COMBINED PUNISHMENT AND REWARD FEEDBACK DURING SEQUENCE LEARNING

COMBINED PUNISHMENT AND REWARD FEEDBACK THROUGH TRANSITION SCHEDULES

BY

RAJBIR SINGH SIDHU, B.Sc.

A THESIS

SUBMITTED TO THE DEPARTMENT OF KINESIOLOGY

AND THE SCHOOL OF GRADUATE STUDIES

OF MCMASTER UNIVERSITY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF SCIENCE

Master of Science (2021) McMaster University
(Kinesiology) Hamilton, Ontario, Canada

TITLE:	Combined punishment and reward feedback through transition schedules
AUTHOR:	Rajbir Singh Sidhu B.Sc. (Life Sciences), McMaster University, Hamilton, Canada
SUPERVISOR:	Dr. Michael J. Carter, Ph.D.
NUMBER OF PAGES:	xi, 72

Lay Abstract

An important part of learning any skill is receiving information that helps us recognize mistakes and improve our performance, known as feedback. In fact, feedback presented as a punishment or reward has been shown to improve an individual's ability to learn and retain skills, respectively. Therefore, can combining punishment and reward feedback benefit both learning and retention? One way to deliver both types of feedback is using a transition schedule. Some have recommended that transitioning from punishment to reward feedback would be most effective, while others have suggested the reverse order. The current study examined whether the order of receiving punishment and reward feedback affected learning and retention. To test this, subjects either received punishment-to-reward feedback or reward-to-punishment feedback during a key-pressing task. Our results did not find conclusive evidence that the order mattered for learning and retention. Furthermore, the difference between punishment and reward feedback overall was smaller than previously thought. These findings highlight that more studies may need to be conducted to get a better understanding of whether the order of punishment and reward feedback can benefit both learning and retention.

Abstract

Punishment and reward feedback during motor learning tasks appear to have some beneficial impact on learning and retention, respectively. Therefore, it is possible that combining punishment and reward feedback would benefit both learning and retention. Within the sports coaching domain, a combination of punishment and reward feedback schedule has been suggested to improve performance. According to the coaching literature, the most effective approach is providing reward-to-punishment feedback. However, transitioning from punishment-to-reward feedback may be more effective based on the motor learning literature. The present study examined the utility of combining punishment and reward feedback through a transition schedule approach during a serial reaction time task. To test the competing predictions about feedback order, half the participants received punishment-to-reward feedback and the other half received the reverse order. Our results revealed that training response time significantly improved with no significant difference between the order of feedback. However, both types of feedback order did not improve retention during the same-day and delayed post-tests. Yet, the non-significant equivalence test indicates that these findings remain inconclusive. Finally, within-subjects analysis of the punishment and reward conditions found that training significantly improved response time with no difference between them. In this case, the equivalence test was significant, revealing

that the estimated effect was surprisingly small. Overall, the current study failed to find conclusive evidence that the order of a transition feedback schedule matters for learning and retention. However, the difference between punishment and reward conditions may be smaller than previously assumed by motor learning studies.

Acknowledgements

I would first like to thank my supervisor, Mike, for his advice, support, and guidance throughout my masters experience. As my mentor and friend, you have pushed me to become a better researcher and supported me whenever I needed it. It was a pleasure working with you.

I would also like to thank my committee members, Dr. Jennifer Heisz and Dr. Joshua Cashaback, and my external examiner Dr. Jae Patterson for taking time out of their busy schedules to provide insightful comments and recommendations on this thesis.

And a special thanks to my parents, Balbir and Paramjit. You have always believed in me and gave me the freedom to find my own way. It means a lot to know that you are proud of the person that I've become and all that I have accomplished. Finally, I would thank my sister, Jasmeet. Despite all the headaches I might have put you through, you have always been there for me and I am grateful for all you have done.

Declaration of Academic Achievement

Format and organization of thesis

This thesis is prepared in the standard format as outlined in the *Guide for the preparation of master's and doctoral theses* provided by McMaster University's School of Graduate Studies.

Contributions to content of thesis

Rajbir S. Sidhu; Brad McKay; Michael J. Carter

Contributions using CRediT (Contributor Roles Taxonomy)

R.S. Sidhu: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Software, Validation, Visualization, Writing - Original Draft Preparation, Writing - Review & Editing *B. McKay:* Data curation, Formal analysis, Validation, Writing - Review & Editing *A. Poskus:* Investigation *A. Yuen:* Investigation *M.J. Carter:* Conceptualization, Data curation, Funding acquisition, Methodology, Supervision, Writing - Review & Editing

Contents

List of Figures

Chapter 1

Literature Review

1.1 General introduction

Many of us can still recall the terrifying, yet exciting experience of learning to ride a bicycle. Rhythmically pushing down on the pedals with your legs to move forward and navigating the bicycle using the handlebars, while simultaneously balancing on it and avoiding pedestrians on the street. However, with the guidance of our parents and months of practice, we begin to master this motor skill. We start riding the bicycle independently, travelling farther distances, and falling less frequently. Even after a long break (e.g., hours, days, or years), we can retain the ability to ride a bicycle with high proficiency. This entire process of learning a new motor skill is emblematic of a behavioural phenomenon known as motor learning.

Motor learning is an essential part of human behaviour that covers a broad range of domains. As a result, it has been defined in various ways. According to Magill and Anderson (2017), motor learning is described as a change in the capability of executing a skilled movement, through practice, that is inferred from a relatively permanent change in performance. It has also been defined as a set of internal processes that lead to the relatively permanent changes in the ability to produce skilled movements as a result of practice (Schmidt et al. 2019). Alternatively, motor learning can be viewed as an umbrella term consisting of two types of learning: (a) skill acquisition, defined as the formation of a new motor skill through practice, and (b) skill maintenance, the ability to maintain a consistent level of performance of a skill in a fluctuating environment (Krakauer *et al.*, 2019). Finally, the cognitive neuroscience literature refers to motor learning as alterations in the neural architecture of the brain that lead to a change in the ability to perform a movement (Diedrichsen and Kornysheva, 2015).

