's disease tested on and off dopaminergic therapy. *NeuroImage*, *185*, 455–470. https://doi.org/10.1016/j.neuroimage.2018.10.045

Hiebert, N. M., Seergobin, K. N., Vo, A., Ganjavi, H., & MacDonald, P. A. (2014). Dopaminergic therapy affects learning and impulsivity in Parkinson's disease. *Annals of Clinical and Translational Neurology*, *1*(10), 833–843. https://doi.org/10.1002/acn3.128

Hill, M. R., Boorman, E. D., & Fried, I. (2016). Observational learning computations in neurons of the human anterior cingulate cortex. *Nature Communications*, *7*(1), 12722. https://doi.org/10.1038/ncomms12722

Holland, P., Codol, O., & Galea, J. M. (2018). Contribution of explicit processes to reinforcement-based motor learning. *Journal of Neurophysiology*, *119*(6), 2241–2255. https://doi.org/10.1152/jn.00901.2017

Holland, P., Codol, O., Oxley, E., Taylor, M., Hamshere, E., Joseph, S., Huffer, L., & Galea, J. M. (2019). Domain-Specific Working Memory, But Not Dopamine-Related Genetic Variability, Shapes Reward-Based Motor Learning. *The Journal of Neuroscience*, *39*(47), 9383–9396. https://doi.org/10.1523/JNEUROSCI.0583-19.2019

Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*(4), 679–709. https://doi.org/10.1037/0033-295X.109.4.679

Holroyd, C. B., & Krigolson, O. E. (2007). Reward prediction error signals associated with a modified time estimation task. *Psychophysiology*, *44*(6), 913–917. https://doi.org/10.1111/j.1469-8986.2007.00561.x

Holroyd, C. B., Pakzad-Vaezi, K. L., & Krigolson, O. E. (2008). The feedback correct-related positivity: Sensitivity of the event-related brain potential to unexpected positive feedback. *Psychophysiology*, *45*(5), 688–697. https://doi.org/10.1111/j.1469-8986.2008.00668.x

Holroyd, C. B., & Umemoto, A. (2016). The research domain criteria framework: The case for anterior cingulate cortex. *Neuroscience & Biobehavioral Reviews*, *71*, 418–443. https://doi.org/10.1016/j.neubiorev.2016.09.021

Hosp, J. A., Pekanovic, A., Rioult-Pedotti, M. S., & Luft, A. R. (2011). Dopaminergic Projections from Midbrain to Primary Motor Cortex Mediate Motor Skill Learning. *Journal of Neuroscience*, *31*(7), 2481–2487. https://doi.org/10.1523/JNEUROSCI.5411-10.2011

Huang, V. S., Haith, A., Mazzoni, P., & Krakauer, J. W. (2011). Rethinking Motor Learning and Savings in Adaptation Paradigms: Model-Free Memory for Successful Actions Combines with Internal Models. *Neuron*, *70*(4), 787–801. https://doi.org/10.1016/j.neuron.2011.04.012

Iannaccone, R., Hauser, T. U., Staempfli, P., Walitza, S., Brandeis, D., & Brem, S. (2015). Conflict monitoring and error processing: New insights from simultaneous EEG–fMRI. *NeuroImage*, *105*, 395–407. https://doi.org/10.1016/j.neuroimage.2014.10.028

Inzlicht, M., Shenhav, A., & Olivola, C. Y. (2018). The Effort Paradox: Effort Is Both Costly and Valued. *Trends in Cognitive Sciences*, *22*(4), 337–349. https://doi.org/10.1016/j.tics.2018.01.007

Izawa, J., & Shadmehr, R. (2011). Learning from Sensory and Reward Prediction Errors during Motor Adaptation. *PLOS Computational Biology*, *7*(3), e1002012. https://doi.org/10.1371/journal.pcbi.1002012

Jahanshahi, M., Wilkinson, L., Gahir, H., Dharmaindra, A., Dharmarinda, A., Dharminda, A., & Lagnado, D. A. (2010). Medication impairs probabilistic classification learning in Parkinson's disease. *Neuropsychologia*, *48*(4), 1096–1103. https://doi.org/10.1016/j.neuropsychologia.2009.12.010

