Motor Expectancy: The Modulation of the Reward Positivity in a Reinforcement Learning Motor Task

by

ROBERT TRSKA

B.Sc., Dalhousie University, 2015

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE

In Interdisciplinary Studies

Supervisory Committee

Motor Expectancy: The Modulation of the Reward Positivity in a Reinforcement Learning Motor Task

By

ROBERT TRSKA

B.Sc., Dalhousie University, 2015

Supervisory Committee

Dr. Olav Krigolson, School of Exercise Science, Physical Health and Education
Supervisor

Dr. Clay Holroyd, Department of Psychology
Co-Supervisor

Abstract

An adage posits that we learn from our mistakes; however, this is not entirely true. According to reinforcement learning theory, we learn when the expectation of our actions differs from outcomes. Here, we examined whether expectancy driven learning lends a role in motor learning. Given the vast amount of overlapping anatomy and circuitry within the brain with respect to reward and motor processes, it is appropriate to examine both motor control and expectancy processes within a singular task. In the current study, participants performed a line drawing task via tablet under conditions of changing expectancies. Participants were provided feedback in a reinforcement-learning manner, as positive (✓) or negative (x) based off their performance. Modulation of expected outcomes were reflected by changes in amplitude of the human event-related potential (ERP), the reward positivity. The reward positivity is thought to reflect phasic dopamine release from the mesolimbic dopaminergic system to the basal ganglia and cingulate cortex. Due to the overlapping circuitry of reward and motor pathways, another human ERP, the bereitschatftspotential (BP), was examined. The BP is implicated in motor planning and execution; however, the late aspect of the BP shares similarity with the contingent negative variability (CNV). Current evidence demonstrates a relationship between expectancy and reward positivity amplitude in a motor learning context, as well as modulation of the BP under difficult task conditions. Behavioural data supports prior literature and may suggest a connection between sensory motor prediction errors working in concert with reward prediction errors. Further evidence supports a frontal-medial evaluation system for motor errors. Additionally, results support prior evidence of motor plans being formed upon target observation and held in memory until motor execution, rather than their formation before movement onset.

TABLE OF CONTENTS

List of Tables

List of Figures

Acknowledgements

There are many people and sources who have made this thesis possible. Firstly, I would like to thank Olav Krigolson for taking me in under his guidance and providing me with opportunities I never thought would be possible. Second, I would like to thank Clay Holroyd for offering his insights and direction for my thesis, as well as hosting many interesting and informative meetings during my study. Third I would like to thank Chad Williams and Cameron Hassall, to which their coding expertise has made the following experiment possible. Fourth I would like to give a heartfelt and deep thanks to the Neuroeconomics lab. I have been part of this team for many years, spanning from my undergraduate to my master's degree, from coast to coast. I am glad to have had the privilege of being part of such a fine team, and I look forward towards many years of jolly cooperation. Lastly, I would like to thank my family to which none of this would have been possible. Through easy and tough times, I knew I was always able to depend on my parents (Anna and Les), my siblings (Richard and Nicole), and relatives (Mike, Sophie, and Stella) for continual support.

CHAPTER 1: A REVIEW OF REINFORCEMENT LEARNING, REWARD EXPECTATION, AND MOTOR LEARNING

1.1 Introduction

The intent of the following overview aims to visit past and current literature in the fields of motor control and reinforcement learning. We will first establish prior theories and adaptations of reinforcement and reward expectation learning, followed by a glimpse at the neural anatomy and mechanisms that facilitate these processes. In addition to this, we will examine these mechanisms with a motor control and learning context in mind. Highlighted landmarks include a mid-frontal region of the brain with integrative and modulatory properties for many cognitive and behavioural systems, known as the anterior cingulate cortex[1] (ACC: Paus, 2001). After its role in learning and motor control is established, several other cortical and subcortical structures will be examined. As the ACC has many efferent and afferent projections to a rich variety of brain regions, a secondary focus will be placed on the primary motor cortex (Goldburg, 1985), supplementary motor area (Goldburg, 1985), premotor regions (Haggard, 2008), and the numerous nuclei of the basal ganglia (Groenwegen, 2003). These regions all play a fundamental role in motor learning. Interestingly, the basal ganglia also subsume an influential role in reward processing behaviours (Tachibana & Hikosaka, 2012), with overlapping activity in reward processing and motor behaviours (Kawagoe, Takikawa, & Hikosaka, 1998). Due to the robust connectivity of the basal ganglia with ACC and motor regions, a following section will be dedicated to their roles in reinforcement and motor learning. Outlined will example incidence of

[1] Current nomenclature has shifted away from the term "Anterior Cingulate Cortex" in lieu of "Anterior Midcingulate Cortex" due to division of regions (Vogt, 2005; Hoffstaeder et al., 2013). For consistency's purpose of this thesis, what is referred as Anterior Cingulate will in fact reference the Anterior Midcingulate.

sensitivity to reward and expectation (Cromwell & Schultz, 2003), as well as highlighting overlap in motor control systems and reward, and associated neural substrates of expectancy.

A myriad of imaging techniques may be used to record expectancy-related activation within the brain (fMRI: Hauser et al., 2014; cellular: Shidara & Richmond, 2002). In the present study electroencephalography (EEG) will be employed (Holroyd & Coles, 2002) to index expectancy effects in participants. From this, electrical activity generated by the brain can be recorded, thus allowing us to produce inferences on levels of activation. To fully appreciate the nuance of EEG, we will establish what is being record and how it is recorded. A review of the methodology of EEG will discuss the physiology and applications of this neuroimaging technique. Topics of this review will involve human event related potentials (ERPs) and their derivative components: the reward positivity[2] for cataloging expectation effects, and motor potentials for motor related activity in the brain. Thus, in broad terms the current thesis will be investigating the connection between motor learning and reward expectation processes. To facilitate this goal, we will utilize EEG in a motor reward expectancy context, and examine the interactions of motor behaviour, expectancy, and levels of cortical activation, as indexed by component analysis.

1.2 Learning and Expectancy

To first understand expectancy, we must establish some general concepts of learning. To begin, we may examine the insights of John Stuart Mill (c.f. utilitarianism, Mill, 1863), in which the principle of utility is applied to human behaviours. Foundational to this principle is the notion of action, and the amount of utility (happiness) it may provide; that is, we as human beings are

[2] The reward positivity is a component with many terms. Classically it has been referenced as the feedback related negativity (FRN), however due to novel approaches in its assessment and functionality, terminology has shifted towards the new nomenclature. As such, any reference to FRN will be termed reward positivity in the current thesis. See Proudfit (2015) or Krigolson (2017) for further detail.

continually striving to maximize our utility or freedom from pain and to avoid noxious situations we may experience. Indeed, this philosophy appears to exist in practice within the domains of learning. Demonstrated many years later, Thorndike's law of effect postulates that reward following a behaviour will serve to strengthen the likelihood of said action and that it would be repeated (Thorndike, 1911). With this in mind, future adaptations of utilitarianism and Thorndike's law are evident across several learning types. To begin we will discuss classical conditioning.

Classical conditioning is a form of learning such that unassociated stimuli are linked together by some event, to produce a conditioned response. This method is best exampled through the historic work of Ivan Pavlov (2010), in which observations of the salivary and gustatory response of dogs to tones indicated an underlying learning mechanism. This learning is exampled by examining the variable responses and stimulus within a paradigm. At the onset, there is a normal unconditioned response (i.e., dog salivating to food). To condition a new response, an unconditioned stimulus (i.e., food to a dog) is paired with a conditioned stimulus (i.e., an audible tone or metronome). This pairing is followed by a conditioned response (i.e., salivation). With sufficient pairings of food and audio tones, a dog is then conditioned to salivate to the tone as there is now a developed association of food with tone. Indeed, this associative type learning is shown to be modulated by rewarding situations (Zamble, 1968), to which rats conditioned to food presentations demonstrated a greater level of excitement anticipation to the learned cue.

Further adaptation of typical classical conditioning paradigms were the models of Rescorla and Wagner (Rescorla, 1972), to which learning was not only predicated on number of pairings but also the predictability of the stimulus. By manipulating the expectancy of stimulus

pairings, it was found that surprising or unexpected events procured larger impacts on behaviour. Mathematically, this model was dependent on several considerations. First was the creation of an original prediction to an outcome (i.e., a light comes on, and the predicted outcome is a $2 bill); the second was the actual event outcome (i.e., a light comes on, and a $20 bill is the outcome). What appears to be the driving force behind learning and behaviour modification is the difference between a prediction and the actual outcome, coined a prediction error[3]. It is seen then that the amount of learning is proportional to the magnitude of prediction errors.

Further modification of prediction error models is also evident in the realm of reinforcement learning. Foundational to this learning is the notion of exploration and exploitation, in that for a system or organism to maximize its utility, it must select potential options (Sutton & Barto, 1998). For example, consider a situation in which you move to a new city for a job opportunity. You have a new place of residence and are familiar with the neighborhood but are grossly unfamiliar with the city where your work is. On your first day to work, you select a commute that you find full of traffic and inevitably arrive to work late. On the second day, you choose a new route, and make a point to avoid the prior commute, finding yourself on time, but just barely. On the third day, you choose another different route, and find yourself with ample time and in fact early; from here on, you determine that this is the best route to travel. Much in line with Thorndike's Law (1911), the reception of beneficial outcomes shapes and influences behaviour. More specifically, an adaptation of reinforcement learning demonstrates a reliance on prediction errors in learning. Termed temporal difference learning, this model encompasses prediction errors and principles of reinforcement, to which magnitude of

[3] It should be of note any reference in the current thesis to prediction error are in relation to reward prediction errors, unless specified otherwise. This is due to overlapping terminology such that sensory prediction errors are also a computation that occur during motor movement (Izawa & Shadmehr, 2011)

prediction errors affects future trials (Sutton & Barto, 1998). To this extent, it is seen that human learning stems not from the commission of arbitrary mistakes, but instead from prediction errors based on feedback.

Curiously, learning from prediction errors also demonstrates an evident overlapping in the behaviour of select cell populations. Research by Schultz, Dayan, and Montague (1997) demonstrated that activation patterns of dopaminergic neurons from the ventral tegmental area (VTA) of the midbrain follow a very similar model to reward-prediction models (Rescorla & Wagner, 1972). To elaborate, phasic activation of these neurons is seen to increase once an association of a cue begins to predict an expected reward. Deviation from this expectation, such as no reward after a learned cue, lead to a suppression of dopaminergic firing rates, such that their activation and inhibition begins following proposed reward expectation models. It is evident that both behaviour and physiology utilize some form of prediction error or reward processing to learn and thus code future or anticipatory responses. Due to the involvement of the mesencephalic dopaminergic system (Montague, Dayan, Sejnowski, 1996) and the basal ganglia (O'Doherty et al., 2004) in learning, it is imperative in the following sections to outline the function and anatomy in regions involved with a motor consideration in mind.

1.3 Anatomy of Motor Learning

The only meaningful way to learn and interact with our physical environment is through the generation of motor movements: every action we commit involves the coordination of brain-to-muscle communication, predominantly by corticospinal or corticobulbar tracts (Porter & Lemon, 1993; Cruccu, Berardelli, Inghilleri, & Manfredi, 1990). By way of these connections we may produce fluid and accurate movements to accomplish goals or influence our surroundings. It is of question then on how motor movements are prepared, and by extension how we may learn

or acquire new motor skills. To facilitate the complexity and variability of movement, many different cortical regions are recruited, each with their own distinct properties contributing to motion planning and execution.

Examination of these regions makes evident a large overlap associated with motor control and reward prediction centers of the brain (Tachibana & Hirosaka, 2012; Balleine, Delgado, & Hirosaka, 2007; Cromwell & Shultz, 2003; Doya, 2000; Graybiel, Aosaki, Flaherty, & Kimura, 1994; Hirosaka, Nakamua, Sakai, & Nakahara, 2002; Isomura et al., 2013). This is unsurprising however, as when considered, goal directed actions are facilitated through generation of motor movements. If we are to obtain and gain what we seek, we must move our bodies to achieve those ends. Commission of movement involves many different cortical and subcortical regions to generate the sufficient motor plan to execute and respond to a desired action. Motor action is a vastly complex process involving several cortical (Ball et al., 1999; Kapogiannis, Campion, Grafman, & Wassermann; 2008), subcortical (Calabresi, Picconi, Tozzi, Ghiglieri, & Filippo, 2014; Doya, 2000; Groenewegen, 2003), and cerebellar (Doya, 2000; Wolpert, Miall, & Kawato, 1998) domains to sufficiently and accurately facilitate an intended action. Working in parallel, these systems appear to have an intimate link to reward processing systems, and thus the following sections investigate pertinent regions involved in both reward processing and motor control.

1.3.1 Cingulate Cortex

The human cingulate cortex lies towards the medial walls of the brain, following the body of the corpus callosum. Although there is some contention to its division, the most anterior portion of the cingulate falls most medial to the frontal lobes (Bush et al., 2002). As stated previously, the ACC possesses many integrative functions, receiving a variety of different

cortico-cortico, cortico-thalamo, and cortico-basal ganglionic projections, as well as a large sum of input from the mesolimbic dopaminergic system (Paus, 2001).

In term of influence over motor regions, the ACC extends many projections to prefrontal cortices, ranging from supplementary motor regions (SMA) to premotor regions (PMC: Goldburg, 1985). Cingulate motor areas (CMA), located deeper into the cingulate sulcus, possess reciprocal connections to the primary motor cortex (M1) and SMA (Devinsky, Morell, & Vogt, 1995). Pertaining to cortico-ganglionic projections, the ventral striatum receives cortical input from the ACC (Cummings, 1993), with the ventral striatum consisting of nucleus accumbens (nACC), and ventromedial portions of the caudate nucleus (CN) and putamen (Put: Cummings, 1993).

The cingulate cortex has also been associated as a site of reward processing. Several experiments, utilizing source localization methods and functional magnetic resonance imaging (Hauser et al., 2014: fMRI) have placed the generation of ACC prediction errors in the dorsal ACC (dACC). A study by Amiez, Joseph, and Procyk (2005) investigated ACC activation in Rhesus monkeys, with electrode probe placement at the dorsal bank of the ACC. In line with theories positing the dACC in reward prediction, Amiez and colleagues demonstrated dACC activity depended on several factors such as reward magnitude and probability, in addition to not being bound to any specific modality (i.e. vision, hearing, touch). Shidara and Richmond (2002) also placed reward expectation activity within the dorsal aspect (ventral bank of the anterior cingulate sulcus) of the ACC in primates. Many other studies and reviews have also localized the ACC as a site of reward processing by way of EEG (Devinsky et al., 1995; Hajcak, Moser, & Holroyd, 2007; Holroyd, Dien, & Coles; 1998; Holroyd, Krigolson, Baker, Lee, & Gibson, 2009; Holroyd & Krigolson, 2007; Holroyd & Yeung 2012; Krigolson, & Holroyd, 2007a; Krigolson,

Hassall, & Handy, 2014; Margulies et al., 2007; Miltner, Braun, & Coles, 1997; Paus, 2001; Proudfit, 2015).

Furthermore, the ACC in on itself may be a region dedicated to much more than merely reward expectation processing. A review by Holroyd and Yeung (2012) attempted to reconcile several outstanding questions with regards to ACC function. The predominant thought now being the ACC plays a role in maintaining sustained and effortful behaviours, supported in the context of a hierarchical learning scheme (Botvinick, Niv, & Barto, 2009). Recently, fMRI analysis demonstrated the ACCs involvement in task maintenance and goal orientation within a hypothetical "task space" (Holroyd et al., 2018). With this framework, the ACC may now satisfy its purported role as motivating and sustaining behaviours, but also functioning as a source of reward prediction processing and motor learning (Paus, 2001).

1.3.2 Motor and Associated Cortex

Considering the volition of movement, it is without doubt that an error-evaluative system is necessary to work in concert with other motor regions. M1 is the principle executor of motor movements, with large pyramidal neuron axons descending the corticospinal tract (Porter & Lemon, 1993). Outlined by Vogt, Finch, and Olson (1992), primary motor regions are shown to receive input from cingulate regions of the brain and are involved in skeletomotor contributions. CMA, located deeper into the cingulate sulcus, possess reciprocal connections to M1 and SMA (Devinsky et al., 1995). Further projections into M1 come from a melange of prefrontal, parietal, cerebellar, and basal ganglion circuits. At its most basic form, M1 receives projections from rostral motor-preplanning cortices such as the SMA and presupplementary motor area (preSMA: Haggard, 2008). Both the premotor regions receive input from the basal ganglia (Doya, 2002).

M1 also receives input from a caudal aspect, with involvement from parietal networks projecting into PMC and later into M1. This parietal involvement is implicated in object-action responses, and other object-oriented movements (Haggard, 2008). In addition to execution of movement via pyramidal cell tracts, there is evidence of higher-order learning and response selection in M1. An investigation by Kapogiannis et al. (2008) utilized intracortical inhibition during a gambling task, with consequential cortical augmentation being displayed during large reward presentation. Due to the wide distribution of dopaminergic projections, it is considered that projections from VTA leading into M1 may play a role in the observed cortical augmentations.

1.3.3 Basal Ganglia

Further dopaminergic projections from the VTA are also evident within the basal ganglia. The basal ganglia consist of an aggregation of various nuclei, with its divisions including the striatum, pallidum (globus pallidus externa and interna), subthalamic nuclei, and the substantia nigra (Groenwegen, 2003). Of note, the basal ganglia are connected locally and through projections in the case of the substantia nigra, which resides in the mesencephalon. Further subdivisions of specificity can be applied to the various regions of the basal ganglia; the striatum is composed of three subnuclei: the caudate nucleus, putamen, and nucleus accumbens. As noted, the pallidum consists of an internal (GPi) and external (GPe) segment; the internal portion more medial relative to the externa. Each of these segments contribute their own circuitry to motor control and movement, as well as reward processing. Pertaining to cortico-ganglionic projections, the ventral striatum receives cortical input from the ACC (Cummings, 1993); the ventral striatum consisting of nACC, and ventromedial portions of the CN and putamen (Put: Cummings, 1993). In addition, the ventral pallidum (VP) has been implicated in reward processes (Tachibana & Hikosaka, 2012).