Despite the variety of motor learning definitions, there are common features among them. First, motor learning is the result of *practice and/or experience*; it leads to the acquisition of a skilled movement and increases the capability of executing it. Second, motor learning *cannot be directly observed*. Given that learning is likely an internal process, occurring at multiple levels within the nervous system, it is often inferred from changes in performance (Magill and Anderson, 2017; Schmidt *et al.*, 2019). For example, an individual's initial attempts at riding a bicycle often produces awkward movements and leads to many falls. However, through practice, their movements and coordination will gradually improve. Therefore, by observing these changes in performance, it can be inferred that an individual is learning. Third, motor learning is *relatively permanent*, that is, the learner should be capable of repeatedly executing the movement, even after a period without any practice (Diedrichsen and Kornysheva, 2015; Magill and Anderson, 2017; Schmidt *et al.*, 2019).

Motor learning can be further described based on three different stages that support the generation of a skilled movement (Krakauer *et al.*, 2019; Wolpert *et al.*, 2011). The first stage, referred to as goal selection, requires the ability to process sensory information from the environment and form a movement goal. The next stage, action selection, an appropriate movement to accomplish that goal is chosen, and finally, the action execution stage involves the performance of the movement with accuracy and precision. Consequently, improvements in any of these stages can be characterized as motor learning (Krakauer *et al.*, 2019).

1.2 Measurement of motor learning

Given that motor learning is defined in a variety of ways, different methods have been used to infer learning from performance in the laboratory setting. In the field of kinesiology, motor learning is inferred based on the degree of permanence of the motor skill (Schmidt *et al.*, 2019). A common method to measure this feature is a delayed retention test, where the learners execute a practiced skill following a period of inactivity or rest (i.e., minutes, hours, or days). If there is a high degree of permanence, it is inferred that the individual has learned the motor skill (Kantak and Winstein, 2012; Magill and Anderson, 2017; Schmidt and Bjork, 1992). Therefore, learning and retention in kinesiology are used interchangeably.

In contrast, the motor neuroscience literature measures motor learning during the time the individual is practicing the skill (Diedrichsen and Kornysheva, 2015). In this case, learning is inferred based on changes in performance during the practice period. Specifically, if there is a significant improvement from the first practice attempt to the last, it is inferred that learning has occurred (Shmuelof *et al.*, 2012; Diedrichsen and Kornysheva, 2015). In contrast, a retention test assesses how well the motor skill is encoded, stored, and retrieved in the central nervous system. Therefore, learning and retention in the field of motor neuroscience are considered distinct components of motor behaviour.

Overall, motor learning is a complex process that can be measured using various methods. In this thesis, learning and retention are operationally defined in accordance with the cognitive neuroscience literature as it is research predominantly from this area that has motivated this thesis. That is, performance changes during training provide an index of learning and performance changes on a delayed post-test following a period of no practice provide a measure of retention.

1.3 Motor learning paradigms

The behavioural processes and cortical substrates supporting motor learning have been studied experimentally using a variety of paradigms in a laboratory setting. Despite the different paradigms employed, these simple laboratory-based tasks are presumed to provide fundamental insight into the learning processes involved with more complex real world-tasks.

Although these simple tasks have enhanced our understanding of the principles of motor learning, it is important to mention that some have argued that they do not generalize to real-world tasks. It has been suggested that paradigms with more complex tasks are likely required to generalize the findings to real world-tasks (Wulf and Shea, 2002). However, quantifying the complexity of a range of tasks, with different movement outcomes and task demands, is a challenge because it can be very subjective across different researchers. For the purpose of this review, the following sections will focus on motor adaptation and sequence learning paradigms.

1.3.1 Motor adaptation

The first method developed to record motor adaptation, referred to as a prism adaptation task, was in 1867 by Hermann von Helmholtz (Helmholtz, 2013). In prism adaptation (Harris, 1963; Redding and Wallace, 1993), participants wore goggles that contained wedge prisms, resulting in the lateral displacement of their visual field to the left or right, and were instructed to point with their index finger toward visual targets directly in front of them. Despite prism adaptation providing a clear demonstration of motor adaptation (Martin *et al.*, 1996), it was replaced by other tasks that allowed for more accurate control of a perturbation.

Another approach that has been developed is the force-field adaptation task (Shadmehr and Mussa-Ivaldi, 1994). During this task, participants hold the handle of a robotic manipulandum and make a reaching movement toward a visual target. In this case, the robotic manipulandum is programmed to apply a force on the hand that is proportional to the velocity of the movement and directly perpendicular to the direction of the movement. Although these forces immediately shifted the path of the reaching movement, participants must adapt to this dynamic perturbation on future attempts to land on the correct target.

One of the more common experimental paradigms that have been extensively used to investigate motor adaptation is the visuomotor rotation task (Pine *et al.*, 1996; Krakauer, 2009). In this paradigm, participants make reaching movements, with the cursor representing the position of the participant's hand, to visually displayed targets. During the reaching movement, a counterclockwise or clockwise rotation is imposed on the cursor, relative to the starting position, and participants must learn to counter the rotation to successfully hit the target.

In all these paradigms, the procedure generally consists of three phases of testing: (a) pre-training phase, used to measure baseline performance prior to a perturbation being introduced; (b) training phase, where a perturbation is introduced and participants learn to adapt during the training trials; and (c) post-training phase, where the perturbation is removed (see Figure 1.1A).

Specifically, during pre-training, participants familiarize themselves with the task and are provided with veridical feedback during the reaching movement. Following pre-training, a perturbation (e.g., lateral displacement of visual field, force-field, or cursor rotation) is imposed on the participant's movement leading to an immediate directional error. However, participants account for the distorted sensory feedback following several trials with the perturbation and learn to gradually modify their movement and compensate for the perturbation. Interestingly, even after the perturbation is removed, participants continue to produce the adaptive movement. But it gradually declines over time and their movements return to baseline performance, a behavioural outcome known as aftereffects (Shadmehr *et al.*, 2010).