Kennerley, S. W., Dahmubed, A. F., Lara, A. H., & Wallis, J. D. (2009). Neurons in the Frontal Lobe Encode the Value of Multiple Decision Variables. *Journal of Cognitive Neuroscience*, *21*(6), 1162–1178. https://doi.org/10.1162/jocn.2009.21100

Kennerley, S. W., Walton, M. E., Behrens, T. E. J., Buckley, M. J., & Rushworth, M. F. S. (2006). Optimal decision making and the anterior cingulate cortex. *Nature Neuroscience*, *9*(7), 940–947. https://doi.org/10.1038/nn1724

Klein-Flügge, M. C., Kennerley, S. W., Friston, K., & Bestmann, S. (2016). Neural Signatures of Value Comparison in Human Cingulate Cortex during Decisions Requiring an Effort-Reward Trade-off. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *36*(39), 10002–10015. https://doi.org/10.1523/JNEUROSCI.0292-16.2016

Kolling, N., Wittmann, M. K., Behrens, T. E. J., Boorman, E. D., Mars, R. B., & Rushworth, M. F. S. (2016). Value, search, persistence and model updating in anterior cingulate cortex. *Nature Neuroscience*, *19*(10), 1280–1285. https://doi.org/10.1038/nn.4382

Krigolson, O. E. (2018). Event-related brain potentials and the study of reward processing: Methodological considerations. *International Journal of Psychophysiology*, *132*, 175–183. https://doi.org/10.1016/j.ijpsycho.2017.11.007

Krigolson, O. E., & Holroyd, C. B. (2007). Hierarchical error processing: Different errors, different systems. *Brain Research*, *1155*, 70–80. https://doi.org/10.1016/j.brainres.2007.04.024

Krigolson, O. E., Holroyd, C. B., Van Gyn, G., & Heath, M. (2008). Electroencephalographic correlates of target and outcome errors. *Experimental Brain Research*, *190*(4), 401–411. https://doi.org/10.1007/s00221-008-1482-x

Kwak, Y., Müller, M. L. T. M., Bohnen, N. I., Dayalu, P., & Seidler, R. D. (2010). Effect of dopaminergic medications on the time course of explicit motor sequence learning in Parkinson's disease. *Journal of Neurophysiology*, *103*(2), 942–949. https://doi.org/10.1152/jn.00197.2009

Lee, E., Seo, M., Monte, O. D., & Averbeck, B. B. (2015). Injection of a Dopamine Type 2 Receptor Antagonist into the Dorsal Striatum Disrupts Choices Driven by Previous Outcomes, But Not Perceptual Inference. *Journal of Neuroscience*, *35*(16), 6298–6306. https://doi.org/10.1523/JNEUROSCI.4561-14.2015

Lemon, N., Aydin-Abidin, S., Funke, K., & Manahan-Vaughan, D. (2009). Locus coeruleus activation facilitates memory encoding and induces hippocampal LTD

that depends on beta-adrenergic receptor activation. *Cerebral Cortex (New York, N.Y.: 1991)*, *19*(12), 2827–2837. https://doi.org/10.1093/cercor/bhp065

Leng, Y., & Zhou, X. (2010). Modulation of the brain activity in outcome evaluation by interpersonal relationship: An ERP study. *Neuropsychologia*, *48*(2), 448–455. https://doi.org/10.1016/j.neuropsychologia.2009.10.002

Leow, L.-A., de Rugy, A., Loftus, A. M., & Hammond, G. (2013). Different mechanisms contributing to savings and anterograde interference are impaired in Parkinson's disease. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00055

Leow, L.-A., Loftus, A. M., & Hammond, G. R. (2012). Impaired savings despite intact initial learning of motor adaptation in Parkinson's disease. *Experimental Brain Research*, *218*(2), 295–304. https://doi.org/10.1007/s00221-012-3060-5

Luft, A. R., & Schwarz, S. (2009). Dopaminergic signals in primary motor cortex. *International Journal of Developmental Neuroscience*, *27*(5), 415–421. https://doi.org/10.1016/j.ijdevneu.2009.05.004

Luppino, G., Matelli, M., Camarda, R. M., Gallese, V., & Rizzolatti, G. (1991). Multiple representations of body movements in mesial area 6 and the adjacent cingulate cortex: An intracortical microstimulation study in the macaque monkey. *The Journal of Comparative Neurology*, *311*(4), 463–482. https://doi.org/10.1002/cne.903110403