Denoting the importance of the globus pallidus in volitional movement, the GPi and GPe facilitate the production of movement through two distinct circuits: the indirect and direct pathway (Graybiel, 2000). Furthermore, damage to these regions not only pronounce difficulties in motor coordination, but also impair general learning capability. This deficit is evidenced by Knowlton, Mangels, and Squire (1996), in which patients with non-dementia Parkinson's were shown to have a marked inability to perform a reinforcement learning task.

1.4 Neural Systems of Reinforcement and Motor Learning

Connectivity of basal ganglia circuits, reward expectation, and motor production have been exampled within primate ventral pallidum (VP: Tachibana & Hikosaka, 2012). Single cell recording of VP cells during a direction-dependent reward bias saccade demonstrated a scaling of activity. VP activity upscaled or downscaled its firing rate in a tonic manner contingent upon expectancy (reward and time of reward). As dopamine is linked with reward processes, it can be speculated that a tonic increase may prime reward-centric circuitry for the next appropriate saccadic movement.

In another saccade-reward dependent task by Kawagoe et al., (1998), primates were investigated in memory-guided learning, measuring activity of the CN. Within the task, reward was provided only when the cue appeared at four possible locations, such that three locations yielded a non-reward and one with reward. Their following results indicated an increase in caudate modulation, dependent on the presented reward schedule. It was observed that reward modulated caudate neuron activation, with an efficacy that opposed the neuron's preferred direction of activation. Cortical input in conjunction with reward-based dopaminergic pathways are a possible explanation for these responses.

As it is known, deviations from expected outcomes produce neural signals, which later

modify our behaviours for subsequent events. To be specific, the previously mentioned dopaminergic system is distributed to and from many different regions of the brain, including the VP of the basal ganglia (Tachibana & Hirosaka, 2012), dorsal striatum (Balleine et al., 2007; Cromwell & Shultz, 2003; Doya, 2000; Graybiel et al., 1994; Hirosaka et al., 2002; Isomura et al., 2013), medial frontal cortex (MFC: Potts et al., 2006) and VTA (Schultz, 1998; Schultz, Carelli, & Wightman; 2015). Among these regions the striatum seemingly holds a large body of neurons relevant to the expectation process.

To further example this, single cell recordings of Macaque anterior striatum (Cromwell & Schultz, 2003) were observed under conditions of expectancy and variable reward magnitudes (drops of juice from low, medium, and high). As such, this study had demonstrated these cells held a high degree of reward sensitivity within the anterior striatum, as striatal neurons were able to distinguish minute differences in reward magnitudes. It is evident then that select populations of neurons within the basal ganglia (Cromwell & Schultz, 2003) and ACC (Shidara & Richmond, 2002) are sensitive to levels of expectation.

It must be noted that much of the reviewed literature has held a heavy emphasis on cellular recordings. While offering a high degree of spatial and temporal clarity, such methodology is not lucrative to human research as these techniques are invasive and costly. With that we will discuss a valid alternative in the field of electroencephalography to index expectation effects and motor behaviours.

1.5 Electroencephalography and Event-Related Potentials

When a neuron is activated, a difference in electrical charge is present on either side of its cellular membrane, represented as a postsynaptic or presynaptic event. In an individual neuron, detecting these changes is difficult, and requires often invasive methods to record from a single

or small cluster of cells (Kawago et al., 1998). However; a single neuron, despite generating electrical activity of its own, is not an accurate representation of the complexity within the human brain. Thus, we must consider aggregations of neurons, and observe their behaviours. The behaviour of neuron clusters activating together produce an additive effect of their respective charges, such that if they were to activate in response to a presented stimulus or event, they may summate charges dependent on their proximity to one another (Luck & Kappenman, 2011). The summation of neural activation will generate an electrical dipole (see figure 1), known as the equivalent current dipole, of varying magnitude (Luck, 2014). The strength of this dipole is entirely dependent on the number of neurons firing (thousands to millions) and the alignment of each neuron's individual dipole. Should they be oriented in the correct manner, the electrical signal summates (Luck, 2014), and the consequential signal may be captured as voltage by a recording apparatus. By placing an electrode on the scalp and amplifying the signal, we can observe the electrical activity generated by the brain (Berger, 1929). Of course, as the brain is a dynamic and constant entity, the signal generated will appear as lines of random noise with no coherent pattern. However, by increasing the number of electrodes utilized, focusing on experimental design, and how the data is analyzed, otherwise hidden neural processes become apparent. One such method is observing the event-related potential (ERP), where an event stimulus is presented, and activity is recorded in response to this event. By recording the response to the presented stimulus, we can compare activity before its presentation and after its presentation for an indication of any neural activations.

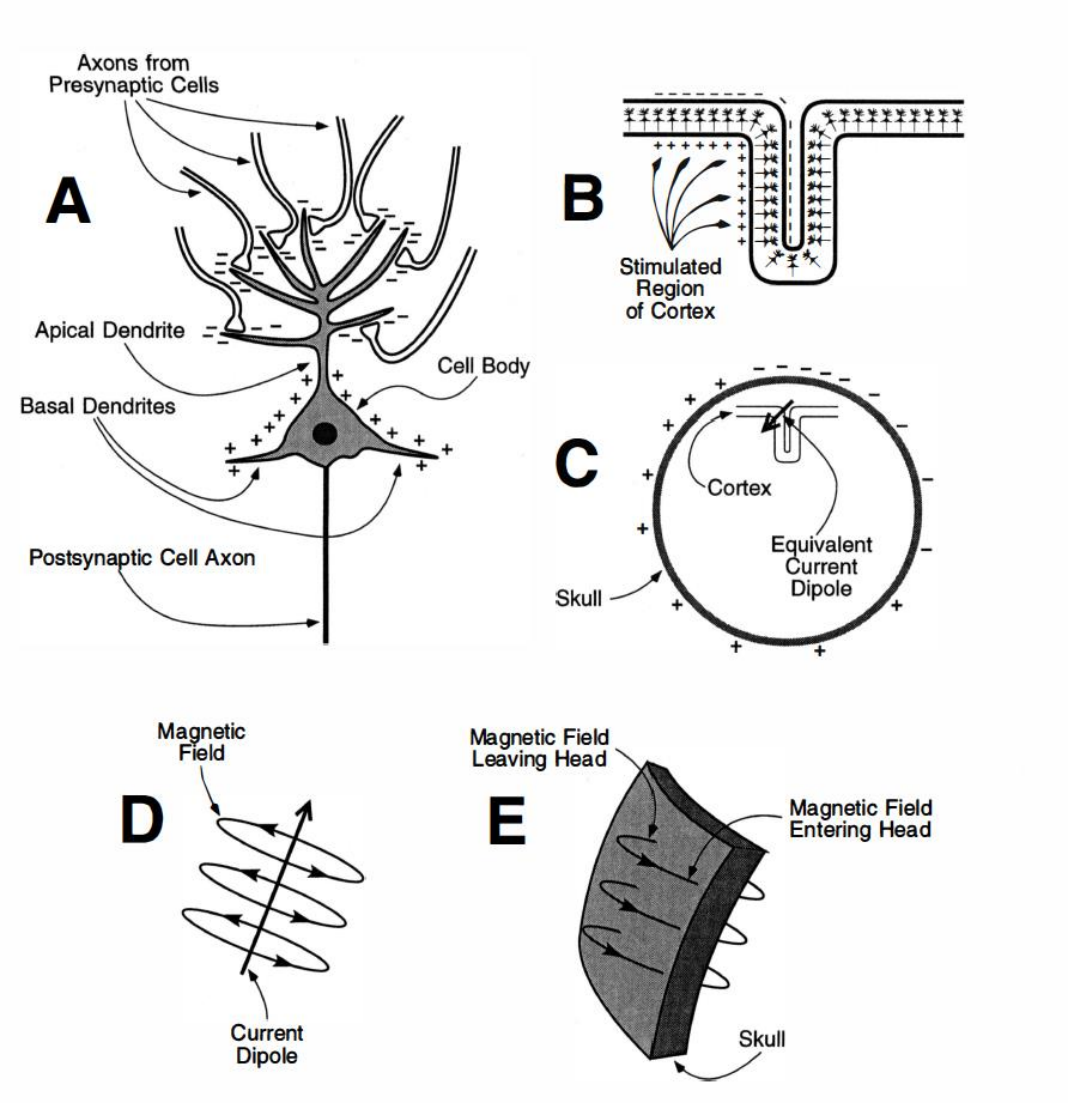

Figure 1. A. Orientation of neuron with direction of polarity during neurotransmission. B. A sample of folded cortex with predominantly pyramidal cells: innervation of this region equates to dipole summation in the orientation of cells. C. The total summation of dipoles across the skull. D. Visualization of the left-hand rule via magnetic field. E. Orientation of magnetic fields produced by dipole summation relative to skull position. Figure adapted from Luck & Girelli, 1998.

These activations are compared to a control or baseline stimulus, such as a fixation point that participants look at in between stimulus presentations. These ERPs are presented as positive or negative deflections, measured in voltages, and contained within each ERP maybe a subset of

components. An ERP component is best described as a characteristic waveform with a distinct latency and polarity, often tied to a specific stimulus, event, or situation.

A key example of a component and its associated stimulus is the reward positivity. Elicited through violations in expectancy, the reward positivity appears when an outcome is different than what is expected to a participant. Imagine for an instance that you are participating in a basic gambling task. You may be asked to select one of two options, one where the chances of winning a small prize is 80%, and another where the chances of winning a large prize is 20%. To keep things simple, we might say the stimulus will be the feedback provided after their response of either a win or a loss. Following the rational of expectancy, by selecting the high-risk, high-gain option you will expect to lose due to the probability difference. However, on the 20% chance of receiving the higher reward, a discrepancy of expectation and result occurs, and thus the reward positivity will occur. To fully appreciate the nuance of this ERP component, we first must look back to its history in the following section.

1.6 Event Related Potentials: The Error Related Negativity (ERN) and Reward Positivity

In 1993, Gehring et al. observed a component seemingly linked to the commission of errors. Participants were asked to respond with their left or right hand based on a modified flanker task (e.g. compatible and incompatible letter arrays: HHHHH, HHSHH, SSSSS, SSHSS: Eriksen & Eriksen, 1974). Focusing on frontal-medial electrode sites (Fz, FCz Cz,), an ERP component was evident whenever participants committed an error during the task. This component, termed the error related negativity (ERN) appeared as a negative deflection, and increased in amplitude when accuracy was emphasized, and an error was made. Conversely, amplitude of this component diminished when response speed was emphasized over task accuracy. Results of this study suggested an innate error-detection system, independent of any

sensory input and localized to frontal-medial regions of the brain. Often found to be time-locked to the participant's response rather than the stimulus itself, the peak amplitudes of the ERN typically occurred between 80ms to 100ms (Gehring, Liu, Orr, & Carp, 2012). However, deviations from this latency have been demonstrated in a series of driving tasks (Krigolson & Holroyd, 2007b; Colino et al. 2017), where participants were asked to complete a simple tracking experiment via steering wheel. Abrupt corners in the driving task were found to elicit an error response in a time range of 100ms to 120ms, a latency longer than the typical ERN.

There is some contention that the reported ERN is in fact another component related to error processing. Following the same frontal-medial electrode positions, this other component is observed on error commission, but strictly based on the presentation of feedback. Termed the feedback related negativity (FRN) or reward positivity (Proudfit, 2015; Krigolson, 2017), this feedback driven component follows the same topography and morphology as the ERN. Of the main differences between the reward positivity and ERN is component latency, with peak values occurring at 230ms to 330ms (Miltner et al., 1997). Rather than being locked on the participant's response like the ERN, instead a positive deflection is observed on feedback given towards their performance. Following the Gehring et al.'s (1993) evaluation of a human error detection system, a wide breadth of research has arose concerning the details of this feedback component. In 1997, a time estimation task conducted by Miltner et al. (1997) demonstrated the reward positivity as being within the same location as the ERN. In addition to a source consistent with previous ERN literature (Gehring et al., 1993), the error signal generated appear to be non-discriminatory to feedback stimuli, such that reward positivity signals were found from auditory, visual, and tactile feedback. Following suit with the ubiquitous error-detection of this

component, Holroyd et al. (1998) demonstrated that error commission from hands or feet had no difference in the activity of the ERN.

1.6.1 Reward Positivity: Neural Basis of Human Error Processing

Of considerable relevance is the seminal paper produced by Holroyd and Coles (2002). In their comprehensive review, a theoretical framework was proposed to establish the connection between the ERN, reward positivity, and mesencephalic dopaminergic reward systems. More specifically, it was suggested that the ERN and reward positivity serve as a prediction error by way of the ACC. To validate their hypothesis, two EEG experiments were conducted, the first of which involved a probabilistic learning task. The main purpose of this task was to investigate both the ERN and reward positivity within the same context, and such parse out any meaningful relationship between the two components. Participants were presented with a two-choice gamble in which responses were done in a left-key, right-key input fashion. Stimulus mappings were presented in six possible outcomes, dependent on several variables. The first stimulus was mapped to left button inputs, and thus rewarded participants 100% of the time; right button inputs were met with a penalty (100% left mapping). A flipped right-button mapping constituted the second stimulus, in which rewards were mapped to right-button inputs, with penalty on left-button presses (100% right mapping). For each left and right button, another pair of stimulus were presented, in which the stimulus rewarded participants irrespective of button input in 50% of responses. Likewise, a similar mapping followed suite with a penalty in 50% of responses (50% mappings). Lastly, two stimuli corresponded to rewards and penalty offered 100% of the time, independent of button response (always correct, always incorrect).

Results demonstrated two distinct behaviours in ERN and reward positivity amplitudes, such that as participants began learning the 100% Mapping conditions, a corresponding decrease

in reward positivity amplitude followed. Incidentally, response locked ERNs remained large in amplitudes, suggesting a shift from a feedback-mediated system to an internal action-error monitoring system. In 50% Mapping conditions however, reward positivity amplitudes were predominant with little to no response-locked activity present. Amplitudes of always correct and incorrect conditions also decreased, demonstrating the condition had no meaningful evaluative properties, and thus less neural activity.

The second experiment (Holroyd & Coles, 2002) utilized a modified Flanker Task, with the center array following the standard S and H pattern. Consistent H or S lines were considered as compatible; however, a manipulation of stimulus frequency was also introduced. That is, half of participants saw a stimulus with a central H on 10% of trials and with a central S on 40% of trials. The remaining half were shown the opposite of these weightings. By doing so, four stimulus conditions were created: infrequent compatible, infrequent incompatible, frequent incompatible, frequent compatible.

Despite their original assumption that amplitudes of the ERN would be lessened on error trials locked to incompatible conditions over compatible conditions, results demonstrated an increased ERN amplitude within frequent incompatible trials. This was at odds with a comparatively smaller infrequent compatible and infrequent incompatible trials, and apparent variability in the ERN amplitude was connected to participant's overall accuracy ratings. The more accurate a participant was, the larger the ERN was during error commission. Characteristics of the ERN now assume traits found within prior temporal difference learning models (Sutton & Barto, 1998), and thus are indicative of a prediction error. By attributing the reward positivity towards a prediction error standpoint, and localizing it to the ACC, the reward

positivity now reflects a function of reward expectancy processes, and reinforcement learning. Studies of this component reflecting reinforcement learning models are many and varied.

1.6.2 Reward Positivity: Reinforcement Learning and Expectancy

A review by Walsh & Anderson (2012) summarized the growing body of evidence that the reward positivity act as a function of reinforcement learning, based off four key features. Of the first, the reward positivity presents itself as a quantitative prediction error (Holroyd & Coles, 2002), secondly; the reward positivity is highly responsive to rewards and the predictor of rewards, a point analogous to phasic dopamine release (Schultz, 1998); third, reward positivity amplitudes and consequential behaviours modulate with experience, and fourth, this system appears to be best engaged during action and awareness. These four characteristics belie the appropriateness of the reward positivity to catalog reward prediction errors, to which a further examination must be taken.

In a modified time-estimation experiment (Miltner et al., 1997), Holroyd and Krigolson (2007) introduced two additional levels of difficulty to the task design. Much like the base study, participants were expected to estimate the duration of one second after an auditory cue and provide a response via key input. Feedback was provided in the form of a ✓or x based on a correct or incorrect response, respectively. Unknown to the participants, their performance in the task manipulated the window of response, such that it was often that a second was not actually a second. Instead, on successes the window of one second decreased in response time, whereas failures widened this window. Under normal conditions, the rate of increase or decrease was equivalent such that participants were correct 50% of trials. Conversely, in the additional blocks, participants were introduced to success rates of 25% (hard difficulty) or 75% (easy difficulty). Much in line with reinforcement models (i.e. Rescorla & Wagner, 1972), a linear trend in reward

positivity amplitude and task difficulty was seen. Under hard conditions, reward positivity amplitudes were larger and scaled downwards with less difficult estimates, such as the easy condition.