1.3.2 Neural correlates of motor adaptation

There are several learning mechanisms that have been proposed to play an important role in adaptation including forward model-based learning (Shadmehr *et al.*, 2010) and more recently, direct policy-based learning (Hadjiosif *et al.*, 2021). A popular theory is that adaptation is driven by sensory prediction error, which is the difference between the predicted and the actual movement outcome. In particular, the sensory prediction

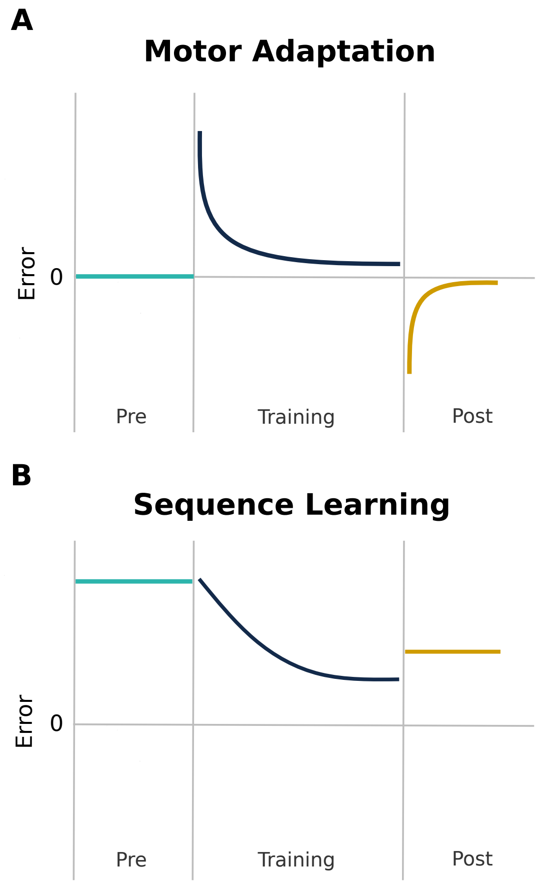

Figure 1.1: Illustrative performance curves found in adaptation learning tasks and sequence learning tasks. **(A) Adaptation learning.** Adaptation tasks are used to examine how individuals maintain their performance when a systematic perturbation is introduced into their learning environment. At the start of the training phase, the perturbation (e.g., force field, cursor rotation, etc) is introduced and causes an immediate directional error. However, participants gradually modify their movements to compensate for the perturbation, leading to a reduction in error. At the start of the post-training phase, the perturbation is removed and this results in a directional error of similar magnitude as in training, but it is now in the opposite direction (i.e., aftereffect). The error decreases rather quickly and baseline performance levels are re-established. **(B) Sequence learning.** Sequence learning tasks examine how individuals learn to produce a series of discrete actions as quickly and accurately as possible. In pre-training, individuals familiarize themselves with the task and their baseline performance is established under control conditions (e.g., no feedback). The training phase is when individuals experience the practice variable being investigated in the experiment (e.g., feedback schedule). Often two or more manipulations of the practice variable are used to identify effective motor learning interventions. With the exposure to the practice variable, error in performance gradually decreases in an exponential fashion over time. Individuals then perform the post-training phase, which is identical to the pre-training phase. The post-training phase usually happens following an extended period of no practice (e.g., 24 hours) to assess the relative permanence of what was learned during the training phase. In the post-training period, error has ideally stabilized at better-than-baseline levels due to the effects of training.

error of the movement outcome updates an internal forward model. A forward model is a system that predicts motor commands based on errors from previous movements (Shadmehr *et al.*, 2010). In addition, it is widely suggested that forward model-based learning is dependent upon specific regions of the brain (Taylor and Ivry, 2014; McDougle *et al.*, 2016). A brain structure that is believed to be involved in computing sensory prediction error and updating an internal forward model is the cerebellum (Izawa *et al.*, 2012; Taylor and Ivry, 2014).

Tseng *et al.* (2007) examined adaptation in healthy controls and participants with spinocerebellar ataxia, a degenerative neurologic disease, during a visuomotor rotation task. Compared to the healthy control, the results revealed that the spinocerebellar ataxia participants exhibited significantly larger errors and a lower rate of adaptation. Furthermore, the severity of the spinocerebellar ataxia participant's symptoms correlated with their adaptive capabilities. Participants with more severe cerebellar symptoms made smaller corrections of their movement errors from trial-to-trial, and thus, adapted and learned at a lower rate. However, this does not mean that the cerebellum is solely responsible for motor adaptation. Instead, multiple distributed cortical structures are likely involved, such as the prefrontal, premotor, and parietal cortex, functional integrated in order to guide adaptive behaviour (Krakauer *et al.*, 2004; Diedrichsen, 2005; McDougle *et al.*, 2016). Overall, these findings demonstrate that the cerebellum plays an essential role and that it could be involved in computing sensory prediction errors to update an internal forward model during adaptation (Shadmehr *et al.*, 2010; Tseng *et al.*, 2007).

1.3.3 Sequence learning

Another type of paradigm that is commonly utilized in the laboratory setting to study the behavioural characteristics of motor learning are sequence learning tasks.

One type of sequence learning task that has been employed in a laboratory setting is a simple sequence task. During this task, participants are presented with a single sequence of actions, typically 4-6 buttons on a keyboard, and instructed to execute that sequence as fast and accurately as possible. The main finding is that the execution of the sequential movements improves across practice trials, that is, participants' performance becomes faster and more accurate (Robertson, 2007).

An additional laboratory-based task that has been developed is a discrete sequence production task (Verwey, 1999). In this task, participants learn to execute two or more discrete sequences by responding to successive stimuli associated with buttons on a keyboard. At first, participants must process each successive stimulus to determine the correct sequence of actions. However, following several practice trials, they begin to execute the entire sequence in response to the first stimulus, without processing later stimuli. Therefore, once participants learn each discrete sequence order, this task examines the ability to select and produce the appropriate sequence, as fast as possible, from a group of sequences.

The most common task that has been used to investigate sequence learning is the serial reaction time task (Nissen and Bullemer, 1987). In the vast majority of serial reaction time tasks, a visual stimulus is presented at one of the four locations arranged horizontally on the center of a computer screen. Each stimulus is associated with one of four buttons on a keyboard directly below the position of the stimulus. At the start of each trial, one of the stimuli at the four locations is cued by a signal, and participants are instructed to press the corresponding button as accurately and fast as possible. After the response is made, reaction time, used as an index of performance, is recorded and then there is a fixed delay before the next visual stimulus is presented. A version of the serial reaction time task will be used in this thesis.