Lydall, E. S., Gilmour, G., & Dwyer, D. M. (2010). Rats place greater value on rewards produced by high effort: An animal analogue of the "effort justification" effect. *Journal of Experimental Social Psychology*, *46*(6), 1134–1137. https://doi.org/10.1016/j.jesp.2010.05.011

MacDonald, P. A., MacDonald, A. A., Seergobin, K. N., Tamjeedi, R., Ganjavi, H., Provost, J.-S., & Monchi, O. (2011). The effect of dopamine therapy on ventral and dorsal striatum-mediated cognition in Parkinson's disease: Support from functional MRI. *Brain: A Journal of Neurology*, *134*(Pt 5), 1447–1463. https://doi.org/10.1093/brain/awr075

MacLean, S. J., Hassall, C. D., Ishigami, Y., Krigolson, O. E., & Eskes, G. A. (2015). Using brain potentials to understand prism adaptation: The error-related negativity and the P300. *Frontiers in Human Neuroscience*, *9*. https://doi.org/10.3389/fnhum.2015.00335

Manohar, S. G., Muhammed, K., Fallon, S. J., & Husain, M. (2019). Motivation dynamically increases noise resistance by internal feedback during movement.

Neuropsychologia, *123*, 19–29. https://doi.org/10.1016/j.neuropsychologia.2018.07.011

Marinelli, L., Crupi, D., Di Rocco, A., Bove, M., Eidelberg, D., Abbruzzese, G., & Felice Ghilardi, M. (2009). Learning and consolidation of visuo-motor adaptation in Parkinson's disease. *Parkinsonism & Related Disorders*, *15*(1), 6–11. https://doi.org/10.1016/j.parkreldis.2008.02.012

Matsumoto, M., Matsumoto, K., Abe, H., & Tanaka, K. (2007). Medial prefrontal cell activity signaling prediction errors of action values. *Nature Neuroscience*, *10*(5), 647–656. https://doi.org/10.1038/nn1890

Molina-Luna, K., Pekanovic, A., Röhrich, S., Hertler, B., Schubring-Giese, M., Rioult-Pedotti, M.-S., & Luft, A. R. (2009). Dopamine in Motor Cortex Is Necessary for Skill Learning and Synaptic Plasticity. *PLOS ONE*, *4*(9), e7082. https://doi.org/10.1371/journal.pone.0007082

Murphy, P. R., Robertson, I. H., Balsters, J. H., & O'connell, R. G. (2011). Pupillometry and P3 index the locus coeruleus-noradrenergic arousal function in humans. *Psychophysiology*, *48*(11), 1532–1543. https://doi.org/10.1111/j.1469-8986.2011.01226.x

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

Parker, N. F., Cameron, C. M., Taliaferro, J. P., Lee, J., Choi, J. Y., Davidson, T. J., Daw, N. D., & Witten, I. B. (2016). Reward and choice encoding in terminals of midbrain dopamine neurons depends on striatal target. *Nature Neuroscience*, *19*(6), 845–854. https://doi.org/10.1038/nn.4287

Paus, T. (2001). Primate anterior cingulate cortex: Where motor control, drive and cognition interface. *Nature Reviews Neuroscience*, *2*(6), 417–424. https://doi.org/10.1038/35077500

Pekny, S. E., Izawa, J., & Shadmehr, R. (2015). Reward-dependent modulation of movement variability. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *35*(9), 4015–4024. https://doi.org/10.1523/JNEUROSCI.3244-14.2015

Pettigrew, J. D. (1982). Pharmacologic control of cortical plasticity. *Retina (Philadelphia, Pa.)*, *2*(4), 360–372.

Pettigrew, J. D., & Kasamatsu, T. (1978). Local perfusion of noradrenaline maintains visual cortical plasticity. *Nature*, *271*(5647), 761–763. https://doi.org/10.1038/271761a0

Pfabigan, D. M., Alexopoulos, J., Bauer, H., & Sailer, U. (2011). Manipulation of feedback expectancy and valence induces negative and positive reward prediction error signals manifest in event-related brain potentials. *Psychophysiology*, *48*(5), 656–664. https://doi.org/10.1111/j.1469-8986.2010.01136.x