Although many studies appear consistent with their findings, it is much in the nature of science to provide confounding results. Regarding the scaling of expectancy, many models of reinforcement learning predict a linear trend with respect to prediction errors and the reward positivity, however this may not entirely be true. Williams et al. (2017) provided evidence of a different function all together. Utilizing a time estimation task derived from Holroyd and Krigolson (2007), an additional two levels of difficulty were introduced to participants. As such, participants were afforded experimental blocks where their expected success rates fell into the following percentages: 10, 25, 50, 75, 90. Results demonstrated a very similar trend across the the core difficulties (25, 50, 75) towards Holroyd and Krigolson's (2007), however once expectancies fell towards the extreme ends of success rates an interesting trend became apparent. Rather than continue as theoretical models would predict, linearly, results demonstrated a sigmoidal function, or a nonlinear trend such that there was no difference between the upper-end and lower-end difficulties (10% - 25%, 75% - 90% successes). Curiously this does not invalidate any reinforcement models, per se, but instead elucidate how these functions occur within a biological system. Indeed, there is speculation that such nonlinear trends are in fact represented by already existing theories and observations, the most telling being the dopaminergic reward system (Fiorillo, Tobler, & Schultz., 2003; Schultz, 2016; Schultz et al., 2015; Stauffer et al., 2014; Stauffer et al., 2016).

1.7 Motor Control and Learning

With regards to motor learning, several considerations may be taken towards the interaction of reinforcement learning and motor control. Behaviourally, reward was also shown to accelerate motor learning (Nikooyan & Ahmed, 2015). In a visuomotor task, participants were asked to rapidly move a cursor to a target via robotic arm. Participants experienced an abrupt rotational perturbation and were required to learn compensation techniques for this interruption. To facilitate learning, participants were provided with either feedback or no visual feedback based on their attempts. In non-visual feedback groups, they were instead offered reward feedback in three manners: no feedback, linear feedback, and cubic feedback; these dimensions influenced a trial score out of 1000, which was shown to participants. Participants were asked to maximize their trial scores. Of the group conditions, participants only receiving reward feedback (and no-visual feedback) were still able to learn the task adeptly; however, the combination of both visual feedback and reward feedback greatly accelerated motor learning rates.

However, contrary to the findings of Nikooyan & Ahmed (2015), negative feedback provided an accelerated rate of learning in another visuomotor task (Galea et al., 2015). Their results suggested motor adaptation and skill retention are affected by reward and punishment; rather, negative feedback (with or without monetary loss) accelerated learning rates, whereas positive feedback (reward) influenced memory retention.

Lastly, a study by Izawa and Shadmehr (2011) investigated the distinction of two prediction-error models that occur during motor movement. The first is dependent on the sensory consequences of our actions relative to the expected outcome, and the other is based on previously outlined reward prediction processes. The classical thought assumed motor

adaptations occur on a sensory-only basis; however, it is now seen that adaptations take advantage of both sensory and reward prediction errors.

1.7.1 Electroencephalography and Motor Learning

To further assess motor control and learning, investigation into underlying cortical processes must be done. As it has been established, neuro-correlates of learning can be indexed by the reward positivity (Holroyd & Coles, 2002, Holroyd & Krigolson, 2007, Williams et al., 2017); however, additional studies have investigated prediction errors in a motor context. For example, a study completed by Holroyd and Krigolson (2006), points towards a hierarchical system to mediate levels of error during a motor task. Participants were asked to play a computerized tracking task in which a joystick was used to maintain a cursor within two barriers. As trials progressed, barriers flanking either side of the screen and shifted in a predictable manner from left to right. Any contact with these barriers constituted a tracking error. The inter-stimulus intervals were shown as straightaway sections where barriers did not move, and overall successful performance for the task was the maintenance of the cursor to the center of the screen. Within 20% of these straightaway sections, an event termed a ‘difficult corner’ occurred, such that the barrier shifted rapidly, and an inevitable hit would occur. Within these difficult corners 50% were ‘locked’, allowing the computer program to take control of the cursor and thus preventing contact with the barrier. Collision with the barriers elicited a response consistent with prior literature, demonstrating an error evaluation within frontal-medial regions of the brain as the ERN or reward positivity, and reflects a higher-level of error processing. Another distinct result of the study involved investigation of parietal components, suggesting a low-level error processing system working in concert to ensure consistency and accuracy in movements.

To further evidence motor control being mediated through high and low-level systems, an expansion of the prior study was conducted, utilizing a similar continuous tracking task (Krigolson & Holroyd, 2007a). To examine if predictive cues can influence error processing, participants were again asked to perform a continuous tracking task. Like its predecessor (Krigolson & Holroyd, 2006), barriers would move, and unexpected error trials would occur. A distinction from the previous study is within both the frequency of the rapid barriers (from 20% to 40%), and the types presented to participants. A total of four possible variants were available, dependent on similar conditions to the previous study (locked and unlocked), with two novel stimulus probing predictive cues (predicted and unpredicted).

Unpredicted trials functioned as the original study, such that only locked or unlocked corners would occur. Within predictive cue trials, coloured signs would appear above the barriers 500ms prior to perturbation. These cues (either green or blue) served as an indication as to what type of trial was about to occur (locked or unlocked); however, no prior information was given to participants as to what each cue meant and thus had to be learned. A further examination in predictive cues examined internal models of motor commands, such that in lieu of sudden shifts in barriers, participants experienced a significant decrease in joystick responsiveness, resulting in an inevitable error with the barrier. As above, decreases in joystick sensitivity were met as predicted or unpredicted, reliant upon the participant's ability to learn and distinguish the cues. External predictive information (visual) and predictive error comparisons between internal motor command and related behaviour elicit ERN-like waveform; however, more was left to be assessed about lower level evaluative systems.

Expanding upon this gap and shifting from continuous driving tasks, a later study (Krigolson & Holroyd, 2007b) further examined the differences of high and low-level error in a

motor context. In a cursor movement task, participants were asked to move a cursor horizontally to a target from a starting position. In this, participants were given three possible conditions, the control being an unmanipulated movement of the cursor. The following two conditions involved a manipulation of the target location in that as participants committed a movement the target disappeared and appeared simultaneously to a different location. In some instances, participants were able to detect and correct for this target perturbation; however, in another instance participants were unable to correct their movement as response to the cursor was prevented. It is of importance to note participants were naïve to which trial they were in. It was speculated that correctable responses would present as a function of lower-level, parietal processing involved with online control (Woodworth, 1899; Goodale et al., 1986). Conversely, uncorrectable errors would represent a high-level function, such that an expected event has not performed as intended and thus a prediction error would be generated (Holroyd & Coles, 2002). Indeed, it appears that the inability to correct for movements did result in the reward positivity, as reflected by the generation of a prediction error. Conversely it is mentioned that the ERN was not produced during the task; this by the logic of prediction errors makes sense, as the ERN is thought to be generated via generation of efference copies of a given motor command and not dependent on external feedback or expectancy.

Further extrapolating on higher level errors, a following study by Krigolson, Holroyd, Van Gyn, and Heath (2008) examined the neural correlates of targets and outcomes. Utilizing a mirrored apparatus (Held & Gottlieb, 1985), participants performed an aiming task via stylus and WACOM tablet. Trials started in the bottom-left corner, cuing participants to move to a presented target on an auditory cue in a left-right fashion. Three possible conditions were provided to participants in a similar fashion to Krigolson & Holroyd (2007b); a control condition

with no target perturbation, a correct condition in which the target shifts to a different location on movement onset, and a blocked condition in which the cursor suddenly became locked to the horizontal axis with no means of vertical axis correction (always missed). Results demonstrated several supporting and contrary results: the prevention of corrective movement (blocked trials) did not elicit error evaluative systems; however, it appears the medial-frontal system is dependent on endpoint outcomes, and thus remains consistent with hierarchical control processes.

1.7.2 Motor Potentials

In addition to the reward positivity, there is evidence of other frontal-medial component modulation during movement tasks. In a memory-guided reaching task (Krigolson et al., 2012), participants were positioned in a setup as established in Held & Gottlieb (1958). Utilizing a WACOM tablet for input, participant arms were occluded with the cursor position as an indication of their limb relative in space. Two target positions were afforded to participants (proximal and distal) with two further conditions (visual and memory). Under visual guided conditions, participants were presented either a proximal or distal target and could move to the target after a tone; conversely, under memory guided conditions, vision of the target was occluded after a preview period, and thus participants had to approximate the location of the target.

Several corresponding components maximal at frontal-medial sites were observed and were speculated to be produced by the SMA (Shibasaki, Barrett, Halliday & Halliday, 1980; Shibasaki & Hallett, 2006). Overall, a recorded decrease in motor potential amplitude was seen in memory guided reaches that correlated with a decrease in peak acceleration and an increase in undershooting the target. Attribution to error may be linked to a reduced activity of SMA and

associated premotor regions (Shibasaki & Hallett, 2006), and thus influencing force production, affecting kinematic profiles.

1.7.3 Bereitschaftspotential

Of further interest are other associated motor components: the bereitschaftspotential (BP). First recorded in 1965 (Kornhuber & Deecke), the properties of this component were seen preceding motor related activity. Occurring as an initial negative drift with a maximal peak shortly before movement onset, a greater frontal activation site than occipital was observed, with a bilateral spread over midline electrodes. The amplitude of this component was also seen to diminish when participants either became 'indifferent' to the target, or through disinterest of task (i.e. excess repetition of a simple task leading to boredom).

Indeed, the BP is a movement related cortical potential. An overview by Shibasaki and Hallett (2006) established many key factors of this component. As we begin to produce an intended movement, there is a gradually increasing potential that typically occurs two seconds before movement onset. This graded potential is found to maximize at mid-centro-parietal regions of the brain in a symmetrical manner and is widely distributed. As this gradient progress, there is a sharp increase of the BP approximately 400ms before movement commission, with the late BP maximal on the contralateral hand or limb of movement. Amplitude of the late component of the BP is also seen to be modulated through preplanned or spontaneous movements, such that spontaneous movements may only present within 500ms of movement onset. The BP produced from such movements are also smaller in amplitude relative to preplanned movements (Libet, Gleason, Wright, & Pearl, 1983).

In addition to component attributes, source localization methods of the BP point towards many motor-related cortical regions. Those involved with its formation include: M1, primary

somatosensory cortex (S1), PMC, and SMA: a majority implicating a frontal-medial origin for the BP (Ball et al., 1999; Brunia, Boxtel, & Böcker, 2012). Involvement of the SMA is still a contested subject; however, a study utilizing subdural electrodes, within the interhemispheric fissure of epileptic participants demonstrated a purportedly clear BP produced by the SMA within a range of many different movement types (i.e. foot movement, hand movement, and tongue extension: Ikdea, Luders, Burgess, & Shibasaki, 1992). Additional intracranial research also implicated the cingulate cortex in the contribution of BP-like components as well, such that the anterior caudal aspect of the gyrus, in addition to the anterior and posterior cingulate produced signals with some varied consistency (Rektor et al., 1998).

According to Taylor (1978), research concerning the BP utilized abrupt and simple movements with no learning requirement. In the corresponding study, participants were asked to press a series of six buttons every 20 seconds with no error. Focusing on electrodes Fz, Cz, C3, and C4, the BP was found to increase at all sites as participant performance increased over time (i.e. a reduction in reaction time). There was a consistent trend found between performance improvement and change in the BP, suggesting a learning response regarding the BP and task proficiency. Prior to Taylor (1987), a study conducted by McAdam and Seales (1969) observed BP amplitudes under a rewarding condition relative to baseline. Under control conditions, participants were asked to click a button periodically and to refrain from any counting or tracking systems. Identical to the control, participants under the reward condition were offered a monetary gain for each correct click. Feedback was provided on correct clicks in the form of an auditory cue and withheld for incorrect responses. Analysis on BP and condition revealed a significant effect of reward and BP amplitude on the contralateral side to the clicking hand. Similarly, improvements of BP may be observed through similar neurofeedback paradigms. A

study by Fumuro et al., (2013) demonstrated an increased BP in patients with Parkinson's disorder, and the first to demonstrate self-generated improvement without aid of pharmaceuticals or surgery. Given the involvement of the overlapping cortical regions of motor control and reward, it would appear sensible to examine the interaction of this component in relation to an expectancy task. As prior literature has established its effect on reward (McAdam & Seales, 1969), and biofeedback conditions (Fumuro et al., 2013), the interaction of the BP and reward positivity may elucidate a pattern in how these regions may facilitate motor learning and execution.

1.8 Conclusions

Considering the prior literature, the purpose of the current thesis is to examine whether reinforcement learning systems within the medial-frontal cortex play a role in motor learning. More specifically, I aim to demonstrate that violations of expectancy scale within conditions in a motor context. To accomplish this, a motor expectancy task was developed upon a modified time estimation paradigm, as completed by Holroyd and Krigolson (2007) and Williams et al. (2017). It is predicted that the amplitude of the reward positivity would be smallest when the difference between expected and actual outcomes are smallest. An example of such would be hitting a target when expecting to hit a target, and conversely, missing a target when expecting to miss. Furthermore, I also predict the reward positivity to be maximal during unpredicted events, such as expecting to miss a target and hitting instead.

Due to its role in preceding motor movements and modulation during reward-type conditions, I will also investigate the bereitschaftspotential (BP: McAdam & Seales, 1969; Shibasaki & Hallett, 2006). It is also predicted that the BP will increase in amplitude in a negative direction under more rewarding conditions, such as expected outcomes. Expected

results of the reward positivity fall in line with prior literature (Holroyd & Coles, 2002; Holroyd & Krigolson, 2007; Williams et al., 2017), and coupled with the BP I aim to demonstrate a connection between expectation effects and motor learning.

CHAPTER 2: MOTOR EXPECTANCY: THE MODULATION OF THE REWARD POSITIVITY IN A REINFORCEMENT LEARNING MOTOR TASK

2.1 Introduction

In our day-to-day lives, the assessment of our actions is key to learning: we tend to avoid errors and repetition of previous errors. However, the notion of learning from our 'mistakes' is a misnomer; rather, we learn when an outcome differs from our expectations (Rescorla & Wagner, 1972; Sutton, 1988; Sutton & Barto, 1990). The discrepancy between actions and outcomes is known as a prediction error and serves as a function of expectancy. Succinctly, a small prediction error would be generated if a committed action (i.e. throwing a dart) meets an expectation (i.e. hitting a bullseye). In contrast, should an action-outcome fail to meet expectation (i.e. hitting the outer ring rather than the intended bullseye), a larger prediction error would be generated. Consequently, the next following dart throw may be joined with a behavioural correction, such as aiming higher or throwing with more force to account for previous error.

These prediction errors may be categorically measured through use of electroencephalography (EEG). By indexing human event related potentials (ERPs), we can observe patterns of activation centered around times in which expectancies would be violated. Indeed, this is the case, as shown by the reward positivity (Proudfit, 2015; Krigolson, 2017), a frontal-medial component associated with the anterior cingulate cortex (ACC: Paus, 2001). This is most exampled through a seminal study conducted by Holroyd and Coles (2002), as the component was shown to model after phasic dopamine release of the ventral tegmental area (VTA: Schultz et al., 1997). Pertinent to the reward positivity is its apparent sensitivity to expectancy effects through prediction errors. The ubiquity in error detection of this system has

been demonstrated across several modalities. For example, prediction errors measured by the reward positivity were observed across auditory, visual, and tactile stimuli (Miltner et al., 1997). Furthermore, it appears this system does not discriminate in the way errors are committed, such that motor errors commissioned by the hands or feet presented no discernable difference in ERN activity (Holroyd, Dien, & Coles, 1998).

It is evident then that the reward positivity serves as a measure of general prediction errors but is also sensitive to variable differences in expectancy. Such is demonstrated through modified time estimation tasks conducted by Holroyd and Krigolson (2007) and Williams et al. (2017), to which participant expectancies were manipulated across conditions. It was posited that when expectations were met (i.e. easy blocks), reward positivity amplitudes would be decreased; conversely, when expectations were not met (i.e. hard blocks), reward positivity amplitudes would be increased. Indeed, when a method of difference subtraction was applied, reward positivity amplitudes were shown to scale in a linear fashion in accordance to perceived expectancy.

Despite this demonstrable scaling in reward positivity in expectancies, there is little examining its interaction in a motor context. Prior literature has established that motor-oriented tasks involve prediction error processing, shown through a series of motor-related tasks (Krigolson & Holroyd, 2006; Krisolson & Holroyd, 2007; Krigolson et al., 2012). It is presumed that, in a motor sense, corrective behaviours are mediated through 'hierarchical' levels of error detection. Specifically, it is thought that high-level processing constitutes activity of the reward positivity, whereas lower-level errors are mediated through components associated with ballistic control. Given the niche between expectancy and motor control, the current study involves a motor expectancy task. Participants performed overt motor movements (drawing a line to a

target), to which condition difficulty was manipulated to tailor expectancy. These manipulations were analogous to prior expectancy literature (Holroyd & Krigolson, 2007; Williams et al., 2017) such that three expectancies were established (expected, control, unexpected). It was hypothesized that the reward positivity would scale with respect to perceived expectancy, and that expected outcomes would produce the smallest amplitudes. Conversely, we hypothesized that the largest amplitudes would be the product of unexpected events.

Furthermore, it was speculated that expectation effects would modulate the bereitschaftspotential, such that component amplitude would be largest during conditions of perceived reward (McAdam & Seales, 1969). We further hypothesized that participant hit frequency and movement times will also decrease during unexpected conditions, as accompanied with a smaller bereitschaftspotential (Perri, Berchicci, Lucci, Spinelli, & Russo, 2016). To this end, we expect to observe a similar pattern of results as established within prior literature, translated towards a motor-movement task.

2.2 Methods

2.2.1 Participants

Twenty-five undergraduate students were recruited in the following study. Five participants were removed due to not learning the task sufficiently. The data of three participants were removed from post-experiment analyses − due to an excessive number of artifacts in the EEG data. One participant presented with an inverted bereitschaftspotential and was removed from the dataset. Of the remaining participants, 17 were used (females = 11, mean age = 21.17, [22.68, 19.65]).