The order of the visual stimuli can be presented as a fixed sequence, in which the stimuli appear according to a repeating pattern, or random sequence, where the stimuli are randomly presented. By doing so, it provides a more specific measure of motor skill learning (Nissen and Bullemer, 1987; Robertson, 2007). Specifically, the average reaction time typically decreases in an exponential manner across a serial reaction time task (Nissen and Bullemer, 1987). But recently, there is a large body of evidence that suggests that motor skill learning during sequence reaction time task occurs from: (a) improvements in the execution of the individual sequence elements regardless of the pattern, referred to as sequence-independent skill learning, and (b) larger improvements in the execution of the sequence elements in a specific pattern, known as sequence-dependent skill learning (Robertson, 2007; Steel *et al.*, 2016; Robertson, 2007). As a result, by comparing the fixed sequence reaction to the random sequence reaction, the difference provides a measure to distinguish between those two forms of learning (Robertson, 2007).

In the sequence learning paradigms, the procedure also typically consists of three phases of training. First, a pre-training phase is conducted where participants familiarize themselves with the task and establish their baseline performance. Next, participants complete the training phase to measure their changes in performance. During the training phase, reaction time gradually decreases in an exponential fashion over time. Finally, a post-training phase is performed where participants often display some degree of permanence of the performance level produced during the training phase (see Figure 1.1B).

1.3.4 Reaction time and response time

A major limitation of a serial reaction time task is that reaction time, the primary measure of the task, is often misinterpreted (Chen *et al.*, 2018; Krakauer *et al.*, 2019). Reaction time is a measure of time between the onset of a stimulus and the initiation of the movement. Yet, the measure that is actually being recorded is response time. Response time is the interval between the stimulus onset and the completion of the action. In this case, the interval from the presentation of the visual stimuli to completion of the button press. Response time can be further divided into reaction time and movement time. Movement time refers to the period between the initiation of the movement and completion of the movement. As a result, changes in response time can be associated with either a reduction in reaction time, movement time, or a combination of both (Magill and Anderson, 2017; Schmidt *et al.*, 2019). With respect to the sequence learning task, a reduction in reaction time likely characterizes improvements in selecting the appropriate movement (i.e., improvements in action selection), while a reduction in movement time may represent improvements in the ability to execute the sequence movement (i.e., improvements in action execution) (Diedrichsen and Kornysheva, 2015). However, the current methodology is unable to distinguish between improvements in these components. Therefore, it is difficult to interpret the underlying component responsible for the overall improvements. Altogether, it is clear that the development of a methodology, that can isolate these measures, is required to gain a deeper understanding of the results (Chen *et al.*, 2018; Wong *et al.*, 2015). From this point forward, the outcome measure of the serial reaction time task will be referred to as response time.

1.3.5 Neural correlates of sequence learning

Identifying cortical regions in relation to sequence learning, as described above, has been well studied to better understand how these structures contribute to sequence learning. For example, Doyon *et al.* (1996) examined the cortical structures associated with the learning of a serial reaction time task. Using positron emission tomography, an imaging technique that captures and measures the metabolic activity of the brain after a radioactive tracer has been absorbed into the bloodstream, the changes in cerebral blood flow, in response to changes in neural activity, was measured during learning. Analysis of the positron emission tomography scans revealed a significant increase in cerebral blood flow, which represents higher neural activity, in the ventral striatum and dentate nucleus of the cerebellum. As a result, these changes in cerebral blood flow during learning suggest that both the cerebellum and striatum play a major role in the learning of a serial reaction time task.

Furthermore, imaging studies have also shown that the cerebellum and striatum mediate learning during different stages of a serial reaction time task (Doyon *et al.*, 2002; Ungerleider, 2002; Penhune and Steele, 2012). In fact, Doyon *et al.* (2002) investigated, using a functional magnetic resonance imaging scan, the structural correlates of motor skill learning in a serial reaction time task and the timing of their neural activity. The functional magnetic resonance imaging data, which captures the metabolic activity of the brain by measuring the changes in the blood oxygenation level due to neural activity, revealed a significant increase in cerebellar activity in the early stages of learning. However, neural activity decreased in the cerebellum near the end of the training phase. In contrast, the striatum only showed significant activity near the end of training. Taken together, these results suggest that the cerebellum may be critical for the initial learning phase when the motor skill is being established, while the striatum may play a larger role during the later stages of learning and potential for the long-term retention of the skill (Doyon *et al.*, 2002; Ungerleider, 2002).

In fact, it has been proposed that the cerebellum receives sensory and motor information and contributes to error correction during the early phase of learning (Penhune and Steele, 2012). However, as learning proceeds, the contribution from the cerebellum decreases and the striatum, which is postulated to learn the association between the visual stimuli and required response, begins to compensate and account for the bulk of the learning (Penhune and Steele, 2012). In addition, as the movement is practiced over the course of training, the motor cortex also begins to compensate and store the representation of the learned movement sequence.

1.4 Feedback characteristics and motor learning

An important aspect of motor skill learning is the information that individuals receive regarding their movement after it has been completed. This type of information related to our movements is broadly defined as feedback. The role of feedback is to provide information that is essential not only for correcting errors, but also for improving the execution of the movement on the next attempt. As a result, feedback has been extensively investigated in the motor learning field and can be divided into two main categories: intrinsic feedback and augmented feedback (Salmoni *et al.*, 1984;

Magill and Anderson, 2017; Schmidt *et al.*, 2019).

Intrinsic feedback is performance-related information that learners naturally receive about their movement from various sensory modalities. For example, golfers know they missed the hole after a putt because they receive visual feedback that the ball did not go in the hole. In addition, they also receive auditory feedback when they hear the ball drop, or not drop, in the hole.

Another type of feedback that is provided about the movement in addition to inherent feedback is augmented feedback. Augmented feedback is information from an external source, such as coaches or instructors, and supplements intrinsic feedback. For instance, a golf coach can provide additional information to a golfer about their technique during the swing or the outcome of the putt. Augmented feedback can be further classified as knowledge of performance or knowledge of results (Schmidt *et al.*, 2019). The following sections will provide an overview of these two types of augmented feedback.