Poe, G. R., Foote, S., Eschenko, O., Johansen, J. P., Bouret, S., Aston-Jones, G., Harley, C. W., Manahan-Vaughan, D., Weinshenker, D., Valentino, R., Berridge, C., Chandler, D. J., Waterhouse, B., & Sara, S. J. (2020). Locus coeruleus: A new look at the blue spot. *Nature Reviews Neuroscience*, *21*(11), 644–659. https://doi.org/10.1038/s41583-020-0360-9

Polich, J. (2007). Updating P300: An Integrative Theory of P3a and P3b. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, *118*(10), 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019

Porter, B. S., Hillman, K. L., & Bilkey, D. K. (2019). Anterior cingulate cortex encoding of effortful behavior. *Journal of Neurophysiology*, *121*(2), 701–714. https://doi.org/10.1152/jn.00654.2018

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

Quattrocchi, G., Monaco, J., Ho, A., Irmen, F., Strube, W., Ruge, D., Bestmann, S., & Galea, J. M. (2018). Pharmacological Dopamine Manipulation Does Not Alter Reward-Based Improvements in Memory Retention during a Visuomotor Adaptation Task. *ENeuro*, *5*(3). https://doi.org/10.1523/ENEURO.0453-17.2018

Reuter, E.-M., Pearcey, G. E. P., & Carroll, T. J. (2018). Greater neural responses to trajectory errors are associated with superior force field adaptation in older adults. *Experimental Gerontology*, *110*, 105–117. https://doi.org/10.1016/j.exger.2018.05.020

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004). The Role of the Medial Frontal Cortex in Cognitive Control. *Science*, *306*(5695), 443–447. https://doi.org/10.1126/science.1100301

Rothenhoefer, K. M., Costa, V. D., Bartolo, R., Vicario-Feliciano, R., Murray, E. A., & Averbeck, B. B. (2017). Effects of Ventral Striatum Lesions on Stimulus-Based versus Action-Based Reinforcement Learning. *The Journal of Neuroscience: The*

Official Journal of the Society for Neuroscience, *37*(29), 6902–6914. https://doi.org/10.1523/JNEUROSCI.0631-17.2017

Sales, A. C., Friston, K. J., Jones, M. W., Pickering, A. E., & Moran, R. J. (2019). Locus Coeruleus tracking of prediction errors optimises cognitive flexibility: An Active Inference model. *PLOS Computational Biology*, *15*(1), e1006267. https://doi.org/10.1371/journal.pcbi.1006267

Salgado, H., Köhr, G., & Treviño, M. (2012). Noradrenergic "tone" determines dichotomous control of cortical spike-timing-dependent plasticity. *Scientific Reports*, *2*, 417. https://doi.org/10.1038/srep00417

Sambrook, T. D., & Goslin, J. (2015). A neural reward prediction error revealed by a meta-analysis of ERPs using great grand averages. *Psychological Bulletin*, *141*(1), 213–235. http://dx.doi.org.proxy1.lib.uwo.ca/10.1037/bul0000006

Sambrook, T. D., & Goslin, J. (2016). Principal components analysis of reward prediction errors in a reinforcement learning task. *NeuroImage*, *124*, 276–286. https://doi.org/10.1016/j.neuroimage.2015.07.032

Santesso, D. L., Dillon, D. G., Birk, J. L., Holmes, A. J., Goetz, E., Bogdan, R., & Pizzagalli, D. A. (2008). Individual differences in reinforcement learning: Behavioral, electrophysiological, and neuroimaging correlates. *NeuroImage*, *42*(2), 807–816. https://doi.org/10.1016/j.neuroimage.2008.05.032

Sara, S. J., & Segal, M. (1991). Plasticity of sensory responses of locus coeruleus neurons in the behaving rat: Implications for cognition. *Progress in Brain Research*, *88*, 571–585. https://doi.org/10.1016/s0079-6123(08)63835-2

Savoie, F.-A., Thénault, F., Whittingstall, K., & Bernier, P.-M. (2018). Visuomotor Prediction Errors Modulate EEG Activity Over Parietal Cortex. *Scientific Reports*, *8*(1), 12513. https://doi.org/10.1038/s41598-018-30609-0

Savoie, F. -A., Hamel, R., Lacroix, A., Thénault, F., Whittingstall, K., & Bernier, P. -M. (2019). Luring the motor system: impact of performance-contingent incentives on pre-movement beta-band activity and motor performance. *Journal of Neuroscience*, *39*(15), 2903-2914.