Recruitment was facilitated through University of Victoria's online experimental sign up system. Participants were reimbursed through course credit in select psychology courses. All

participants had normal-to-corrective vision, and no known psychological or psychiatric illness. Prior to the experiment, all participants completed informed consent forms in addition to demographic questionnaires consisting of questions regarding age, sex, and handedness. Experimental procedures were approved by the Human Research Ethics Board at the University of Victoria before recruitment and data collection, and followed the ethical standards outlined in the 1964 Declaration of Helsinki.

2.2.2 Procedure

Participants were seated comfortably in front of a 22'' ViewSonic VX2252MH LCD (1920 x 1080 pixels) computer monitor in which the experimental procedure was displayed. Inputs were done via a WACOM Intuos 4: PTK 1240 tablet (Wacom Co., Ltd) and stylus pen. Participants were instructed to keep the tip of the stylus in contact with the tablet always during the task. Overlaid on participant hands was a screen, blocking vision of both hands and input tablet (see figure 2 for reference).

Prior to the main experiment, participants were asked to complete a practice version of the task to familiarize themselves with the procedure. Practice sessions consisted of 40 trials. Each trial of the practice task involved a target appearing at four possible locations, tied to each corner of the screen. This target indicated the corner in which participants were to draw a line from the center position to the center of a hidden 'target square' behind the target cue. Participants were informed of this target square but were given no information of its dimensions or properties (see figure 3 for reference). Upon ending their movement, feedback was provided to participants in the form of a ✓ (hit) or x (miss). The hidden square increased or decreased in size at an equivalent rate 25% of its size within a current trial, dependent on their performance.

For example, if a square was 100x100 pixels, a hit would generate a square that is 75x75 pixels for the next trial.

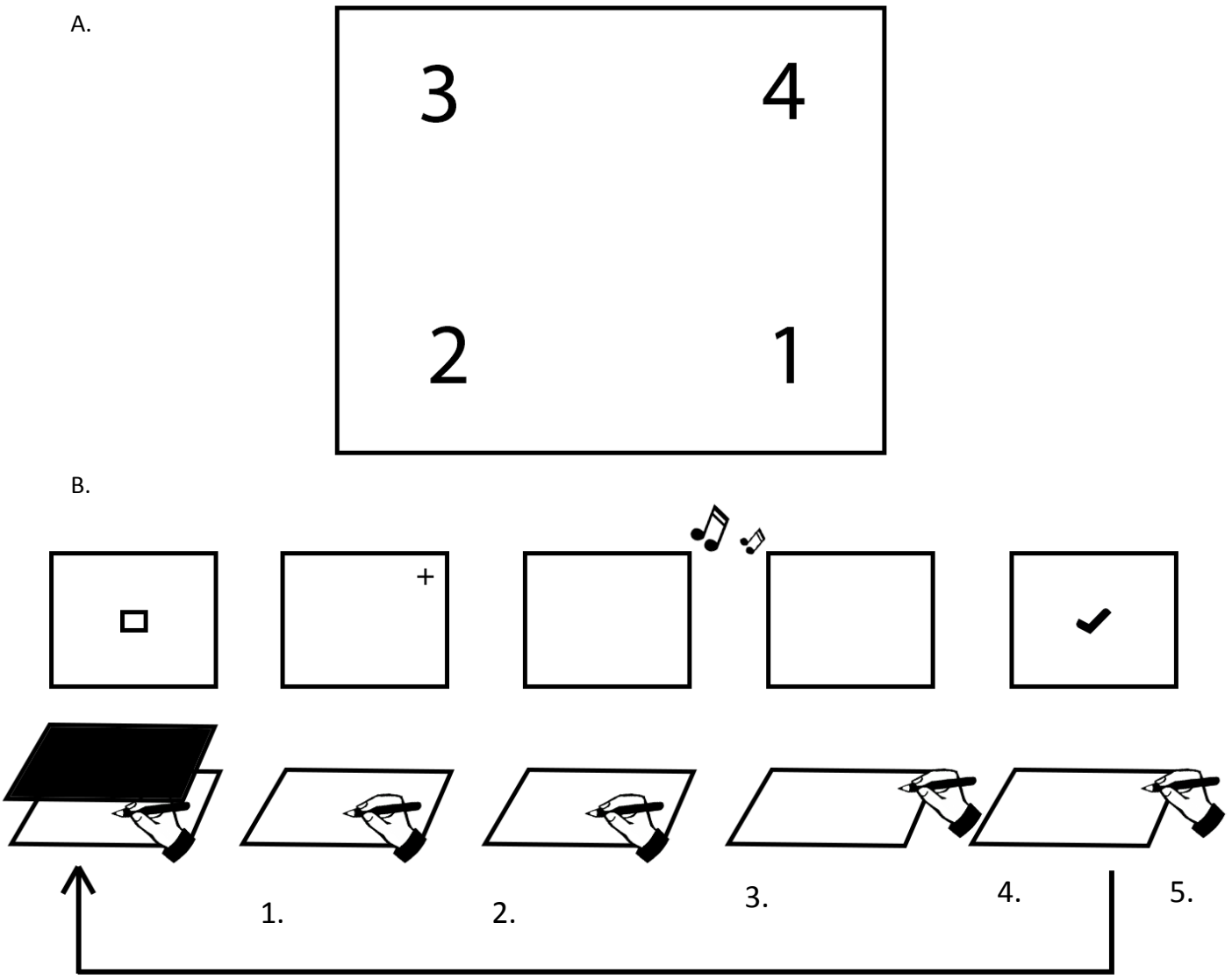

Figure 2. A visual depiction of the motor-expectancy paradigm. (A) Describes the stimulus quadrants and their ordering for later analysis. Numbers listed in pixels: (1) 1440-270, (2) 480-270, (3) 480-810, (4) 1440 – 810. Note: numbers are not associated with any stimulus presented to the participant. (B) Black square depicts a screen obscuring the participant hands. (1) Participants begin each trial in the center of the home target. (2) A Target appears in any quadrant randomly. (3) Participant waits for auditory cue to move. (4) Participant moves to target. (5) Feedback on performance is provided.

Conversely, a miss would generate a square that was 125x125 pixels for the following trial. This rate is identical to the control condition within the final experimental run. Participants did not continue onto the main experiment unless they achieved approximately 20 out of 40 trials, or 50% success rate. The experiment took place in a lit and sound-dampened room within the Neuroeconomics Laboratory at the University of Victoria.

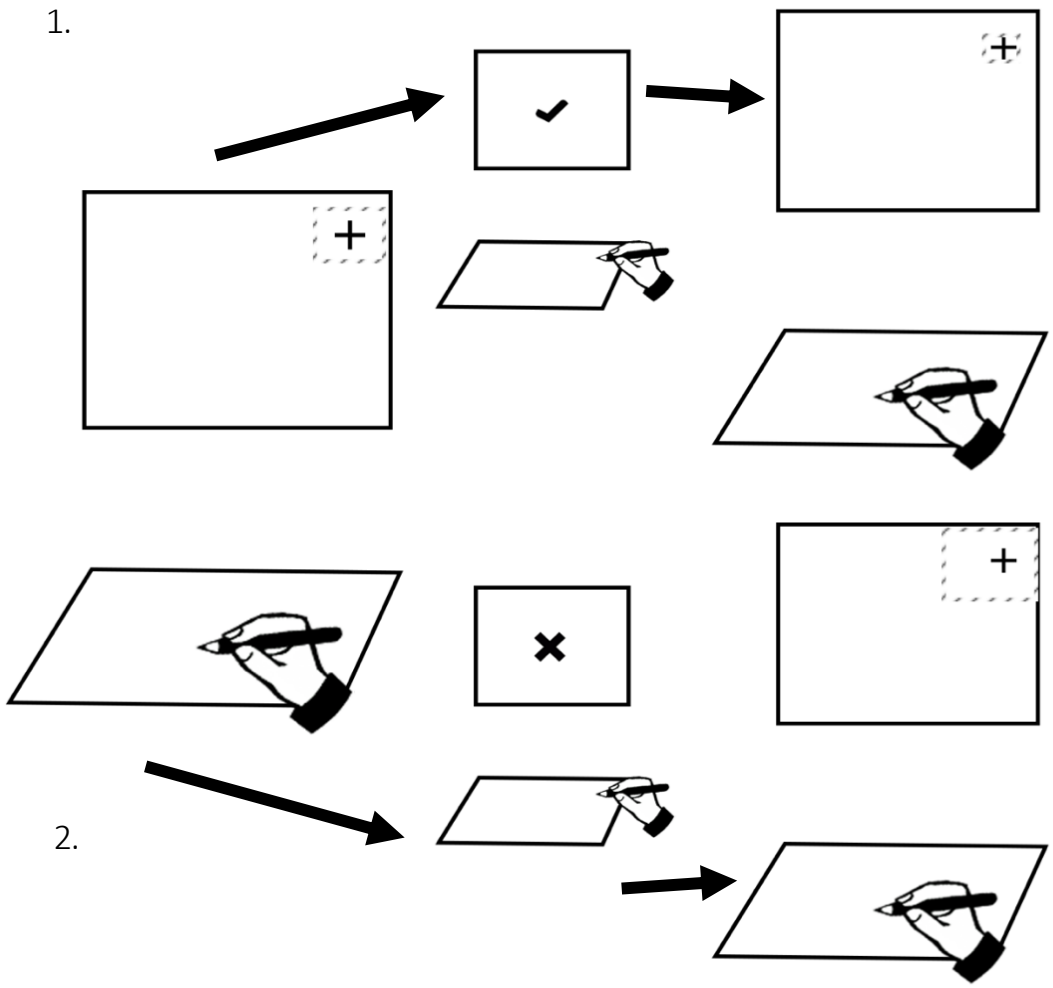

Figure 3. Visual depiction of target square (hatched lines) rate of change. Target square was not visible to participants; however, participants were informed of its presence. (1) Upon successful trials (Hits), area of target square decreases on the following trial. (2) Upon unsuccessful trials (Misses), area of target square increases on the following trial. Rate of change dependent on block. Depicted squares not representative of actual position or size change.

2.2.3 Experimental Task

The experiment was coded in MATLAB (Version 9.1.0, Mathworks, Natick, USA), utilizing the Psychophysics Toolbox extension (Brainard, 1997). In the present task, participants were asked to draw a line to a cued target location, dependent on a randomly prompted corner of the screen. Invisible to the participant was an additional 'target square' in which participants were asked to approximate the center of it, using the original target cue as reference for quadrant and location.

To explain the task, participants were provided written and verbal instruction prior to starting, in addition to the practice trials. Participants were made aware of the length of the experiment, in addition to block and trial numbers; however, participants were naïve to which block they were in. Block design adhered to a similar design as Holroyd & Krigolson (2007) and Williams et al. (2017) with the following discrepancies: like Williams et al. 2017, all conditions were counterbalanced across participants and provided an initial training phase; however, like Holroyd and Krigolson (2007) only three block conditions types were presented to participants. Each block consisted of 75 trials, and there were two blocks of each type to a total of six blocks, 450 trials. Blocks were counterbalanced to ensure randomness. A summary of block design and behavioural outcomes may be seen in Table 1.

With each block, the 'target square' changed its area at a variable rate, dependent on participant success. Initial target square starting size was 150x150 pixels. In the control block, rate of change was dependent on the following: hits decreased target square size by 25% (i.e.: 100x100 pixels decreases to 75x75 pixels on the next trial). Misses increased target square size by 25% (i.e.: 100x100 pixels increases to 125x125 pixels on the next trial). In the expected correct block, the rate of change was dependent on the following: hits decreased target square size by 7% (i.e.

100x100 pixels decreases to 93x93 pixels on the next trial), and misses increased target square size by 100% (i.e. 100x100 pixels increases to 200x200 pixels). In the unexpected block, the rate of change was dependent on the following: hits decreased target square size by 60% (i.e. 100x100 pixels decreases to 40x40 pixels), and misses increased target square size by 40% (i.e. 100x100 pixels increases to 140x140 pixels). By establishing these variable rates of change, overall success rates were able to be modulated, and thus a manipulation of expectancy.

Table *1. Table of manipulations according to condition with 95% confidence intervals. Area of target is listed as the percentage of change in area of the target square. Probabilities describe the percent hits or misses per block condition.*

Condition	Area of Target: Correct	Area of Target: Incorrect	Correct Probability	Incorrect Probability	Correct Actual	Incorrect Actual
Expected	-7%	+100%	~75%	~25%	66%: [59, 72]	34% [28 40]
Control	-25%	+25%	~50%	~50%	44% [43, 45]	56% [54, 57]
Unexpected	-60%	+40%	~25%	~75%	26% [26, 27]	74% [72, 74]

At experiment onset, participants were assigned a randomly selected block difficulty. At the start of each trial, a ‘home target’ square appeared at the center of the screen (100x100 pixels). Participant cursors were invisible everywhere except within the bounds of the home target. Placement of the cursor within the ‘home target’ square indicated readiness for the next trial, and participants were asked to place their cursor within the center of the square. To

facilitate consistency, a small crosshair was placed in the center of the square as a visual guide to the home target's true center.

While inside the home target cursor movement was visible with a residual tracer (i.e. a small trail of fading circles) to indicate participant cursor position for orientation purposes. Upon centering the cursor, the trial would begin, and the home target square was removed, presenting a target cue at a random corner of the screen. Target quadrant locations adhered to the following x-y coordinates in pixels: (1) 1440-270, (2) 480-270, (3) 480-810, (4) 1440 – 810 (quadrant distribution is outlined in Figure 2). Presented targets appeared on screen for 900 – 1100 ms. The targets then disappeared for a 400 – 600ms period, followed by an auditory cue indicating that the participant should make their aiming movement to the target region. If a participant left the home target prior to the auditory cue, they were informed that they had left the target too soon, and the trial ended. If participants left the target at the appropriate time (i.e. immediately after the auditory cue), they were provided feedback based on their performance. Like the practice session feedback was provided to participants in the form of a ✓ (hit) or x (miss). Feedback delay was 400 – 600ms and presented on screen for 750ms. Trials were self-paced, in which the next trial did not begin until participants returned to the center of the home target.

Prior to data collection, experimental variables were calibrated and adjusted according to the behavioural data of six pilot participants through trial and error. Adjustments to the rate of change in target squares yielded the final experimental manipulation.

2.2.4 Data Acquisition

Electroencephalographic data were recorded in a standard 10-20 layout, distributed over 64 electrode sites (ActiCap, Brainproducts, GmbH, Munich Germany). Electrodes were referenced to a common ground, with electrode impedances kept below a threshold of 20 kΩ to

ensure clarity of signal. Data sampling range was at 500Hz, amplified, and filtered through an antialiasing low-pass filter of 8kHz. A DATAPixx 2 stimulus unit (VPixx, Vision Science Solutions, Quebec, Canada) was used to ensure temporal coincidence of event-markets with experimental stimuli.

2.2.5 Data Processing

Data were processed offline with Brain Vision Analyzer 2 software (Version 2.1.1, Brain Products, GmbH, Munich, Germany) following standard methods, as per the Neuroeconomics laboratory (http://www.neuroeconlab.com/data-analysis.html). First, sufficiently noisy or faulty electrodes were removed. EEG data were then re-referenced to an average of the mastoid process electrodes (TP9 and TP10). Data were next filtered via a dual-pass Butterworth filter, with passband of 0.1 Hz to 30Hz. A 60 Hz notch filter was utilized to filter ambient electrical noise. Data were then segmented into 3000ms epochs (-1000ms to 2000ms range) based from block condition (expected correct, control, unexpected correct). Contained within each condition epoch was all markers of interest for a given condition. Segmented data was then subject to an independent component analysis (ICA). Components related to ocular artifacts were removed, and data were reconstructed. Following ICA, removed channels were re-introduced via interpolation, utilizing a spherical splines method.

A baseline correction was then applied, dependent on the component of interest. For the reward positivity, a baseline of -200ms to 0 was utilized, whereupon data were then further segmented based off its condition and marker. For instance, under the expected correct condition, data would form an epoch of -200ms to 600ms for each marker denoting a hit, and a separate epoch would be formed for markers denoting a miss. These markers were quantified based off a participant successfully hitting the target or not. This process was repeated for each block

condition (expected correct, control, unexpected correct). A different baseline correction was applied to the BP, ranging from -500ms to -300ms. This window was chosen as the late BP was decidedly the most suited component of analysis. Prior literature lists the late BP as a range -500ms prior to movement onset, and is associated with task precision, discreteness, and complexity; many of which are relevant factors in the current task (Shibasaki & Hallett, 2006).

Epochs were then formed from the data encompassing a -500ms to 500ms window, centered around a marker involving the detection of participant movement onset. This process was repeated for each block condition as well. An artifact rejection algorithm was then used to reject any segments of data with gradients 10 μV/ms or with a 100 μV absolute difference within the segment (82% [76%, 87%] segments kept). ERP waveforms for each condition and marker of interest were then computed by averaging the segmented EEG data for each electrode and participant. Next, difference waveforms were constructed in a manner like Holroyd and Krigolson (2007). The difference waveforms were derived as such: (1) Control difference would be hits versus misses in the Control condition. (2) Hits during the unexpected correct block versus misses in the expected correct block (thus unexpected hits vs. unexpected misses), (3) Hits in the expected correct block versus misses in the unexpected correct block (expected hits v. expected misses).

To determine magnitude of activation of the reward positivity, the component was quantified for each participant as the maximum peak of the difference wave at channel FCz via visual inspection and previous literature (Holroyd & Coles, 2002). A following maximum peak search with a 200ms window around the peak of the grand average difference wave (150 – 350ms) quantified the maximum peak activation of the reward positivity (Hajcak, Moser, Holroyd & Simons, 2006; Proudfit, 2015). A similar method was employed for the BP; however,

as the BP is a negative-oriented component, a minimum peak search was applied to 100ms bins across the length of the entire waveform with at channel Cz (Shibasaki & Hallett, 2006).