1.4.1 Knowledge of performance

Knowledge of performance is information that is provided to the learner about their movement characteristics that led to a performance outcome. This type of information is often presented by expert instructors or motion analysis software and helps the learner make the appropriate corrections to their movement characteristics, such as the positioning, timing, and speed of their movement. For example, a golfer might receive feedback from their coach about the placement of their hands on their club in order to improve their shot. In fact, researchers have shown that knowledge of performance feedback enhances learning of a range of motor tasks (e.g., Lindahl, 1945; Newell *et al.*, 1983; Young and Schmidt, 1992).

1.4.2 Knowledge of results

A great deal of research on motor skill learning and feedback has focused on information that learners receive regarding the outcome of their movement in relation to the task goal, known as knowledge of results (Salmoni *et al.*, 1984; Sigrist *et al.*, 2013).

Research studies investigating the impact of knowledge of results feedback during motor skill learning directly emerged from the work of Thorndike (1927) in experimental psychology. Thorndike (1927) proposed that knowledge of results feedback was an important component of learning and strengthening the bond between a behaviour and the task goal. To explore this, participants were assigned to either a binary-knowledge of results group or non-knowledge of results group and were randomly presented with strips of paper with different lengths between 3 and 27 cm. The goal of the experiment was to estimate their lengths using a known 10 cm strip of paper as the reference. Following the seven blocks of training, it was shown that the constant error in the binary-knowledge of results group reduced by 61% compared to the -7% reduction in the non-knowledge of results group. Therefore, providing evidence that knowledge of results feedback can facilitate learning by strengthening the bond between a behaviour and an environmental goal.

Within the motor learning research field, there are also different types of knowledge of results manipulations. One type of manipulation that has received considerable attention over the years is the frequency of providing knowledge of results (Bilodeau and Bilodeau, 1958; Salmoni *et al.*, 1984; Winstein and Schmidt, 1990). In fact, frequency of knowledge of results can be further subdivided into absolute frequency and relative frequency. Absolute frequency is the absolute number of times feedback is presented to the learner over the course of practice. In contrast, relative frequency refers to the percentage of trials that feedback is provided to learners.

Another type of manipulation that has been examined is the precision of knowledge of results (Salmoni *et al.*, 1984; Luft, 2014). There are three major categories of knowledge of results precision feedback based on the content of their information. Binary knowledge of results feedback only informs the learner about whether the outcome of their movement was correct or incorrect in relation to the task goal (e.g., a golfer receiving feedback about whether they got the ball in the hole or missed). Graded knowledge of results feedback provides different degrees of error based on categories (e.g., a golfer being informed that their shot was 'short', 'long', 'left', or 'right' of the hole). Lastly, finely graded knowledge of results feedback also provides the degree of error but with numerical units (e.g., a golfer being informed that they missed the hole by 2.5 cm of the left). Thus, graded knowledge of results feedback provides directional information whereas finely graded knowledge of results feedback provides both direction and magnitude information.

To investigate the relationship between the precision of knowledge of results and motor learning, Bennett and Simmons (1984) instructed participants to perform a linear positioning task using a slide bar. During the acquisition trials, participants were randomly divided into four groups: no-knowledge of results, binary knowledge of results, finely graded knowledge of results, and irrelevant knowledge of results feedback, followed by a retention test without feedback. Overall, the study revealed that participants receiving finely graded feedback produced significantly lower absolute error scores compared to the other three groups during acquisition and retention trials. Overall, studies have demonstrated that the presentation and manipulation of knowledge of results feedback regarding the outcome of a movement is an important aspect of improving the execution and learning of a motor task (Thorndike, 1927; Bennett and Simmons, 1984; Magill *et al.*, 1991; Sparrow and Summers, 1992; Weeks and Kordus, 1998).

1.5 Learning from feedback

Over the past several decades, researchers have begun to investigate the main theoretical aspects of learning in relation to feedback. It has been proposed that the major components are: (a) *the learning processes*, which refers to the behavioural mechanisms that support motor learning, and (b) *feedback characteristics*, specifying the form in which feedback is presented after the movement is complete. Below, I provide an overview of two common types of learning processes, known as error-based learning and reinforcement learning, as well as a discussion of feedback characteristic manipulations.

1.5.1 Error-based and reinforcement learning

During error-based learning, the sensorimotor system compares the movement outcome and movement goal, and estimates the direction and magnitude of the error, known as the sensory prediction error (Wolpert *et al.*, 1995). By doing so, the sensorimotor system knows not only whether the movement missed the goal, but also the specific way it was missed; allowing it to adjust future movement attempts and minimize the sensory prediction error.

A simpler type of learning process that is used to guide learning is a reinforcement learning process. During this process, the motor system learns which movements to execute in order to maximize the occurrences of rewards or minimize punishments (Kaelbling *et al.*, 1996; Sutton and Barto, 2018). For this type of learning process, the punishment and reward signals provide less information, that is, it only indicates the success or failure of the movement and does not provide information about the direction or magnitude of the error. As a result, the motor system adjusts the learner's movement on the next attempt, to maximize reward or minimize punishment, through trial and error (Luft, 2014).

1.5.2 Feedback characteristics

The feedback that learners receive related to their movements can be presented based on: performance content, motivational value, or a combination of both (Luft, 2014). Feedback based on performance content is often presented as either binary, graded, or finely graded error information. For example, an individual throwing darts at a bullseye could be informed that they either hit/missed the bullseye, were too far to the left, or were 2 cm to the left from the bullseye, respectively. In contrast, feedback according to motivational values relies on reward or punishment signals (e.g., gaining or losing money) in the form of binary information to enhance learning. In this case, the individual throwing darts would earn $1 (i.e., reward signal) for hitting the bullseye or lose $1 (i.e., punishment signal) for missing the bullseye. Finally, it can be presented as a combination of performance content and motivational information. For instance, after the dart has been thrown, they can receive finely graded error information as a monetary reward (e.g., earning $2 for being 2 cm from the bullseye) or punishment (e.g., losing $2 for being 2 cm from the bullseye).