Schultz, W. (2016). Dopamine reward prediction error coding. *Dialogues in Clinical Neuroscience*, *18*(1), 23–32.

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

Seo, H., & Lee, D. (2007). Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *27*(31), 8366–8377. https://doi.org/10.1523/JNEUROSCI.2369-07.2007

Seo, M., Lee, E., & Averbeck, B. B. (2012). Action selection and action value in frontal-striatal circuits. *Neuron*, *74*(5), 947–960. https://doi.org/10.1016/j.neuron.2012.03.037

Shadmehr, R., Huang, H. J., & Ahmed, A. A. (2016). A Representation of Effort in Decision-Making and Motor Control. *Current Biology*, *26*(14), 1929–1934. https://doi.org/10.1016/j.cub.2016.05.065

Shenhav, A., Cohen, J. D., & Botvinick, M. M. (2016). Dorsal anterior cingulate cortex and the value of control. *Nature Neuroscience*, *19*(10), 1286–1291. https://doi.org/10.1038/nn.4384

Shima, K., & Tanji, J. (1998). Role for cingulate motor area cells in voluntary movement selection based on reward. *Science (New York, N.Y.)*, *282*(5392), 1335–1338. https://doi.org/10.1126/science.282.5392.1335

Showers, M. J. (1959). The cingulate gyrus: Additional motor area and cortical autonomic regulator. *The Journal of Comparative Neurology*, *112*, 231–301. https://doi.org/10.1002/cne.901120118

Sidarta, A., Vahdat, S., Bernardi, N. F., & Ostry, D. J. (2016). Somatic and Reinforcement-Based Plasticity in the Initial Stages of Human Motor Learning. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *36*(46), 11682–11692. https://doi.org/10.1523/JNEUROSCI.1767-16.2016

Skvortsova, V., Palminteri, S., & Pessiglione, M. (2014). Learning to minimize efforts versus maximizing rewards: Computational principles and neural correlates. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *34*(47), 15621–15630. https://doi.org/10.1523/JNEUROSCI.1350-14.2014

Smith, E. H., Banks, G. P., Mikell, C. B., Cash, S. S., Patel, S. R., Eskandar, E. N., & Sheth, S. A. (2015). Frequency-Dependent Representation of Reinforcement-Related Information in the Human Medial and Lateral Prefrontal Cortex. *Journal of Neuroscience*, *35*(48), 15827–15836. https://doi.org/10.1523/JNEUROSCI.1864-15.2015

Sokhadze, E. M., Casanova, M. F., Casanova, E. L., Lamina, E., Kelly, D. P., & Khachidze, I. (2017). Event-related Potentials (ERP) in Cognitive Neuroscience Research and Applications. *NeuroRegulation*, *4*(1), 14–14. https://doi.org/10.15540/nr.4.1.14

Soltani, M., & Knight, R. T. (2000). Neural origins of the P300. *Critical Reviews in Neurobiology*, *14*(3–4), 199–224.

Stanton, P. K., & Sarvey, J. M. (1985). Blockade of norepinephrine-induced long-lasting potentiation in the hippocampal dentate gyrus by an inhibitor of protein synthesis. *Brain Research*, *361*(1–2), 276–283. https://doi.org/10.1016/0006-8993(85)91299-5

Strick, P. L., Dum, R. P., & Picard, N. (1998). Motor areas on the medial wall of the hemisphere. *Novartis Foundation Symposium*, *218*, 64–75; discussion 75-80, 104–108. https://doi.org/10.1002/9780470515563.ch5

Summerside, E. M., Shadmehr, R., & Ahmed, A. A. (2018). Vigor of reaching movements: Reward discounts the cost of effort. *Journal of Neurophysiology*, *119*(6), 2347–2357. https://doi.org/10.1152/jn.00872.2017

Swainson, R., Rogers, R. D., Sahakian, B. J., Summers, B. A., Polkey, C. E., & Robbins, T. W. (2000). Probabilistic learning and reversal deficits in patients with Parkinson's disease or frontal or temporal lobe lesions: Possible adverse effects of dopaminergic medication. *Neuropsychologia*, *38*(5), 596–612. https://doi.org/10.1016/s0028-3932(99)00103-7