2.2.6 Data Analysis

Behavioural data were comprised of participant hit frequency, participant response time, quadrant end-point, X-axis constant and variable error, Y-axis constant and variable error, and radial error and variability. Participant hit frequency was derived from the total number of trial hits and converted to a percentage. Similarly, participant response time was derived from the time it took for participants to hit a target based on movement onset from hearing the auditory cue. Quadrant end-point was defined as where participants ended during misses, measured in x-y coordinates. X-axis and Y-axis constant errors were computed as the difference of participant endpoints and actual target cue location. X-axis and Y-axis variable error was calculated as the standard deviation of the X-axis and Y-axis constant error. To compute radial error, X-axis and Y-axis residuals (difference of endpoint and target) were then computed into a vector unit. Radial variable error was then computed as the standard deviation of the radial error. Means for each variable were generated for each participant based on each quadrant and condition. These behavioural variables were then later subject to two-way repeated measures ANOVA, with an $\alpha = 0.05$ requirement. All measures of confidence were set at 95%.

To confirm the presence of reward positivity, single-samples t-tests were used to compare the max peak amplitudes of each expectancy's difference wave against zero ($\alpha = 0.05$). This was done under the assumption that a p-value greater than 0.05 would indicate no statistical difference between a null effect and the component of interest. To compare effects of quadrant and condition for the reward positivity, a two-way repeated measures ANOVA was conducted with a criterion of $\alpha = 0.05$. With regards to the bereitschaftspotential, a one-way repeated

measures ANOVA was conducted across conditions, followed by a two-way repeated measures ANOVA examining effects of condition and bins. All significance testing was conducted at $\alpha = 0.05$. All measures of confidence were set at 95%. Additionally, all post-hoc examinations were conducted with a Bonferroni correction.

2.3 Results

Endpoint values were averaged together and plotted according to quadrant across conditions (figure 4).

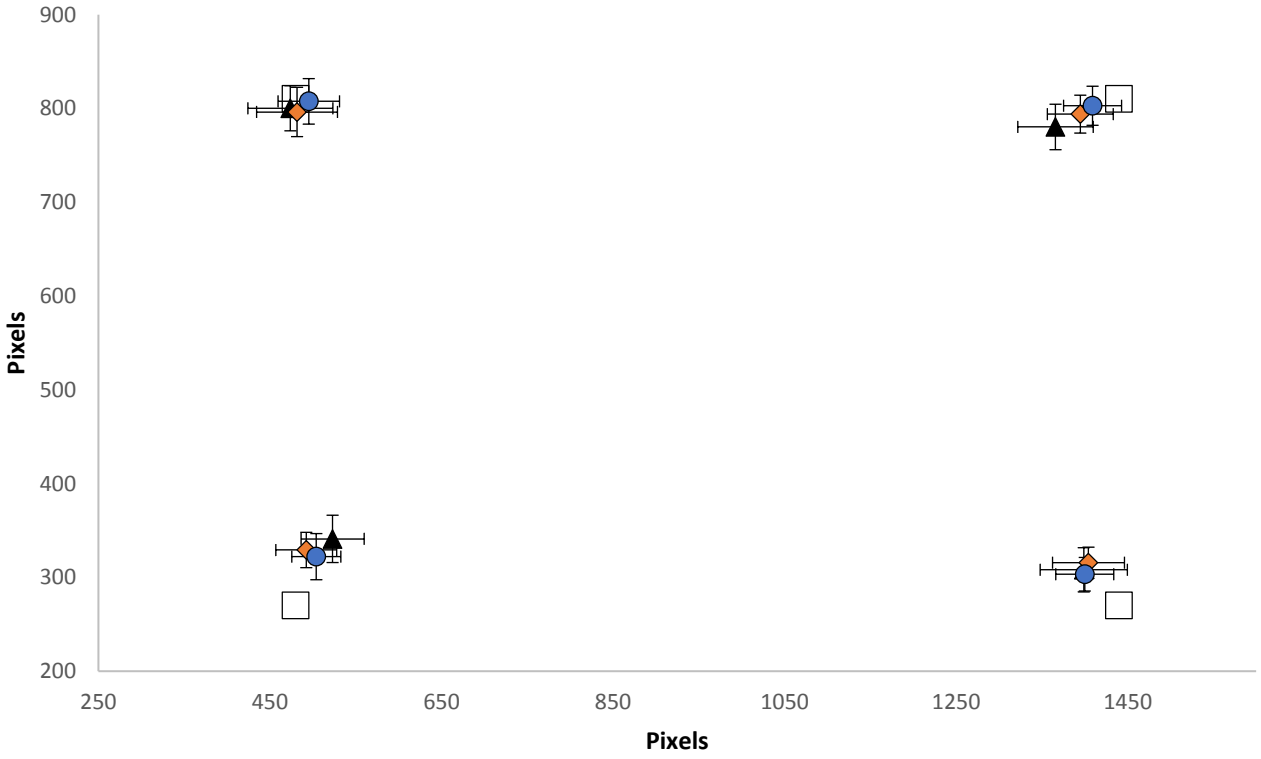

Figure 4. Distribution of quadrant endpoints per condition. (1) triangles indicate expected condition type, (2) diamonds indicate control condition type, (3) circles indicate unexpected condition type. Square denotes target cue location on screen per quadrant. Bars depict 95% Confidence with respect to X and Y axis.

A two-way repeated measures ANOVA was applied to hit frequency with respect to condition and quadrant (condition: expected correct, control, unexpected correct; quadrant: 1, 2, 3, 4). Results demonstrated an effect of expectancy on participant hit frequency, $F(13, 26) = 110.38$, p

< 0.05, η^2_p = 0.55, and no effect of quadrant, $F(13, 39) = 0.78$, $p > 0.05$, η^2_p = 0.03. No observed interaction was seen between condition and quadrant, $F(6, 78) = 1.51$, $p > 0.05$. Post hoc analysis of hit frequency and expectancy demonstrated a difference between expected correct and control conditions, $t(13) = 24.35$, $p < 0.05$, a difference between control and unexpected correct conditions, $t(13) = 51.17$, $p < 0.05$, and a difference between expected correct and unexpected correct conditions, $t(13) = 51.17$, $p < 0.05$.

A following two-way repeated measures ANOVA was applied to participant's response time with respect to condition and quadrant (condition: expected correct, control, unexpected correct; quadrant: 1, 2, 3, 4). Results demonstrated an effect of expectancy on participant movement time, $F(13, 26) = 7.36$, $p < 0.05$, $\eta^2_p = 0.08$, and no effect of quadrant. $F(13, 39) = 0.78$, $p > 0.05$, $\eta^2_p = 0.0.3$. No interaction of quadrant and condition was observed, $F(6, 78) = 0.145$, $p > 0.05$. A following post hoc between movement time and conditions revealed a difference in expected correct and control conditions, $t(13) = 6.501$, $p < 0.05$, and a difference in times for expected to unexpected correct, $t(13) = 8.248$, $p < 0.05$. Relationships between condition and time to target is seen in figure 5.

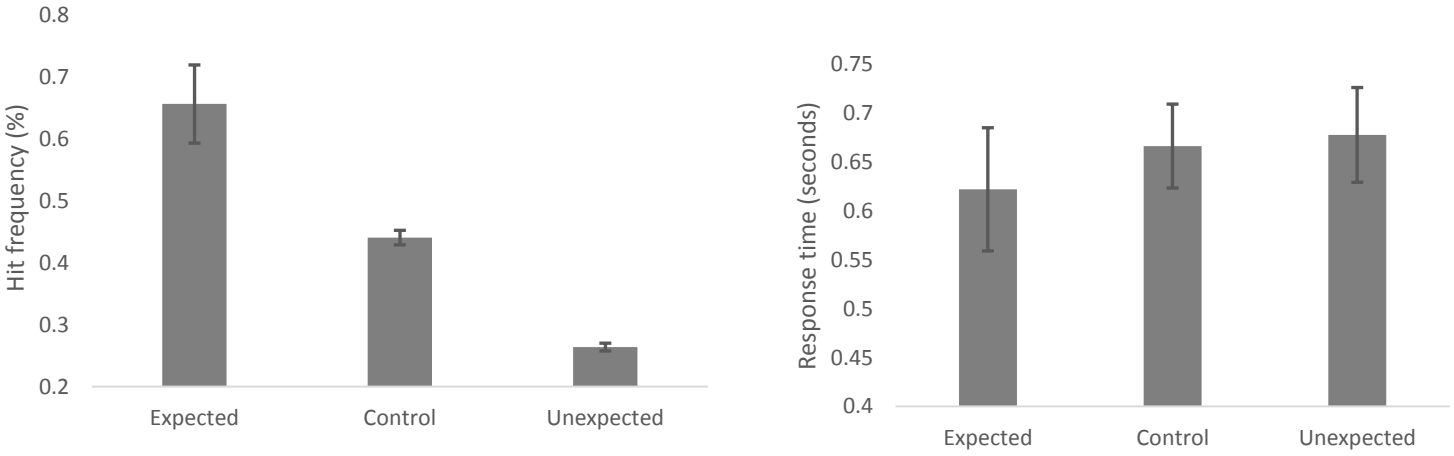

Figure 5. Behavioural data of hit frequency (left) and response time (right) under condition types. Bars depict 95% confidence intervals.

See figure 6 and 7 for summary of Constant X/Y(CEX/CEY) and Variable Error X/Y (VEX/VEY). A two-way repeated measures ANOVA was applied to CEX with respect to

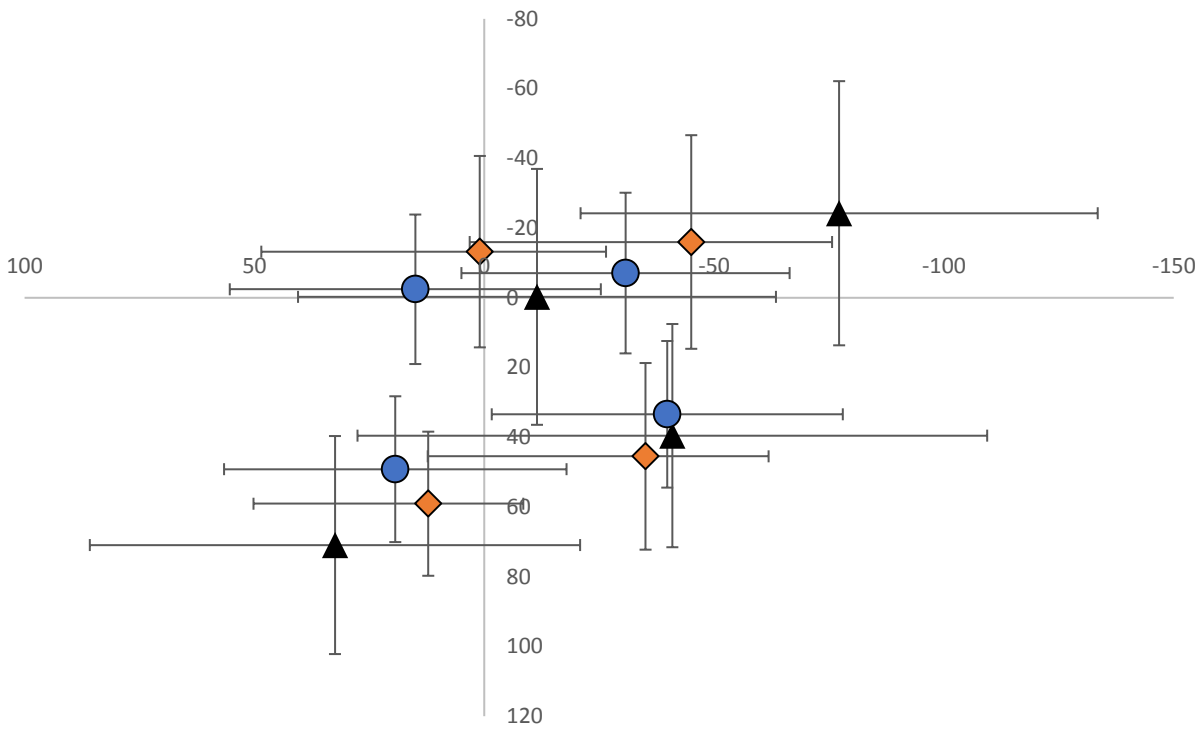

Figure 6. Distribution of X and Y constant error) across experimental conditions. Quadrant distribution oriented as shown in figure 2. Bars demonstrate 95% confidence on both X and Y axis, respectively. (1) triangles indicate expected correct condition type, (2) diamonds indicate control condition type, (3) circles indicate unexpected correct condition type.

condition and quadrant (condition: expected correct, control, unexpected correct; quadrant, 1, 2, 3, 4). Results demonstrated an effect of expectancy on CEX, $F(2, 32) = 3.49$, $p < 0.05$, $\eta^2_p = 0.0057$, and an effect of quadrant on CEX, $F(3, 48) = 3.16$, $p < 0.05$, $\eta^2_p = 0.11$. Post hoc analysis of CEX demonstrate a difference between quadrant 1 and quadrant 2, $t(16) = 2.963$, $p < 0.05$, and a difference between quadrant 2 and quadrant 4, $t(16) = 4.02$, $p < 0.05$. Correction of condition demonstrates no difference between expectancies. A two-way repeated measures ANOVA was then applied to CEY with respect to condition and quadrant (condition: expected

correct, control, unexpected correct; quadrant, 1, 2, 3, 4). Results demonstrated an effect of quadrant, $F(3, 48) = 13.57$, $p < 0.05$, $\eta^2_p = 0.32$. A summary of post hoc results is seen in table 2.

Table *2. Table of constant error (Y-Axis) post hoc comparisons across four quadrants. P-values adjusted via Bonferroni correction.*

Quadrant	1	2	3
1	-	-	-
2	$t(16) = 3.544$, $p < 0.05$	-	-
3	$t(16) = 3.928$, $p < 0.05$	$t(16) = 7.485$, $p < 0.05$	-
4	$t(16) = 6.045$, $p < 0.05$	$t(16) = 9.174$, $p < 0.05$	$t(16) = 0.433$, $p > 0.05$

A later two-way repeated measures ANOVA was applied to VEX with respect to condition and quadrant (condition: expected correct, control, unexpected correct; quadrant, 1, 2, 3, 4). Results demonstrated no effect of quadrant or condition; $F(3, 48) = 0.32$, $F(2, 32) = 0.13$, p

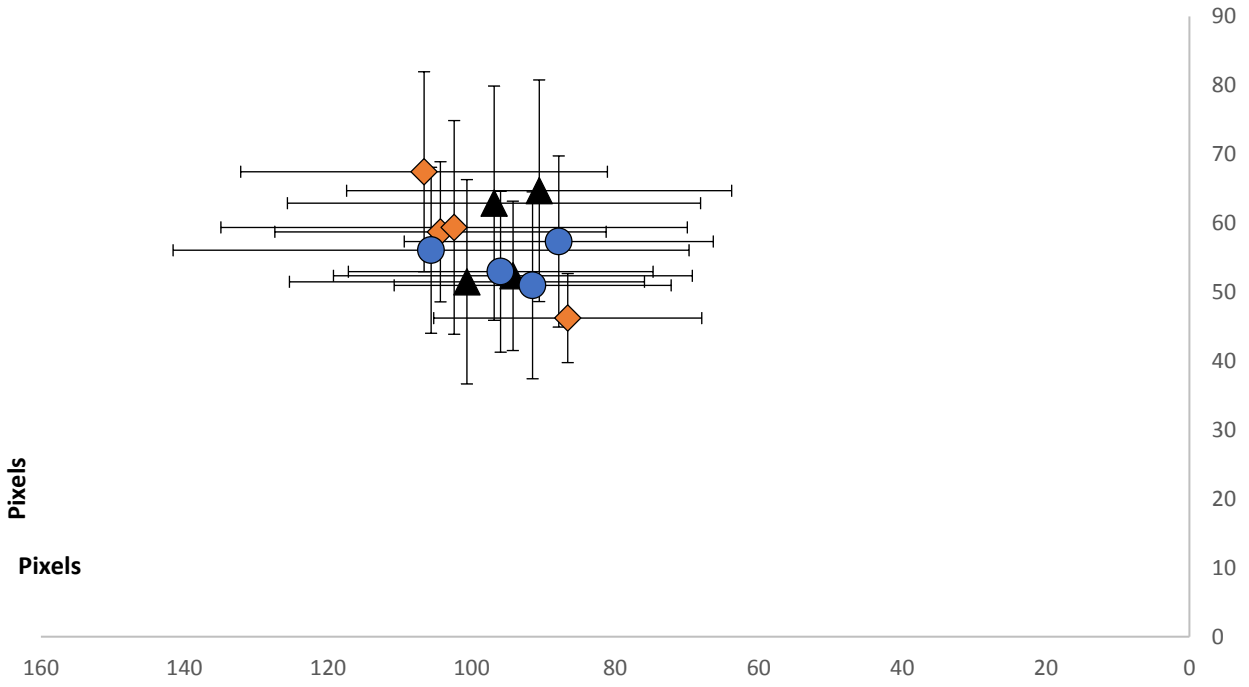

Figure 7. Distribution of X and Y variable error across experimental conditions. Quadrant distribution oriented as shown in figure 2. Bars demonstrate 95% confidence on both X and Y axis, respectively. (1) triangles indicate expected condition type, (2) diamonds indicate control condition type, (3) circles indicate unexpected condition type.

> 0.05. A two-way repeated measures ANOVA was applied to VEY with respect to condition and quadrant (condition: expected correct, control, unexpected correct; quadrant, 1, 2, 3, 4). Results demonstrated an effect of quadrant, $F(3, 48) = 3.04$, $p > 0.05$, $\eta^2_p = 0.03$. Post hoc analysis demonstrated a difference between quadrant and quadrant 4, $t(16) = 2.536$, $p < 0.05$.