For tasks that involve an error-based learning process, feedback is often presented as finely graded error signals to inform the learner about the direction and magnitude of the error (Luft, 2014; Cashaback *et al.*, 2017). However, it is important to emphasize that the same error can result in different adjustments. For example, when a golfer hits the ball to the right of the hole, the motor system can adjust the next movement by changing the movement of the arm, alignment of the shoulders, or through any combination of adjustments. In addition, even if an error occurred, its direction and magnitude might not be very clear. Consider the example of learning to snap your fingers. If you fail to make any sound, it is very difficult to know how you should adjust the positioning of your fingers to make a sound. In these instances, the reinforcement learning process is used to guide learning which uses a different form of feedback characteristic. In a reinforcement learning process, feedback is presented as reward or punishment based on binary information (Izawa and Shadmehr, 2011; Luft, 2014; Cashaback *et al.*, 2017). For instance, a golfer learning to putt will receive a reward feedback, for getting the ball in the hole, or a punishment feedback, for missing the hole. Most importantly, the feedback does not inform the learner about how close or far the shot was from the hole.

In fact, punishment and reward feedback have been shown to influence motor learning when applied in different paradigms and presented in various formats. Below, I highlight some experiments that have focused on motor learning through punishment and reward feedback.

1.6 Punishment and reward feedback

1.6.1 Motor adaptation

Recently, studies have begun to investigate punishment and reward feedback in the domain of motor adaptation to assess the relationship between punishment and reward feedback on motor learning. In fact, it has been proposed that punishment and reward feedback have dissociable effects on motor learning and retention, respectively (e.g., Abe *et al.*, 2011).

To test this, Galea *et al.* (2015) investigated the effects of punishment and reward on learning using a visuomotor rotation task. In particular, participants were instructed to make center-out reaching movements with a cursor toward visual targets. In the first phase, known as the baseline phase, the participants' baseline performance was measured prior to the perturbation being introduced. During the adaptation phase, a 30° counter clockwise visuomotor rotation was imposed onto the cursor and the aim was to adapt their reaching movements to the abrupt perturbation. For each adaptation trial, participants performed the movements with visual feedback of the cursor and endpoint angular error. In addition, the participants were presented with feedback after the completion of each adaptation trial depending on their group. The reward feedback group received money based on their graded endpoint error, the punishment feedback group lost money based on their graded endpoint error, and finally, the random positive feedback group randomly gained money irrespective of graded endpoint error. In the next phase, referred to as the no-vision phase, the perturbation, cursor feedback, and punishment and reward feedback were removed to assess for retention, where a drift back to baseline level reflected the degree of retention.

The results revealed that the punishment feedback group led to faster learning of the perturbation compared to the reward and random positive feedback group. In contrast, the reward feedback group displayed a significantly slower decay rate during the no-vision phase, indicating that reward led to greater retention.

However, given that the visual feedback of the cursor is provided throughout adaptation, participants are also receiving error feedback regarding their reaching movement. As a result, it is still unclear whether the distinct effects on learning and retention are attributable to punishment and reward feedback, respectively. To address this, Song *et al.* (2020) attempted to investigate the effects of punishment and reward feedback during a visuomotor rotation task without visual feedback of the cursor. Following the baseline phase, where the cursor followed the participants' reaching movement, a 50° counterclockwise rotation was imposed on the cursor without visual feedback. At the end of each adaptation trial, participants in the reward group received monetary points based on their endpoint error and the punishment group lost monetary points based on their endpoint error. Consistent with Galea *et al.* (2015), Song *et al.* (2020) found that punishment feedback led to faster adaptation compared to reward feedback. However, when participants were re-exposed to the 50° counterclockwise rotation without visual feedback to assess for retention, referred to as the re-adaptation phase, there was no significant difference between the groups.

There may be several different reasons for this inconsistent finding. A possible explanation is that participants in both studies received a combination of error feedback and punishment and reward feedback. As previously discussed, error feedback is typically presented as finely graded error information, while punishment or reward feedback is based on binary information, such as hitting or missing the target. Based on that distinction, Galea *et al.* (2015) provided a combination of error feedback, through visual feedback of the cursor and endpoint angular error, and punishment and reward feedback. In addition, the participants in both studies gained or lost money based on their endpoint angular error, which generates both types of feedback. Therefore, it becomes difficult to separate the effects of error-based and reinforcement learning processes.

It is important to dissociate these two types of feedback because the motor system likely relies on either feedback to enhance learning (Izawa and Shadmehr, 2011; Cashaback *et al.*, 2017). To examine how these signals contribute to motor learning, Izawa and Shadmehr (2011) manipulated the quality of feedback information during a motor adaptation task. All participants made reaching movements with a robotic hand to move a cursor to a visual target and experienced a gradual visuomotor rotation up to 8°. However, the Error group received full visual feedback of the cursor as well as a reward signal if the cursor hit the target (i.e., high-quality error information), the End-Point Error group only received visual feedback of the cursor at the endpoint of their movement along with a reward signal (i.e., intermediate-quality error information), and the Reward group was only provided with a reward signal without visual error feedback (i.e., low-quality error information). Although all groups adapted to the perturbation, adaptation in the error group was accompanied by a change in the estimation of their perceived hand position following their movement. Specifically, they estimated their hand position to be perturbed by 8.8°±-0.6° of the actual position. In contrast, the reward group displayed no significant change in their estimation and the end-point error group was intermediate relative to those groups. Based on those results, it is predicted that an error and reward signal can both enhance motor adaptation, but learners are much more dependent on a reward signal as the quality of the error signal decreases (Izawa and Shadmehr, 2011).