Swick, D., Pineda, J. A., Schacher, S., & Foote, S. L. (1994). Locus coeruleus neuronal activity in awake monkeys: Relationship to auditory P300-like potentials and spontaneous EEG. *Experimental Brain Research*, *101*(1), 86–92. https://doi.org/10.1007/BF00243219

Taylor, J. A., Klemfuss, N. M., & Ivry, R. B. (2010). An explicit strategy prevails when the cerebellum fails to compute movement errors. *Cerebellum (London, England)*, *9*(4), 580–586. https://doi.org/10.1007/s12311-010-0201-x

Therrien, A. S., Wolpert, D. M., & Bastian, A. J. (2016). Effective reinforcement learning following cerebellar damage requires a balance between exploration and motor noise. *Brain*, *139*(1), 101–114. https://doi.org/10.1093/brain/awv329

Torrecillos, F., Albouy, P., Brochier, T., & Malfait, N. (2014). Does the processing of sensory and reward-prediction errors involve common neural resources? Evidence from a frontocentral negative potential modulated by movement execution errors. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *34*(14), 4845–4856. https://doi.org/10.1523/JNEUROSCI.4390-13.2014

Torta, D. M. E., Castelli, L., Zibetti, M., Lopiano, L., & Geminiani, G. (2009). On the role of dopamine replacement therapy in decision-making, working memory, and

reward in Parkinson's disease: Does the therapy-dose matter? *Brain and Cognition*, *71*(2), 84–91. https://doi.org/10.1016/j.bandc.2009.04.003

Varazzani, C., San-Galli, A., Gilardeau, S., & Bouret, S. (2015). Noradrenaline and Dopamine Neurons in the Reward/Effort Trade-Off: A Direct Electrophysiological Comparison in Behaving Monkeys. *The Journal of Neuroscience*, *35*(20), 7866–7877. https://doi.org/10.1523/JNEUROSCI.0454-15.2015

Vazey, E. M., Moorman, D. E., & Aston-Jones, G. (2018). Phasic locus coeruleus activity regulates cortical encoding of salience information. *Proceedings of the National Academy of Sciences*, *115*(40), E9439–E9448. https://doi.org/10.1073/pnas.1803716115

Vo, A., Seergobin, K. N., & MacDonald, P. A. (2017). Effects of levodopa on stimulus-response learning versus response selection in healthy young adults. *Behavioural Brain Research*, *317*, 553–561. https://doi.org/10.1016/j.bbr.2016.10.019

Vo, A., Seergobin, K. N., & MacDonald, P. A. (2018). Independent effects of age and levodopa on reversal learning in healthy volunteers. *Neurobiology of Aging*, *69*, 129–139. https://doi.org/10.1016/j.neurobiolaging.2018.05.014

Vo, A., Seergobin, K. N., Morrow, S. A., & MacDonald, P. A. (2016). Levodopa impairs probabilistic reversal learning in healthy young adults. *Psychopharmacology*, *233*(14), 2753–2763. https://doi.org/10.1007/s00213-016-4322-x

Walsh, M. M., & Anderson, J. R. (2012). Learning from experience: Event-related potential correlates of reward processing, neural adaptation, and behavioral choice. *Neuroscience and Biobehavioral Reviews*, *36*(8), 1870–1884. https://doi.org/10.1016/j.neubiorev.2012.05.008

Walton, M. E., Croxson, P. L., Behrens, T. E. J., Kennerley, S. W., & Rushworth, M. F. S. (2007). Adaptive decision making and value in the anterior cingulate cortex. *NeuroImage*, *36*, T142–T154. https://doi.org/10.1016/j.neuroimage.2007.03.029

Walton, M. E., Kennerley, S. W., Bannerman, D. M., Phillips, P. E. M., & Rushworth, M. F. S. (2006). Weighing up the benefits of work: Behavioral and neural analyses of effort-related decision making. *Neural Networks: The Official Journal of the International Neural Network Society*, *19*(8), 1302–1314. https://doi.org/10.1016/j.neunet.2006.03.005

Warren, C. M., Hyman, J. M., Seamans, J. K., & Holroyd, C. B. (2015). Feedback-related negativity observed in rodent anterior cingulate cortex. *Journal of Physiology-Paris*, *109*(1), 87–94. https://doi.org/10.1016/j.jphysparis.2014.08.008

Watanabe, T., Mima, T., Shibata, S., & Kirimoto, H. (2021). Midfrontal theta as moderator between beta oscillations and precision control. *NeuroImage*, *235*, 118022. https://doi.org/10.1016/j.neuroimage.2021.118022

Williams, Z. M., Bush, G., Rauch, S. L., Cosgrove, G. R., & Eskandar, E. N. (2004). Human anterior cingulate neurons and the integration of monetary reward with motor responses. *Nature Neuroscience*, *7*(12), 1370–1375. https://doi.org/10.1038/nn1354

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

Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). The neural basis of error detection: Conflict monitoring and the error-related negativity. *Psychological Review*, 931–959.