A following two-way repeated measures ANOVA was applied to radial error values (expectancy: expected correct, control, unexpected correct, quadrant: 1, 2, 3, 4). An effect of expectancy, $F(2, 32) = 4.61$, $p < 0.05$, $\eta^2_p = 0.058$, was observed; however, no effect of quadrant, $F(3, 48) = 2.096$, $p > 0.05$, or interaction between expectancy or quadrant, $F(6,96) = 0.82$, $p > 0.05$. was seen. Post hoc data demonstrated a difference between expected correct conditions and control conditions, $t(16) = 6.279$, $p < 0.05$, as well as a difference between expected correct and unexpected correct conditions, $t(16) = 3.792$, $p < 0.05$. A following two-way repeated measures ANOVA was applied to radial variable error values (expectancy: expected correct, control, unexpected correct, quadrant: 1, 2, 3, 4). No effect of expectancy or quadrant was observed, $F(2, 32) = 4.64$, $p > 0.05$, $F(3, 48) = 1.24$, $p > 0.05$, respectively.

Component Topography

Two main components were examined within this study. First was the reward positivity, in which peak activation is observed within a window 150 – 350ms post stimulus presentation due to observed maximum peak amplitude (Hajcak et al., 2006). The second component examined was the bereitschaftspotential (BP), a negative going potential preceding -1.5 – 0 s before movement onset (Shibasaki & Hallett, 2006). Both components were mid-central, maximal and minimal at electrode sites FCz (reward positivity) and Cz (BP), respectively. Evidence of component activation at appropriate electrode sites is demonstrated topographically (see figure 8).

Reward Positivity

Single-samples t-tests against zero of each difference wave confirmed the presence of the reward positivity: expected correct: $t(16) = 5.540$, $p < 0.05$, control: $t(16) = 10.765$, $p < 0.05$, unexpected correct: $t(16) = 7.307$, $p < 0.05$. Grand average waveforms of each of the three conditions demonstrated a component that is consistent in morphology and timing as the reward positivity

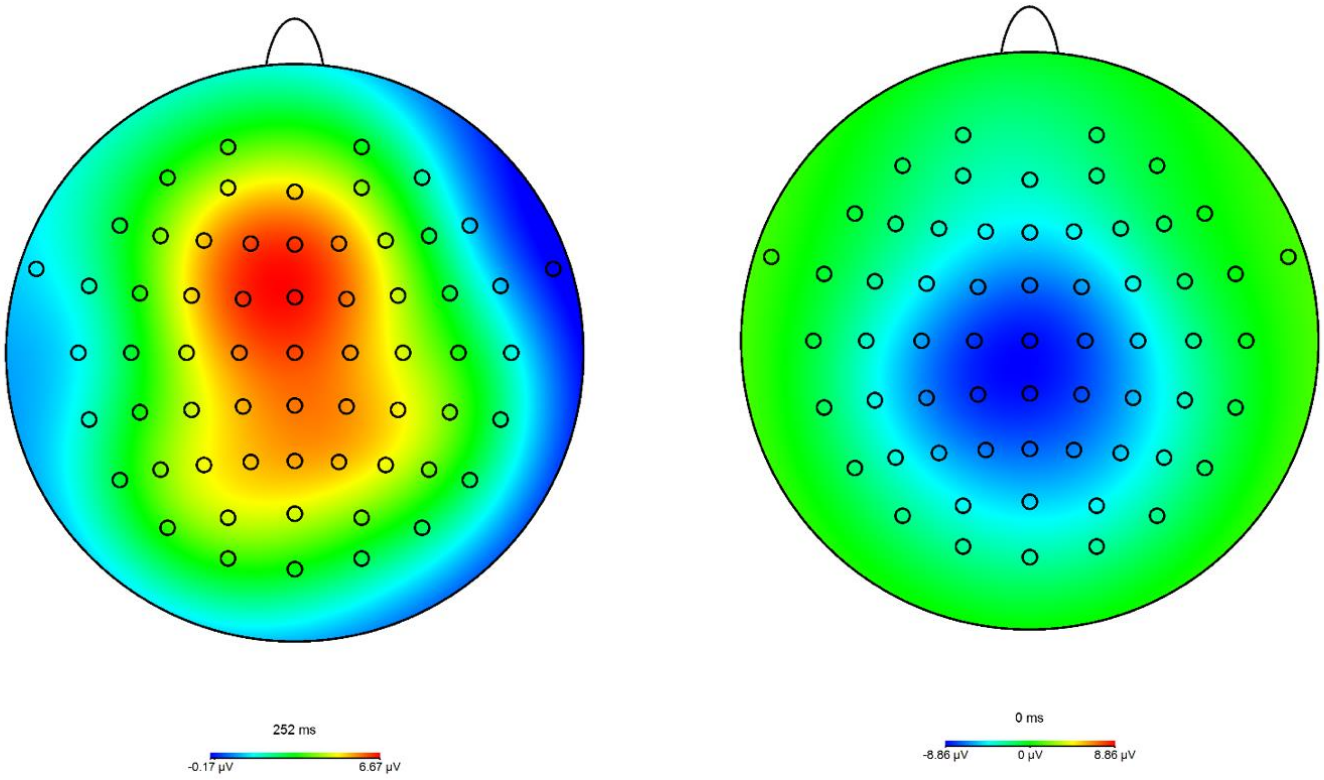

Figure 8. Topographic plots focusing on channels FCz (Left) and Cz (Right). Distributions of activation are consistent with the timing and location of the Reward Positivity and Bereitschaftspotential, respectively.

to condition (condition: expected correct, control, unexpected correct), $F(16, 32) = 6.99$, $p < 0.05$, $\eta^2_p = 0.19$.

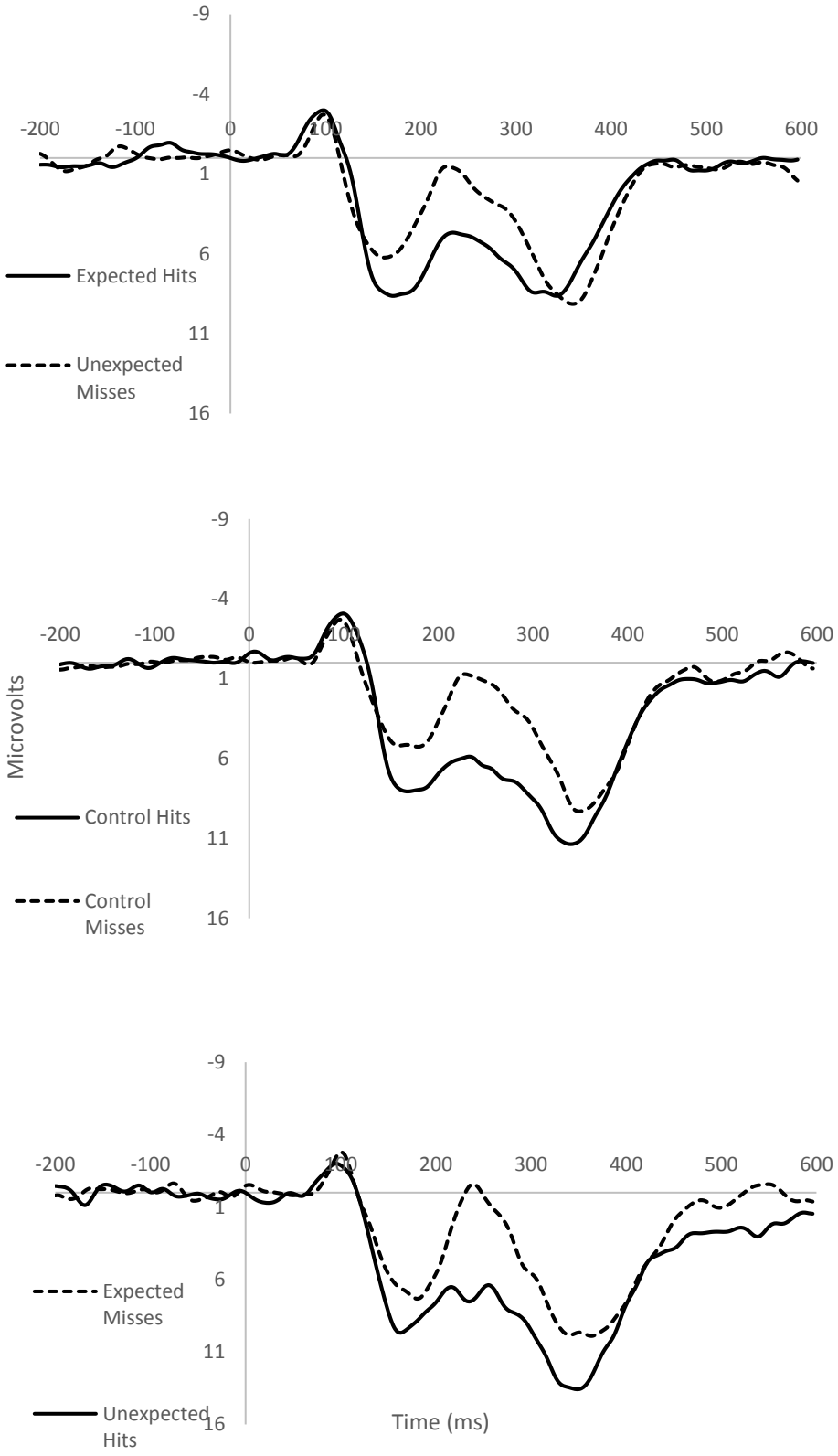

Figure 9. Conditional Waveforms matched to expectancy prior to generation of difference waveforms. Range occurs from -200ms of stimulus onset (feedback in the form of ✓or x) to 600ms post stimulus presentation. Units are measured in µV, with negative plotted upwards as standard convention.

Post hoc comparison of conditions demonstrated an observed difference in expected correct versus unexpected correct conditions, $t(16) = 2.724$, $p < 0.05$. Comparison of unexpected correct to control demonstrate: $t(16) = 2.248$, $p < 0.05$; however, a comparison of expected correct to control conditions revealed no difference, $t(16) = 0.012$, $p > 0.05$ (see figure 9). A Pearson product-moment correlation coefficient was then computed to assess the relationship between hit frequency and the reward positivity. There was a weak positive correlation between expected hits and expected correct reward positivity amplitudes, r = 0.10, n = 17, $p = 0.693$, a weak negative correlation was seen between control hit frequency and control reward positivity amplitudes, r = -.341, n = 17, $p = .180$, and a moderate negative correlation was seen between unexpected hit frequency and unexpected correct reward positivity amplitudes, r = -.429, n = 17, $p = 0.085$.

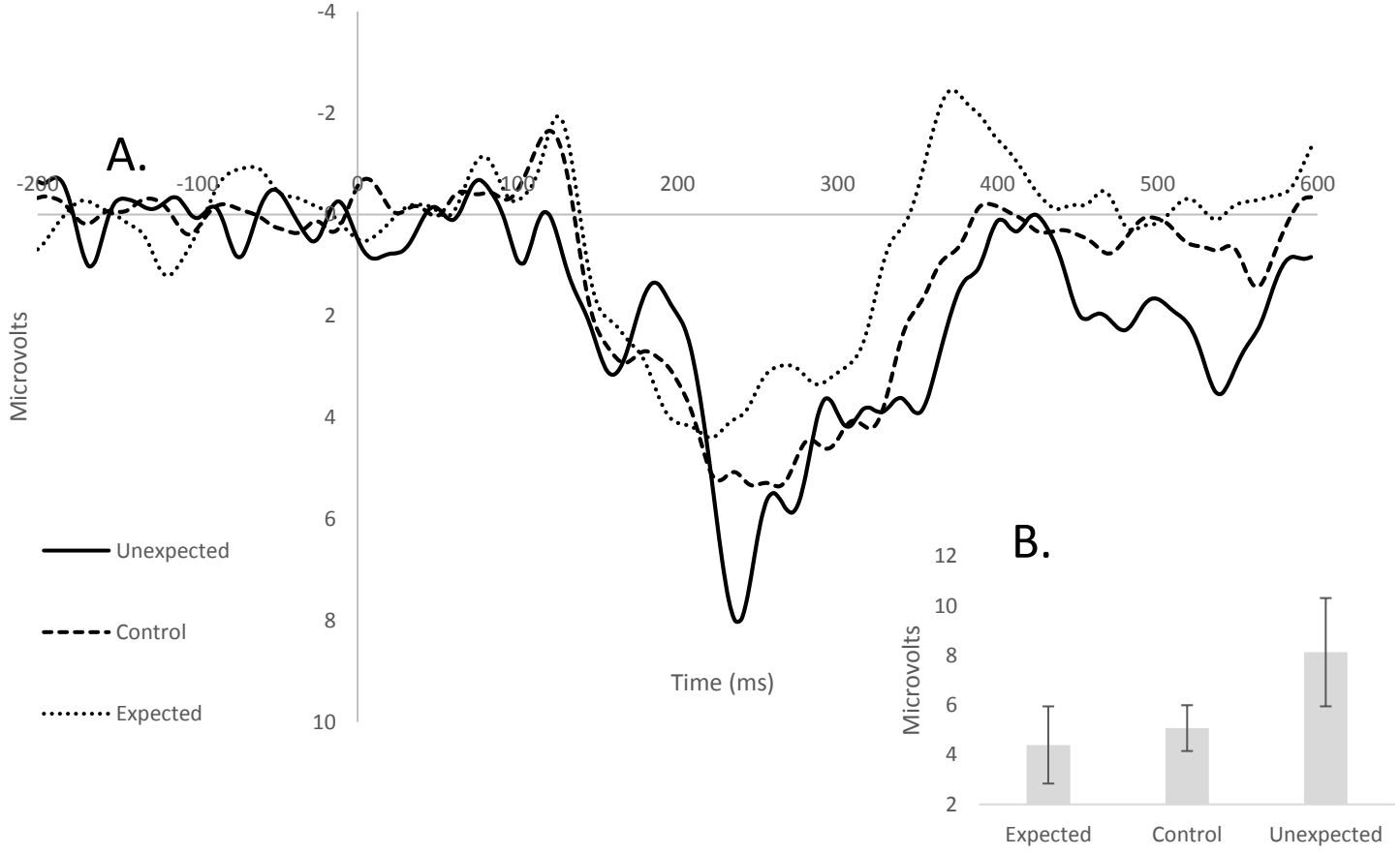

Figure 10. (A) Difference waveforms of the reward positivity across three expectancies. Range occurs from -200ms of stimulus onset (feedback in the form of ✔ or x) to 600ms post stimulus presentation. Units are measured in µV, with negative plotted upwards as standard convention. (B) Peak voltages of each condition. Error bars depict 95% confidence.

Bereitschaftspotential

Grand average waveforms demonstrate a component that is consistent in form and timing with the BP (figure 11). A repeated measures ANOVA was conducted on the BP with respect to condition (condition: expected correct, control, unexpected correct), $F(16,32) = 3.66$, $p < 0.05$, $\eta^2_p = 0.05$. There was no difference observed when comparing expected to control $t(16) = 0.773$, $p > 0.05$, or expected correct to unexpected correct, $t(16) = 0.323$, $p > 0.05$; however, a comparison of control to unexpected correct reveals a difference $t(16) = 2.39$, $p < 0.05$.

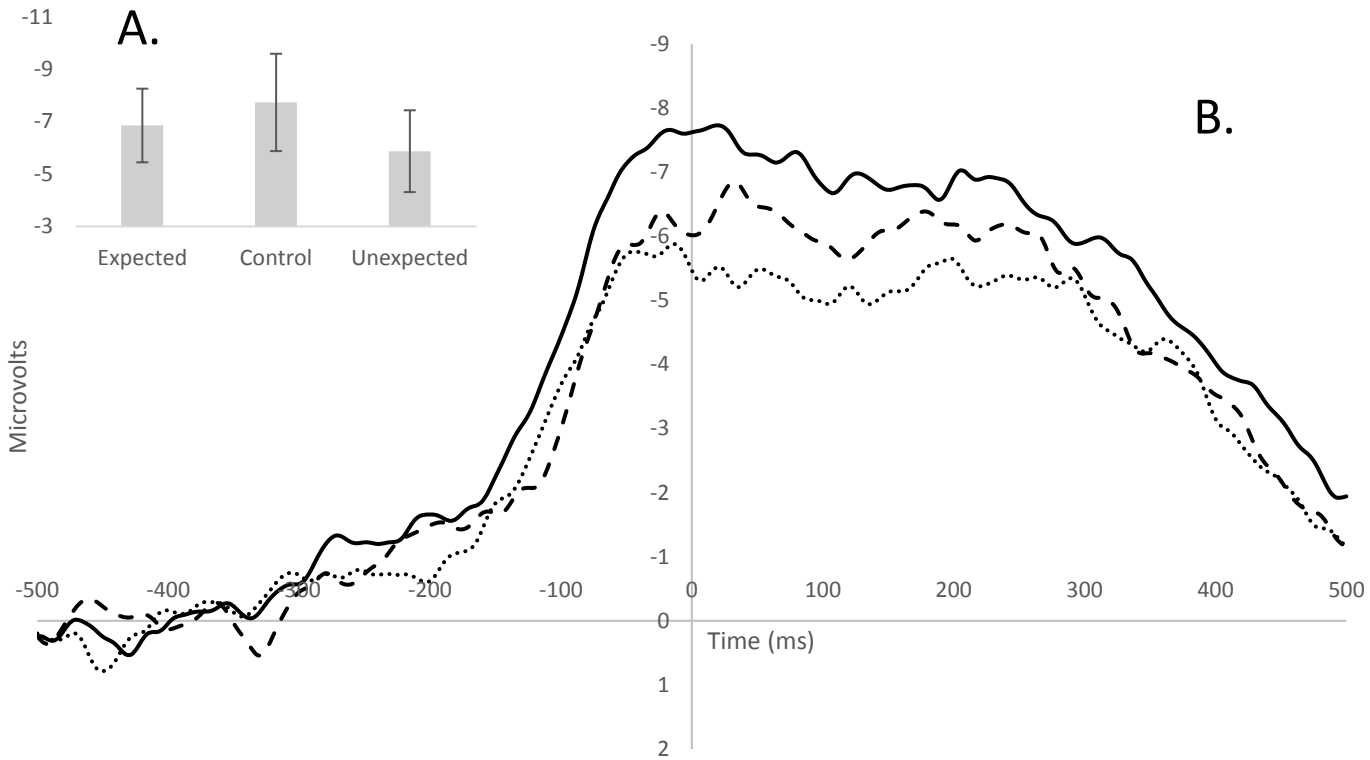

Figure 11. (A) Peak voltage bereitschaftspotential of three block conditions, ranging from -500ms to 500ms. 0ms denotes movement onset. Units are measured in μV, with negative plotted upwards as standard convention. (B) Component waveform of the bereitschaftspotential.