Cashaback *et al.* (2017) expanded this work by investigating the influence of reward signals and error signals, when presented in combination or separately, during a visuomotor reaching task. Specifically, participants reached a visual target using a robotic manipulandum and their hand was laterally shifted perpendicular to their movement. During each trial, the participants were either presented with an error feedback signal (i.e., visual end-point error of the cursor), reward feedback signal (i.e., target size doubled when cursor hit the target), or a combination of both. It was shown that both an error feedback and reward feedback signal enhanced adaptation during a reaching movement. However, learners appeared to rely on an error signal, independent of the information quality, when both feedback signals are provided. Given that Galea *et al.* (2015) and Song *et al.* (2020) failed to dissociate these two types of feedback, participants may have been relying on their endpoint error feedback instead of the punishment and reward feedback. As a result, it is difficult to draw clear conclusions about the supposed dissociable effects of punishment and reward feedback on learning and retention, respectively. In sum, these experiments highlight the importance of designing experiments that do not confound the provision of reinforcement feedback with the provision of error feedback when the goal is to investigate the influence of the former on motor learning.

1.6.2 Sequence learning

Beyond the domain of motor adaptation, researchers have also begun to investigate the behavioural relationship between reinforcement feedback and motor learning within the sequence learning field. For example, Wachter *et al.* (2009) used a serial reaction time task to understand the effects of punishment and reward feedback on learning and retention. Using a similar method to that of Nissen and Bullemer (1987), participants were presented with four visual stimuli and instructed to press one of four buttons associated with the stimuli as quickly and accurately as possible. In addition, the stimuli were presented in either random sequence or fixed repeating sequences to minimize explicit knowledge of the sequence.

The experimental task began with a familiarization period, consisting of four blocks of random sequences without feedback. In addition, each participant's criterion response time was calculated based on their median response time on the last block of familiarization. After familiarization, the training period began where participants were randomly allocated to a reward, punishment, or control group and received feedback based on their criterion response time. Specifically, the reward group received monetary points after each trial if their response time was faster than their criterion response time. In contrast, the punishment group lost monetary points after each trial if their response time was slower than their criterion response time. Furthermore, both groups were presented with either a red or green visual stimulus indicating that their response time was slower or faster than their criterion response time, respectively. The control group received an equal number of green and red stimuli and were informed that they were not based on their performance. Following the training period, participants performed the last four blocks without feedback.

Throughout the experiment, response time gain, the difference between the response time of each trial and the criterion response time, was used as an index of changes in performance (Wachter *et al.*, 2009).

During the training block, punishment feedback led to a significant drop in response time gain, indicating a faster reaction time, compared to the reward and control groups. In addition, all groups displayed an increase in absolute response time gain during the last several blocks without feedback. However, the reward group had a significantly larger absolute response time gain compared to the punishment and control group. This suggested that punishment feedback enhanced the learning component of the serial reaction time task, whereas reward feedback enhanced the retention component.

Steel *et al.* (2016) also investigated the impact of punishment and reward feedback during a serial reaction time task. During familiarization, all the participants performed three blocks of random sequences without feedback and their criterion response time was calculated based on the last block of familiarization. Following familiarization, the feedback period began with the training phase flanked by a pre- and post-training phase. During training, the visual stimuli were presented in six fixed-sequence blocks, whereas pre- and post-training consisted of three blocks in random–fixed–random sequence order. The difference in response time between the average of the two random blocks versus the fixed block was used to determine sequence-specific skill learning. Also, as participants progressed through the feedback period, their criterion response time was continuously re-calculated after each block Similar to Wachter *et al.* (2009), participants were also divided into a reward, punishment, or control group; however, feedback was presented based on their updated criterion response time. To assess retention, participants also performed 1-hour, 24-hours, and 30-day tests after the training period without feedback.

The results revealed that the mean response time for the punishment group significantly decreased across the training phase compared to the reward and control groups. However, all groups displayed a significant reduction in their mean response time during the 1 hour, 24 hour, and 30 day retention tests with no difference between them. These findings demonstrate that punishment feedback likely enhances the learning aspect of a sequence learning task, but neither reward nor punishment feedback enhances the retention component. It also suggests that the benefits of reward feedback may not be as robust as previously suggested by other researchers (Abe *et al.*, 2011; Galea *et al.*, 2015; Wachter *et al.*, 2009).

However, it is important to acknowledge that these results may not be as definitive given the methodological limitations in the motor adaptation studies as discussed in the previous section. Overall, punishment and reward feedback appear to have some beneficial impact on motor learning within the different domains. In fact, providing a combination of punishment and reward feedback may produce the proposed benefits on both learning and retention.

1.6.3 Neural correlates of punishment and reward feedback

Over the past several years, interest has grown in identifying the neural correlates underlying the effects of reinforcement feedback. It is proposed that the distinct behavioural effects of reward and punishment feedback are associated with specific neural networks and cortical structures.

For reward feedback, it has been postulated that the neurotransmitter dopamine, primarily released by dopaminergic neurons in the midbrain, plays an important role in reward-based learning (Wickens *et al.*, 2003; Wachter *et al.*, 2009; Galea *et al.*, 2015). Specifically, these dopaminergic neurons project and release dopaminergic signals in response to a reward stimulus into the striatum, a subcortical structure of the brain known for regulating movement and responses to rewarding stimuli (Wickens *et al.*, 2003; Wachter *et al.*, 2009). This in turn activates the neurons within the striatum which project into the motor regions of the cerebral cortex and influence the motor commands. Most importantly, the continuous recruitment and integration of neurons within these cortical regions over the course of learning can strengthen and refine these neural connections; allowing the motor cortex to generate the appropriate motor commands to gain a reward (Wickens *et al.*, 2003; Bromberg-Martin *et al.*, 2010).

In contrast, punishing stimuli can alter the dopamine concentration in downstream cortical regions and lead to behavioural changes to avoid those stimuli in the future (Bromberg-Martin *et al.*, 2010; Jean-Richard-Dit-Bressel *et al.*, 2018). It is suggested that serotonin, a neurotransmitter mainly produced by neurons originating in the raphe nuclei, plays an important role in regulating behaviour that leads to a punishing experience (Jean-Richard-Dit-Bressel *et al.*, 2018). In fact, the neurons in the raphe nuclei project into the striatum and release serotonin signals that inhibit the reward-dopamine pathway. In addition, the raphe nuclei neurons also project to additional cortical structures within the temporal lobe and prefrontal cortex, commonly implicated in punishment-based learning (Jean-Richard-Dit-Bressel *et al.*, 2018), and influence motor commands to avoid punishing stimuli.