Yokoi, A., & Weiler, J. (2021). *Pupil-linked arousal modulates trial-by-trial motor adaptation in humans* [Preprint]. Neuroscience. https://doi.org/10.1101/2021.04.03.438075

Yoon, T., Geary, R. B., Ahmed, A. A., & Shadmehr, R. (2018). Control of movement vigor and decision making during foraging. *Proceedings of the National Academy of Sciences*, *115*(44), E10476–E10485. https://doi.org/10.1073/pnas.1812979115

Yoon, T., Jaleel, A., Ahmed, A. A., & Shadmehr, R. (2020). Saccade vigor and the subjective economic value of visual stimuli. *Journal of Neurophysiology*, *123*(6), 2161–2172. https://doi.org/10.1152/jn.00700.2019

Zénon, A., Sidibé, M., & Olivier, E. (2014). Pupil size variations correlate with physical effort perception. *Frontiers in Behavioral Neuroscience*, *8*. https://doi.org/10.3389/fnbeh.2014.00286

Zentall, T. R. (2010). Justification of Effort by Humans and Pigeons: Cognitive Dissonance or Contrast? *Current Directions in Psychological Science*, *19*(5), 296–300. https://doi.org/10.1177/0963721410383381

Zhou, Z., Yu, R., & Zhou, X. (2010). To do or not to do? Action enlarges the FRN and P300 effects in outcome evaluation. *Neuropsychologia*, *48*(12), 3606–3613. https://doi.org/10.1016/j.neuropsychologia.2010.08.010

Appendix: Ethics Approval

Western Research

Date: 22 February 2021

To: Dr. Paul Gribble

Project ID: 108682

Study Title: Neural Signatures of Prediction Error in Reinforcement Based Motor Learning

Application Type: Continuing Ethics Review (CER) Form

Review Type: Delegated

REB Meeting Date: 09/March/2021

Date Approval Issued: 22/Feb/2021

REB Approval Expiry Date: 24/Feb/2022

Dear Dr. Paul Gribble,

The Western University Research Ethics Board has reviewed the application. This study, including all currently approved documents, has been re-approved until the expiry date noted above.

REB members involved in the research project do not participate in the review, discussion or decision.

Western University REB operates in compliance with, and is constituted in accordance with, the requirements of the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (TCPS 2); the International Conference on Harmonisation Good Clinical Practice Consolidated Guideline (ICH GCP); Part C, Division 5 of the Food and Drug Regulations; Part 4 of the Natural Health Products Regulations; Part 3 of the Medical Devices Regulations and the provisions of the Ontario Personal Health Information Protection Act (PHIPA 2004) and its applicable regulations. The REB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000940.

Please do not hesitate to contact us if you have any questions.

Sincerely,

The Office of Human Research Ethics

Note: This correspondence includes an electronic signature (validation and approval via an online system that is compliant with all regulations).

Western Research

Date: 24 November 2020

To: Dr. Penny MacDonald

Project ID: 102018

Study Title: Distinguishing the roles of ventral and dorsal striatum in cognition (REB #18517)

Application Type: Continuing Ethics Review (CER) Form

Review Type: Delegated

REB Meeting Date: 1/Dec/2020

Date Approval Issued: 24/Nov/2020

REB Approval Expiry Date: 29/Nov/2021

Dear Dr. Penny MacDonald,

The Western University Research Ethics Board has reviewed the application. This study, including all currently approved documents, has been re-approved until the expiry date noted above.

REB members involved in the research project do not participate in the review, discussion or decision.