A two-way repeated measures ANOVA was applied to BP bin values (expectancy: expected correct, control, unexpected correct, Bin: 1 - 8). An effect of expectancy, $F(2, 32) =$

4.47, $p < 0.05$, $\eta^2_p = 0.025$, was observed. An effect of bin was also seen, $F(7, 112) = 17.63$, $p < 0.05$, $\eta^2_p = 0.26$. See figure 12.

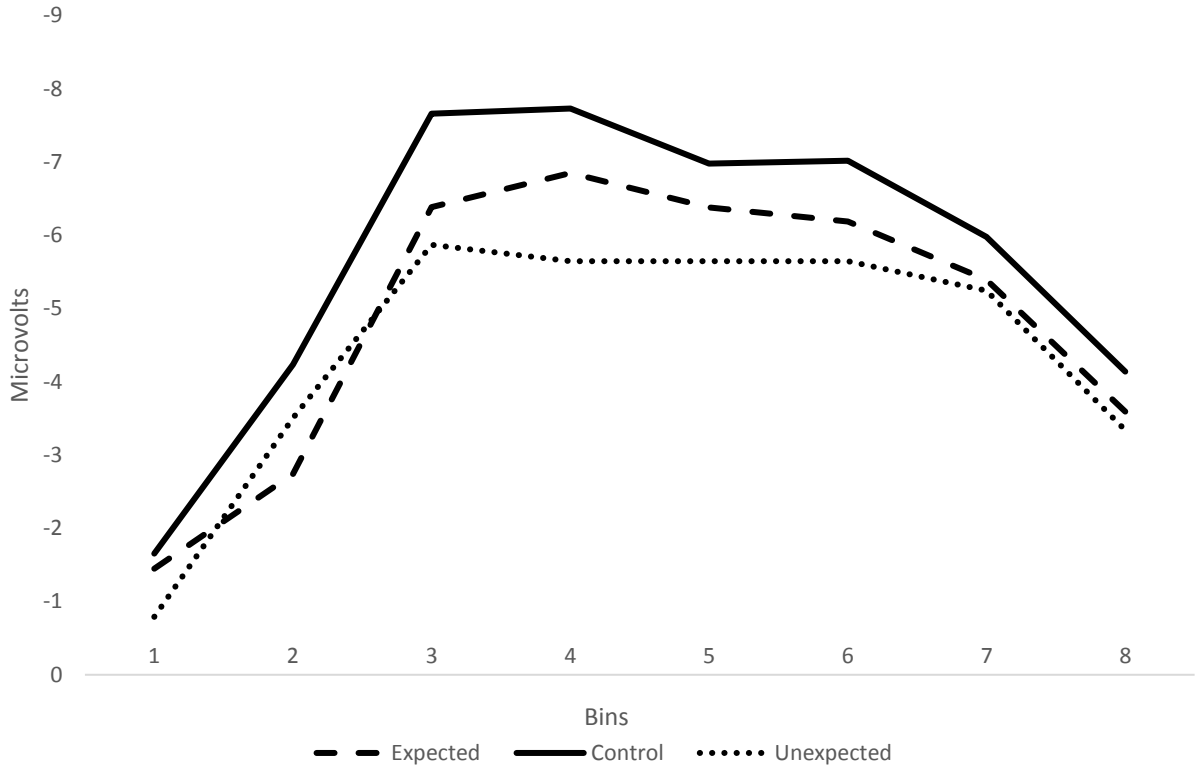

Figure 12. Bereitschaftspotential of three block conditions: solid line represents control condition, dotted line represents expected condition, dashed line represents unexpected condition. Bins compose of 100ms segments across entire timeframe (-500ms to 500ms). Bin 1 corresponds to -500ms to -400ms, bin 2 corresponds to -400ms to -300ms and forward until 400ms to 500ms for bin 8. Units are measured in μV, with negative plotted upwards as standard convention.

Post hoc analysis of conditional data across the binned waveform demonstrated a difference between expected and control conditions, $t(16) = 3.792$, $p < 0.05$, and between unexpected and control conditions, , $t(16) = 9.116$, $p < 0.05$, and with no difference between expected and unexpected conditions, , $t(16) = 0.8654$, $p > 0.05$. Further examination of condition and bin is outlined in table 3.

Table *3. Table of comparisons across bin and condition for the bereitschaftspotential. Bin 2 corresponds to 200ms – 300ms, bin 3 corresponds to 300ms – 400ms, and forward until bin 5. Bold text denotes observed differences within the waveform.*

	Conditional Comparisons		
Bin	Expected correct to Control	Control to unexpected correct	Expected correct to unexpected correct
2	$t(16) = 1.242$, $p > 0.05$	$t(16) = 1.413$, $p > 0.05$	$t(16) = 0.609$, $p > 0.05$
3	$t(16) = 1.718$, $p > 0.05$	**$t(16) = 3.006$, $p < 0.05$**	$t(16) = 0.8109$, $p > 0.05$
4	$t(16) = 1.51$, $p > 0.05$	**$t(16) = 3.283$, $p < 0.05$**	$t(16) = 1.562$, $p > 0.05$
5	$t(16) = 1.435$, $p > 0.05$	$t(16) = 1.918$, $p > 0.05$	$t(16) = 1.083$, $p > 0.05$

2.4 Discussion

In the present study, participants were asked to perform a target task with experimental manipulations in expectancy. Three discrete conditions were presented (expected, control, unexpected), with four possible quadrant targets to hit, per trial. Behaviourally, participants were seen to consistently undershoot the intended target. These findings are consistent with prior literature, as it is known occluding or removing targets from sight remove online control to account for trajectory adjustments (Woodworth, 1899). Indeed, the current paradigm elicits memory-guided reaches by way of target presentation and removal before movement onset. Memory-guided reaches have reliably demonstrated a trend of undershooting targets (Westwood, Heath, & Roy, 2003; Heath, 2005; Krigolson and Heath, 2004; Krigolson et al., 2012).

Because of this fact, participants may have relied on a more proprioceptive learning system than one involved with visual-spatial systems, and thus may have been more dependent on an offline mode of control (Heath, 2005). This study may share overlapping features to that of a targeting task conducted by Heath (2005), and function as that of a memory-guided reach, as mentioned above. For example, in the present study, a target appears on screen, then disappears

and movement onset does not begin until ~400 – 600ms; this is analogous to the "Target Delay 500" condition in Heath (2005) where target disappeared after preview phase (~2000ms) and an audio tone prompted movement 500ms after target was gone. Albeit a more extreme version of the current paradigm, the structure of 'Target-delay-movement' without the assistance of target location may very well utilize similar kinematics as the occluded limb condition.

Modulations of expectancy were reflected in a behavioural manner, as participants experienced a decrease in accuracy as conditional difficulty increased. The electrophysiological data did not reflect a complete modulation of expectancy, however. An evident difference is seen within reward positivity amplitudes between unexpected and expected/control conditions, yet no difference may be said across the expected or control condition. The similarity between the two conditions may be due to an insufficient difference between block difficulties. As shown in Williams et al. (2017), the reward positivity is known to behave in a presumably non-linear fashion and follows a similar sigmoidal pattern of phasic dopaminergic release (Schultz, Dayan, Montague, 1998). Presumably then, the lack of difference may follow a similar pattern to that of the 'Very Easy' and 'Easy' conditions of Williams et al. (2017), such that the sensitivity of neurons either cannot detect the difference between the two conditions, or they were too similar for the system to care. Furthermore, modulation of the reward positivity is evident within the previously mentioned unexpected condition comparisons. Differences in amplitude are consistent with prior reward expectancy tasks (Holroyd & Krigolson, 2007; Williams et al., 2017), and may serve to add further evidence towards a frontal medial system responsible for a hierarchical error-evaluation system for motor movements (Krigolson & Holroyd, 2006; Krigolson & Holroyd, 2007a, b). It is not definitive to state however, as the current study did not evaluate

low-level processing within parietal cortex; however, the frontal-medial response of the reward positivity appears consistent with prior literature.

An expectancy effect has also been shown to affect amplitudes of the BP. A reduction of BP amplitude was observed regarding unexpected designed blocks, with a corresponding increase of time to target ranges and reduced hit frequency. This finding corresponds with evidence of repeated errors result in a decrease in BP amplitude and thus increase the time to target (Perri, 2016). Oddly there was an increase in BP amplitude during the control condition, with no difference observed between expected or unexpected, or unexpected and expected. A point of significance is the comparison of unexpected condition to the control; a slightly unintuitive outcome. There is a possible consideration for how this may occur: Firstly, there is an evident scaling of BP towards expectancy; however, due to the similarity of expected to control reward positivity amplitudes, as stated previously, it may be the BP amplitudes are being obfuscated due to the similarity of the two condition types. This is a possibility as the bereitschaftspotential has been previously observed to increase in amplitude due to levels of motivation and success (McAdam & Seales, 1969), whereas other BP-related components such as the lateralize readiness potential (LRP) have also been shown to demonstrate a modulation of amplitude during commission and observation of motor errors (van Schie, Mars, Cole, & Bekkering, 2004). Compilation of the two block types combined with the innate ambiguity of success in the control trials may be interfering with a discrete difference between condition types. Further evidence of the intricacies of motor-preparatory and reward-evaluation systems are the overlap of basal ganglionic nuclei. Intracranial recordings demonstrate BP activity within the striatum as well as the cortex, suggesting an overlapping of the circuitry involved in these processes (Rektor, Kuba, & Brazdil, 2002).

A more telling story may be the BP recorded was overlapping with another component known as the contingent negative variability (CNV). Although subject to contention, evidence has demonstrated support of the late BP and terminal aspect of the CNV share many similarities (Grünewald, Grünewald-Zuberbier, Netz, Hömberg & Sander, 1979), such that they share many similar psychological and physiological substrates. The CNV and late BP show a similar localization of midfrontal regions, in addition to similar effects of response speed variation in both the BP and CNV. Furthermore, the CNV has demonstrated some sensitivity to expectancy and predictability (Deecke, Niesser, & Ziller, 1980). It is then demonstrated that both the reward positivity and bereitschaftspotential are modulated in some form by expectancy effects, and more specifically expectancy set in a motor-movement scenario.

CHAPTER 3: CONSIDERATIONS & DISCUSSION

3.1 Implications

In the present thesis, we have established a scaling of expectancy towards motor learning. Utilizing an expectancy task derived from prior time estimation paradigms (Miltner, Braun, & Coles, 1997; Holroyd & Krigolson, 2007; Williams et al., 2017), we indexed violations of expectancy through a human ERP, the reward positivity. In association with the reward positivity, we examined an additional motor component, the bereitschaftspotential. In line with prior literature, we have demonstrated a modulation of reward positivity amplitudes with respect to expectations; specifically, we have found maximal amplitudes under the conditions of unexpectedness. Conversely, reward positivity amplitudes were found to be minimal when outcomes matched with expectations; however, a caveat in processing had occurred when no evident difference between expected and control conditions occurred. Further evidence corroborated a decreased amplitude of the BP towards decreased motor behavioural performance, and a perplexing outcome contrary to our original hypothesis. Rather than yielding the largest amplitudes during rewarding situations (McAdam & Seales, 1969), it was found that control conditions demonstrated maximal amplitudes of the BP. Considerations of these discrepancies are examined in the following section.

3.1.1 Discussion

To extrapolate on the reward positivity, the component has been theorized as analogous to the phasic dopamine release of the mesolimbic dopaminergic system (Schultz et al., 1997; Holroyd & Coles, 2002). Given the similarity between perceived expected outcomes and the control condition, it appears participants may have been expecting similar outcomes. This is contrary to behaviour results, as mean accuracy ratings under expected conditions differed from

those of the control by some margin. To reconcile these findings, several options are addressed. The first to be considered is the activation thresholds. As it is known, the mesolimbic dopaminergic system is sensitive to degrees of expectancy and reward values (Schultz et al.,1997; Cromwell & Schultz, 2003), and indeed this is demonstrated within an EEG context, as established by Williams et al. (2017). It may be considered that the perceived similarity of these two conditions to participants, and their consequential violations in expectancy was not enough to elicit a difference between them. This is not an entirely novel finding and follows a similarity when examining the upper-ends of conditional expectancies in Williams et al. (2017). In particular, a sigmoidal trend was observed among condition types leading to no discernable difference between their 'very easy' and 'easy' conditions as well as the 'hard' and 'very hard' conditions. Similarity of the current study conditions (expected and control) may follow a similar pattern.

A second consideration is the reliance of sensory prediction errors in a motor context. As outlined by Izawa and Shadmehr (2011), corrective motor behaviours are not solely mediated through 'reward' prediction errors but are also adjusted through sensory prediction errors as well. Without visual guidance of both the target, cursor, or acting limb, participants may have been heavily dependent on sensory prediction errors in addition to reward prediction errors to modulate future movements. Outlined by Wolpert, Ghahramani, and Jordan (1995), the notion of an internal 'sensor' to continuously evaluate commissioned and ongoing movements is cornerstone to the concepts of motor control. To elaborate, there are two main monitors of motor control: the inverse model, in which a copy of a signal (likened to copying a file on a computer) is sent to various centers of the brain concurrent with its original execution. This 'efferent' copy is thought to be processed by the parietal regions of the cortex, and cerebellum, and is involved

with the coordination and tuning of motor trajectory through inverse dynamics models (Kawato, 1999). In the present experiment, it may be that participants were learning from predominantly proprioceptive information, in that the main relevant sensory feedback other than visual-reward input (feedback on hits or misses) was the cooperation of internal signals and reinforcement feedback.

This integrating of explicit reward prediction errors with sensory prediction errors may be the impetus to allow participants to complete the task with some degree of competency. By doing so this may address the discrepancy in participant accuracy and reward positivity amplitudes. Based on the current results, it may suggest some further interaction in reward-expectation processes and tuning of motor behaviours. Further reinforcement learning concepts, and their interaction with motor control are discussed by Wolpert, Gahramani, and Flanagan (2001). Sufficed to say, the likely scenario is both an overlap of condition similarity and a division of sensory and reward prediction errors to mediate behaviour and response.

In addition to the reward positivity, a second reconciliation must be made. As our original hypothesis regarding the bereitschaftspotential involved a greater amplitude during expected (or more rewarding) situations (McAdam & Seales, 1969), it was quite unintuitive to observe the control condition with the largest amplitude. In further examination of the paradigm structure, it could be posited that the component being elicited is the contingent negative variation (CNV: Gaillard & Naatanen, 1976). The CNV paradigm follows a relatively simple structure, including a preparatory stimulus (S1) to which it indicates participants of an imperative stimulus (S2). This later stimulus is then used as a cue to tell participants to execute some specific action. Examining the structure of the current experiment, there may be similar overlap with the basic CNV task structure. S1 could be considered the quadrant target, and the S2 the auditory tone; however,

there is some disagreement to if the present component is strictly the CNV. In the present study, S1 serves as not only a cue to S2, but also a goal-directed target to which a motor plan must be formed. Given that the task follows a similar behavioural and electrophysiological profile as previous memory-guided reach studies (Krigolson et al., 2012), the activity of the observed component must have a motor-related underpinning.

As the control condition increased and decreased in an equivalent rate, it can be considered to be not learnable in terms of expectancy. Given that prior literature of the CNV demonstrated an an increase in amplitude during situations of in situations of unpredictability and unexpectedness (Deecke et al., 1980), our current results convey a similar structure; however, an account of Gruenwald and colleagues (1978) demonstrated support of the late BP and terminal aspect of the CNV to share many overlapping psychological and physiological properties. To separate the CNV and BP further, a review by Rohrbaugh (1983) has stated the terminal aspect of the CNV is in fact a motor component through various arguments. In particular, several arguments involve an increase in CNV amplitude during tasks which demand larger amounts of muscular effort or rapidity of response. Furthermore, its topographic and morphological similarities to the BP are made clear. Despite its contention, there is an overlap in CNV and BP literature with agreeance and disagreement to what each component constitutes. In the present study, as demonstrated through use of memory-guided reaches and expectation, it may be surmised the observed component is the blend of motor and expectancy related functions acting in tandem.

3.1.2 Application

In brief, applications of expectancy learning in a motor context are broad. To generalize the findings in a more ecological manner, we may consider motor expectation learning in a rehabilitative fashion. There is evidence of levodopa and other similar dopamine antagonists

greatly improving treatment outcomes in different severity of stroke patients (Floel & Cohen, 2010). This is of interest due to the overlapping theme of dopamine and reward expectation-motor control circuitry. Further evidence demonstrates that available dopamine increases the optimization of motor learning in M1 (Molina-Luna et al., 2009). As there is an effect in learning and motor memory in both stroke patients and healthy individuals, perhaps this increase in available dopamine encourages the formation of stronger neural connection in the residual, undamaged neurons. By taking advantage of known methods to index reward expectancies (i.e. reward positivity), utilizing expectancies in a motor fashion in tandem with possible pharmacological interventions may yield more successful therapies (Molina-Luna et al., 2009; Floel & Cohen, 2010).