Western University REB operates in compliance with, and is constituted in accordance with, the requirements of the TriCouncil Policy Statement: Ethical Conduct for Research Involving Humans (TCPS 2); the International Conference on Harmonisation Good Clinical Practice Consolidated Guideline (ICH GCP); Part C, Division 5 of the Food and Drug Regulations; Part 4 of the Natural Health Products Regulations; Part 3 of the Medical Devices Regulations and the provisions of the Ontario Personal Health Information Protection Act (PHIPA 2004) and its applicable regulations. The REB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000940.

Please do not hesitate to contact us if you have any questions.

Sincerely,

The Office of Human Research Ethics

Note: This correspondence includes an electronic signature (validation and approval via an online system that is compliant with all regulations).

Curriculum Vitae

Education

Western University (London, ON): *PhD Neuroscience*
Expected 2021

McGill University (Montreal, QC): *BSc Neuroscience*
2015

- Minor: Cognitive Science

Publications

Palidis DJ, Gribble PL (2020). EEG correlates of physical effort and reward processing during reinforcement learning. Journal of neurophysiology, 124(2), 610–622

Cashaback J, Lao CK, **Palidis DJ**, Coltman SK, McGregor HR, Gribble PL (2019). The gradient of the reinforcement landscape influences sensorimotor learning. PLoS computational biology, 15(3), e1006839

Palidis DJ, Cashaback J, Gribble PL (2019). Neural signatures of reward and sensory error feedback processing in motor learning. Journal of neurophysiology, 121(4), 1561–1574

Palidis DJ, Wyder-Hodge PA, Fooken J, Spering M (2017). Distinct eye movement patterns enhance dynamic visual acuity. PloS one, 12(2), e0172061

Ming W, **Palidis DJ**, Spering M, McKeown MJ (2016). Visual Contrast Sensitivity in Early-Stage Parkinson's Disease. Investigative ophthalmology & visual science, 57(13), 5696–5704

Palidis DJ, McGregor HR, Vo A, Macdonald PA, Gribble PL (2021). Null effects of levodopa on reward- and error-based motor adaptation, savings, and anterograde interference. BioRxiv, doi: https://doi.org/10.1101/2020.11.19.390302 (In press by Journal of Neurophysiology)

Oral Presentations

Palidis DJ, Gribble PL (2020). EEG correlates of reward and effort processing during reinforcement learning. *Southern Ontario Motor Behavior Symposium*. Toronto, Canada

Cashaback J, Lao CK, **Palidis DJ**, Coltman SK, McGregor HR, Gribble PL (2017). The reinforcement landscape influences sensorimotor learning. *Journal of Exercise, Movement, & Sport*. St. Johns, Canada

Poster Presentations

Palidis DJ, Gribble PL (2019). EEG correlates of reward and effort processing in reinforcement learning. NCM, Toyama, Japan

Palidis DJ, Gribble PL (2019). EEG correlates of reward and effort processing in reinforcement learning. RLDM, Montreal, Canada

Palidis DJ, Gribble PL (2019). EEG correlates of reward and effort processing during reinforcement learning. SFN, Chicago, USA

Palidis DJ, Cashaback J, Gribble PL (2018). ERP Correlates of Feedback Processing in Reward and Error Based Motor Learning. CSBBCS, St. John's, Canada

Palidis DJ, Gribble PL (2018). Neural substrates of reward and sensory error based reach adaptation. SFN, San Diego, USA

Palidis DJ, Cashaback J, Gribble PL (2018). Distinct neural signatures of reward and sensory prediction error in motor learning. CAN, Vancouver, Canada

Palidis DJ, Cashaback J, Gribble PL (2018). Distinct neural signatures of reward and sensory prediction error in motor learning. Neural Control of Movement, Santa Fe, USA

Cashaback J, Lao C, McGregor HR, **Palidis DJ**, Coltman SK, Gribble PL (2017). Reinforcement gradient ascent during sensorimotor learning. Neural Control of Movement, Dublin, Ireland

Lao C, Cashaback J, McGregor HR, **Palidis DJ**, Coltman SK, Gribble PL (2017). Ascending the reinforcement gradient during sensorimotor learning. Southern Ontario Neuroscience Associatio, St. Catherines, Canada

Awards

NSERC Post Graduate Scholarship - Doctoral (Value: CAD 63,000). 5/2018 – 4/2021

Ontario Graduate Scholarship (Value: CAD 15,000). Declined. 2018.

NSERC Undergraduate Student Research Award (Value: CAD 4,500). 5/2014 - 8/2014