3.1.3 Limitations

As discussed and considered earlier, there have been some limitations in the experimental design. Optimally the separation of block difficulty would be at a greater distance, reaching the upper-ends of previously established literature. To accomplish this goal would require a fundamental change in the structuring of the experiment. Rather than four possible quadrants, it may be more lucrative to instead focus on one. A single quadrant task with the current expectancy manipulation may be more analogous to the time estimation tasks, which expanded or shrank their window of response in a more controlled manner. Other considerations include the difficulty of the task, discussed as a closed loop task where limited environmental feedback was afforded to the participant; however, this may not be to the detriment of the experiment, and a comparison between motor expectancy modulations of an open loop versus closed loop task may be of interest in terms of motor learning

3.1.4 Conclusion

In this study we have provided evidence of an expectancy-driven reinforcement learning system in a motor context. Modulations of frontal-medial regions indexed by the reward positivity provide support of a higher-level processing in error and reward evaluation. These changes are reflected in behavioural decreases in performance in both accuracy and movement time; the latter supported by a decrease in amplitude of the bereitschaftspotential (BP). Leading support to a motor-system sensitive to expectancy is a difference in BP amplitude, such that the component was most negative in conditions of highest uncertainty. Behavioural and neurophysiological data support results from Krigolson et al., 2012, in which memory-guided reaches are the product of a motor plan stored in memory rather than a motor plan generated prior to movement execution.

References

Amiez, C., Joseph, J. P., & Procyk, E. (2005). Anterior cingulate error-related activity is modulated by predicted reward. Euro*pean Journal of Neuroscience, 21*(12), 3447-3452.

Ball, T., Schreiber, A., Feige, B., Wagner, M., Lücking, C. H., & Kristeva-Feige, R. (1999). The role of higher-order motor areas in voluntary movement as revealed by high-resolution EEG and fMRI. *Neuroimage, 10*(6), 682-694.

Balleine, B. W., Delgado, M. R., & Hikosaka, O. (2007). The role of the dorsal striatum in reward and decision-making. *Journal of Neuroscience, 27*(31), 8161-8165.

Berger, H. (1929). Über das elektrenkephalogramm des menschen. *Archiv für psychiatrie und nervenkrankheiten, 87*(1), 527-570.

Botvinick, M. M., Niv, Y., & Barto, A. C. (2009). Hierarchically organized behavior and its neural foundations: a reinforcement learning perspective. *Cognition, 113*(3), 262-280.

Brunia, C. H., van Boxtel, G. J., & Böcker, K. B. (2012). Negative slow waves as indices of anticipation: the Bereitschaftspotential, the contingent negative variation, and the stimulus-preceding negativity. *In The Oxford handbook of event-related potential components*.

Bush, G., Vogt, B. A., Holmes, J., Dale, A. M., Greve, D., Jenike, M. A., & Rosen, B. R. (2002). Dorsal anterior cingulate cortex: a role in reward-based decision making. *Proceedings of the National Academy of Sciences, 99*(1), 523-528.

Calabresi, P., Picconi, B., Tozzi, A., Ghiglieri, V., & Di Filippo, M. (2014). Direct and indirect pathways of basal ganglia: a critical reappraisal. *Nature neuroscience, 17*(8), 1022.

Colino, F. L., Howse, H., Norton, A., Trska, R., Pluta, A., Luehr, S. J., ... & Krigolson, O. E. (2017). Older adults display diminished error processing and response in a continuous tracking task. *Psychophysiology, 54*(11), 1706-1713.

Cromwell, H. C., & Schultz, W. (2003). Effects of expectations for different reward magnitudes on neuronal activity in primate striatum. *Journal of neurophysiology, 89*(5), 2823-2838.

Cruccu, G., Berardelli, A., Inghilleri, M., & Manfredi, M. (1990). Corticobulbar projections to upper and lower facial motoneurons. A study by magnetic transcranial stimulation in man. *Neuroscience letters, 117*(1-2), 68-73.

Cummings, J. L. (1993). Frontal-subcortical circuits and human behavior. *Archives of neurology, 50*(8), 873-880.

Deecke, L., Niesser, A., & Ziller, B. (1980). Modality (visual and tactile) and stimulus predictability influence contingent negative variation and reaction time. *In Progress in brain research* (Vol. 54, pp. 301-308). Elsevier.

Devinsky, O., Morrell, M. J., & Vogt, B. A. (1995). Contributions of anterior cingulate cortex to behaviour. *Brain, 118*(1), 279-306.

Doya, K. (2000). Complementary roles of basal ganglia and cerebellum in learning and motor control. *Current opinion in neurobiology, 10*(6), 732-739.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & psychophysics, 16*(1), 143-149.

Fiorillo, C. D., Tobler, P. N., & Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science, 299*(5614), 1898-1902.

Fumuro, T., Matsuhashi, M., Mitsueda, T., Inouchi, M., Hitomi, T., Nakagawa, T., ... & Takahashi, R. (2013). Bereitschaftspotential augmentation by neuro-feedback training in Parkinson's disease. *Clinical Neurophysiology, 124*(7), 1398-1405.

Floel, A., & Cohen, L. G. (2010). Recovery of function in humans: cortical stimulation and pharmacological treatments after stroke. *Neurobiology of disease, 37*(2), 243-251.

Gaillard, A. W., & Näätänen, R. (1976). Modality effects on the contingent negative variation in a simple reaction-time task. *In The responsive brain* (pp. 40-45).

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.

Gehring, W. J., Goss, B., Coles, M. G., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological science, 4*(6), 385-390.

Gehring, W. J., Liu, Y., Orr, J. M., & Carp, J. (2012). The error-related negativity (ERN/Ne). *Oxford handbook of event-related potential components*, 231-291.

Goldberg, G. (1985). Supplementary motor area structure and function: Review and hypotheses. *Behavioral and brain Sciences, 8*(4), 567-588.

Goodale, M. A., Pelisson, D., & Prablanc, C. (1986). Large adjustments in visually guided reaching do not depend on vision of the hand or perception of target displacement. *Nature, 320*(6064), 748.

Graybiel, A. M., Aosaki, T., Flaherty, A. W., & Kimura, M. (1994). The basal ganglia and adaptive motor control. *Science, 265*(5180), 1826-1831.

Graybiel, A. M. (2000). The basal ganglia. *Current biology, 10*(14), R509-R511.

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

Grünewald, G., Grünewald-Zuberbier, E., Netz, J., Hömberg, V., & Sander, G. (1979). Relationships between the late component of the contingent negative variation and the Bereitschaftspotential. *Electroencephalography and Clinical Neurophysiology, 46*(5), 538-545.

Haggard, P. (2008). Human volition: towards a neuroscience of will. *Nature Reviews Neuroscience, 9*(12), 934.

Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2006). The feedback-related negativity reflects the binary evaluation of good versus bad outcomes. *Biological psychology, 71*(2), 148-154.

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.

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.

Heath, M. (2005). Role of limb and target vision in the online control of memory-guided reaches. *Motor control, 9*(3), 281-309.

Held, R., & Gottlieb, N. (1958). Technique for studying adaptation to disarranged hand-eye coordination. *Perceptual and Motor Skills, 8*(3), 83-86.

Hikosaka, O., Nakamura, K., Sakai, K., & Nakahara, H. (2002). Central mechanisms of motor skill learning. *Current opinion in neurobiology, 12*(2), 217-222.

Hoffstaedter, F., Grefkes, C., Caspers, S., Roski, C., Palomero-Gallagher, N., Laird, A. R., ... & Eickhoff, S. B. (2014). The role of anterior midcingulate cortex in cognitive motor control: evidence from functional connectivity analyses. *Human brain mapping, 35*(6), 2741-2753.

Holroyd, C. B., Dien, J., & Coles, M. G. (1998). Error-related scalp potentials elicited by hand and foot movements: evidence for an output-independent error-processing system in humans. *Neuroscience letters, 242*(2), 65-68.

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

Holroyd, C. B., Krigolson, O. E., Baker, R., Lee, S., & Gibson, J. (2009). When is an error not a prediction error? An electrophysiological investigation. *Cognitive, Affective, & Behavioral Neuroscience, 9*(1), 59-70.

Holroyd, C. B., Ribas-Fernandes, J. J., Shahnazian, D., Silvetti, M., & Verguts, T. (2018). Human midcingulate cortex encodes distributed representations of task progress. *Proceedings of the National Academy of Sciences, 115*(25), 6398-6403.

Ikeda, A., Lüders, H. O., Burgess, R. C., & Shibasaki, H. (1992). Movement-related potentials recorded from supplementary motor area and primary motor area: role of supplementary motor area in voluntary movements. *Brain, 115*(4), 1017-1043.

Isomura, Y., Takekawa, T., Harukuni, R., Handa, T., Aizawa, H., Takada, M., & Fukai, T. (2013). Reward-modulated motor information in identified striatum neurons. Journal of *Neuroscience, 33*(25), 10209-10220.

Izawa, J., & Shadmehr, R. (2011). Learning from sensory and reward prediction errors during motor adaptation. *PLoS computational biology, 7*(3), e1002012.

Kapogiannis, D., Campion, P., Grafman, J., & Wassermann, E. M. (2008). Reward-related activity in the human motor cortex. *European Journal of Neuroscience, 27*(7), 1836-1842.

Kawagoe, R., Takikawa, Y., & Hikosaka, O. (1998). Expectation of reward modulates cognitive signals in the basal ganglia. *Nature neuroscience, 1*(5), 411.

Kawato, M. (1999). Internal models for motor control and trajectory planning. *Current opinion in neurobiology, 9*(6), 718-727.

Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996). A neostriatal habit learning system in humans. *Science, 273*(5280), 1399-1402.

Kornhuber, H. H., & Deecke, L. (1965). Hirnpotentialänderungen bei Willkürbewegungen und passiven Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale. *Pflüger's Archiv für die gesamte Physiologie des Menschen und der Tiere, 284*(1), 1-17.

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

Krigolson, O. E., Hassall, C. D., & Handy, T. C. (2014). How we learn to make decisions: rapid propagation of reinforcement learning prediction errors in humans. *Journal of cognitive neuroscience, 26*(3), 635-644.

Krigolson, O., & Heath, M. (2004). Background visual cues and memory-guided reaching. *Human movement science, 23*(6), 861-877.

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.

Krigolson, O. E., Heinekey, H., Kent, C. M., & Handy, T. C. (2012). Cognitive load impacts error evaluation within medial-frontal cortex. *Brain research, 1430*, 62-67.

Krigolson, O. E., & Holroyd, C. B. (2006). Evidence for hierarchical error processing in the human brain. *Neuroscience, 137*(1), 13-17.

Krigolson, O. E., & Holroyd, C. B. (2007a). Predictive information and error processing: The role of medial-frontal cortex during motor control. *Psychophysiology, 44*(4), 586-595.

Krigolson, O. E., & Holroyd, C. B. (2007b). Hierarchical error processing: different errors, different systems. *Brain research, 1155*, 70-80.

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.

Libet, B., Gleason, C. A., Wright, E. W., & Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential) the unconscious initiation of a freely voluntary act. *Brain, 106*(3), 623-642.

Luck, S. J. (2014). *An introduction to the event-related potential technique*. MIT press.

Luck, S. J., & Girelli, M. (1998). *Electrophysiological approaches to the study of selective attention in the human brain*.

Luck, S. J., & Kappenman, E. S. (Eds.). (2011). *The Oxford handbook of event-related potential components*. Oxford university press.

Margulies, D. S., Kelly, A. C., Uddin, L. Q., Biswal, B. B., Castellanos, F. X., & Milham, M. P. (2007). Mapping the functional connectivity of anterior cingulate cortex. *Neuroimage, 37*(2), 579-588.

McAdam, D. W., & Seales, D. M. (1969). Bereitschaftspotential enhancement with increased level of motivation. Electroencephalography and clinical *Neurophysiology, 27*(1), 73-75.

Mill, J. S. (1863). *Utilitarianism*. London: Parker, Son and Bourn.

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.

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.

Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *Journal of neuroscience, 16*(5), 1936-1947.

Nikooyan, A. A., & Ahmed, A. A. (2015). Reward feedback accelerates motor learning. *Journal of Neurophysiology, 113*(2), 633-646.

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.

Paus, T. (2001). Primate anterior cingulate cortex: where motor control, drive and cognition interface. *Nature reviews neuroscience, 2*(6), 417.

Pavlov, P. I. (2010). Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex. *Annals of neurosciences, 17*(3), 136.

Perri, R. L., Berchicci, M., Lucci, G., Spinelli, D., & Di Russo, F. (2016). How the brain prevents a second error in a perceptual decision-making task. *Scientific reports, 6*, 32058.

Porter, R., & Lemon, R. (1993). *Corticospinal function and voluntary movement* (No. 45). Oxford University Press, USA.

Potts, G. F., Martin, L. E., Burton, P., & Montague, P. R. (2006). When things are better or worse than expected: the medial frontal cortex and the allocation of processing resources. *Journal of cognitive neuroscience, 18*(7), 1112-1119.

Proudfit, G. H. (2015). The reward positivity: From basic research on reward to a biomarker for depression. *Psychophysiology, 52*(4), 449-459.

Rektor, I., Bareš, M., & Kubová, D. (2001). Movement-related potentials in the basal ganglia: a SEEG readiness potential study. *Clinical neurophysiology*, *112*(11), 2146-2153.

Rektor, I., Kuba, R., & Brázdil, M. (2002). Interictal and ictal EEG activity in the basal ganglia: an SEEG study in patients with temporal lobe epilepsy. *Epilepsia, 43*(3), 253-262.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical conditioning II: Current research and theory, 2*, 64-99.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of neurophysiology, 80*(1), 1-27.

Schultz, W. (2016). Dopamine reward prediction-error signaling: a two-component response. Nature Reviews. *Neuroscience, 17*(3), 183.

Schultz, W., Carelli, R. M., & Wightman, R. M. (2015). Phasic dopamine signals: from subjective reward value to formal economic utility. *Current opinion in behavioral sciences, 5*, 147-154.

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

Shibasaki, H., Barrett, G., Halliday, E., & Halliday, A. M. (1980). Components of the movement-related cortical potential and their scalp topography. *Electroencephalography and clinical neurophysiology, 49*(3-4), 213-226.

Shibasaki, H., & Hallett, M. (2006). What is the Bereitschaftspotential?. *Clinical neurophysiology, 117*(11), 2341-2356.

Shidara, M., & Richmond, B. J. (2002). Anterior cingulate: single neuronal signals related to degree of reward expectancy. *Science, 296*(5573), 1709-1711.

Stauffer, W. R., Lak, A., Kobayashi, S., & Schultz, W. (2016). Components and characteristics of the dopamine reward utility signal. *Journal of Comparative Neurology, 524*(8), 1699-1711.

Stauffer, W. R., Lak, A., & Schultz, W. (2014). Dopamine reward prediction error responses reflect marginal utility. *Current biology, 24*(21), 2491-2500.

Sutton, R. S. (1988). Learning to predict by the methods of temporal differences. *Machine learning, 3*(1), 9-44.

Sutton, R. S., & Barto, A. G. (1990). *Time-derivative models of pavlovian reinforcement*.

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

Tachibana, Y., & Hikosaka, O. (2012). The primate ventral pallidum encodes expected reward value and regulates motor action. *Neuron, 76*(4), 826-837.

Taylor, M. J. (1978). Bereitschaftspotential during the acquisition of a skilled motor task. *Electroencephalography and clinical neurophysiology, 45*(5), 568-576.

Thorndike, E. L. (1911). *Animal intelligence: Experimental studies*. Macmillan.

van Schie, H. T., Mars, R. B., Coles, M. G., & Bekkering, H. (2004). Modulation of activity in medial frontal and motor cortices during error observation. *Nature neuroscience, 7*(5), 549.

Vogt, B. A. (2005). Pain and emotion interactions in subregions of the cingulate gyrus. *Nature Reviews Neuroscience, 6*(7), 533.

Vogt, B. A., Finch, D. M., & Olson, C. R. (1992). Functional heterogeneity in cingulate cortex: the anterior executive and posterior evaluative regions. *Cerebral cortex, 2*(6), 435-443.

Walsh, M. M., & Anderson, J. R. (2012). Learning from experience: event-related potential correlates of reward processing, neural adaptation, and behavioral choice. *Neuroscience & Biobehavioral Reviews, 36*(8), 1870-1884.

Westwood, D. A., Heath, M., & Roy, E. A. (2003). No evidence for accurate visuomotor memory: Systematic and variable error in memory-guided reaching. *Journal of motor behavior, 35*(2), 127-133.

Williams, C. C., Hassall, C. D., Trska, R., Holroyd, C. B., & Krigolson, O. E. (2017). When theory and biology differ: The relationship between reward prediction errors and expectancy. *Biological psychology, 129*, 265-272.

Wolpert, D. M., Ghahramani, Z., & Flanagan, J. R. (2001). Perspectives and problems in motor learning. *Trends in cognitive sciences, 5*(11), 487-494.

Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science, 269*(5232), 1880-1882.

Wolpert, D. M., Miall, R. C., & Kawato, M. (1998). Internal models in the cerebellum. *Trends in cognitive sciences, 2*(9), 338-347.

Woodworth, R. S. (1899). Accuracy of voluntary movement. *The Psychological Review: Monograph Supplements, 3*(3), i.

Zamble, E. (1968). Classical conditioning of excitement anticipatory to food reward: Partial reinforcement. *Psychonomic Science, 10*(3), 115-116.