To investigate the cortical structures involved in reward and punishment feedback, Wachter *et al.* (2009) studied another group of participants using the same serial reaction time procedure as described above. However, an exception to the procedure was that all participants underwent a functional magnetic resonance imaging scan. The purpose of the scans was to measure the participant's brain activity during trials with and without feedback.

Following the analysis of the functional magnetic resonance imaging data, the reward group showed a significant increase in activation in the striatum, nucleus accumbens, and prefrontal cortex during the feedback trials. The punishment group, however, showed a significant decrease in activation in the striatum, but increased activation in the insula, a small cortical structure located within the lateral sulcus, and regions of the prefrontal cortex. The control group showed no difference in activation in those regions throughout the experiment. Therefore, these results demonstrate that the behavioural effects of reward and punishment feedback are likely processed by different neural structures within the brain.

To further investigate the neural mechanisms underlying the effects of reinforcement feedback, Steel *et al.* (2019) examined the changes in the neural connectivity after training with reward and punishment feedback. Using the same sequence learning design as Steel *et al.* (2016), all participants performed a familiarization block, pre-test, training, post-test, and 24 hour retention test, with feedback only being presented during training. In addition, pre- and post-resting state functional magnetic resonance imaging scans were also performed to assess the effects of reward and punishment feedback on neural connectivity.

The analysis of the scans revealed that the neural connectivity for the reward feedback group increased between the premotor cortex with the striatum, cerebellum, and supplementary motor area (located in the dorsomedial region of the frontal lobe), but decreased between the premotor cortex with the medial temporal lobe and inferior frontal gyrus (a small portion of the frontal lobe). For the punishment feedback group, the scans showed the opposite pattern; the neural connectivity decreased between the premotor cortex with the striatum, cerebellum, and supplementary motor area, but increased between the premotor cortex with the medial temporal lobe and inferior frontal gyrus.

Taken together, these findings provide evidence that reward and punishment feedback have differential effects on neural connectivity after training. Most importantly, it also indicates that the differential behavioural effects of reward and punishment feedback may be mediated by the interaction of distinct neural networks and cortical structures (Steel *et al.*, 2019), rather than being associated with a single cortical region within the brain (Wickens *et al.*, 2003; Jean-Richard-Dit-Bressel *et al.*, 2018; Steel *et al.*, 2019).

1.7 Combining punishment and reward feedback for motor learning

Within the sports world, coaches are always attempting to increase their athlete's performance. A common approach that is recommended by coaching guidelines is providing reward and punishment performance feedback during their training (Burton and Raedeke, 2008; Williams and Krane, 2015; Chen *et al.*, 2018). By doing so, the athlete is able to receive valuable information regarding their performance in relation to the task goal and improve the execution of their movement on the next attempt.

The manuals within the sports coaching domain suggest that reward feedback enhances performance by strengthening the bond between the desired movement and the reward (Burton and Raedeke, 2008; Williams and Krane, 2015). In addition, it may also increase the athlete's belief in succeeding at the task, and thus improve their performance. However, if the reward is not based on performance accomplishments or is provided excessively, it can potentially make the athlete feel like they are being manipulated and negatively impact their performance. Alternatively, punishment feedback likely increases an athlete's performance by reducing undesirable behaviour during training. Unfortunately, punishment may lead to a negative attitude towards the coach if it is overly used during training, which can in turn negatively impact the training environment and the athlete's performance.

Despite the potential side effect, the handbooks do not recommend that coaches should only provide reward feedback and avoid punishment feedback. Instead, they suggest using a combination of punishment and reward to improve an athlete's performance (Williams and Krane, 2015). Specifically, according to the coaching literature, reward feedback should be provided early in training and punishment feedback should only be occasionally presented near the end of training. Based on this prediction from the coaching literature, this thesis will investigate the utility of providing both punishment and reward feedback during training using a transition scheduling approach.

Chapter 2

Introduction

Humans display a remarkable ability to learn a variety of motor skills, such as riding a bicycle, swimming the butterfly, or throwing a fastball. An essential component of learning any motor skill is the information learners receive about their movements, known as feedback. The role of feedback is to provide information that not only helps the learner identify their movement error in relation to the task goal, but also helps guide them towards an appropriate movement solution (Magill and Anderson, 2017; Tresilian, 2012). While feedback frequency manipulations have historically received considerable attention in motor learning research (for reviews see Salmoni *et al.*, 1984; Swinnen, 1996; Anderson *et al.*, 2019), the characteristics of how feedback is presented to learners has seen a surge of interest in the last decade (see Luft, 2014, for a review). Feedback characteristics can reflect performance-related information and be provided as binary (e.g., "hit", "miss"), graded (e.g., "too short", "too long"), or finely-graded (e.g., -12.64 cm, 8.03 cm) information. Feedback can also be provided to reflect a motivational characteristic through either reward (e.g., +5 cents) or punishment (e.g.,

-5 cents). Finally, feedback can be provided in a way that combines both performance-related information and motivational value (e.g., monetary gain or loss depends on the amount of error).

The impact of receiving punishment or reward feedback on motor learning has been a popular area of research as of late (for a review see Lohse *et al.*, 2019; Wachter *et al.*, 2009; Abe *et al.*, 2011; Galea *et al.*, 2015; Steel *et al.*, 2016; Izawa and Shadmehr, 2011; Cashaback *et al.*, 2017, 2019). In a visuomotor adaptation task where participants had to learn to update their reaching direction to compensate for a rotation applied to a cursor that represented the movement of their unseen hand, Galea et al. (2015) found that punishment feedback accelerated learning while adapting to the rotation, but it was reward feedback that resulted in greater retention of what was learned during training. Similar findings have also been found during sequence learning, such as the serial reaction time task (e.g., Abe *et al.*, 2011; Wachter *et al.*, 2009). In contrast, Steel *et al.* (2016) found the benefit of punishment feedback on learning the serial reaction time task, but did not find that feedback type differentially impacted short (6 hours) or long term (24 hour and 30 day) retention. Their results not only suggested that punishment feedback may have more reliable effects as compared to reward feedback (also see Song and Smiley-Oyen, 2017; Song *et al.*, 2020), but that punishment feedback effects may be task-dependent—beneficial for serial tasks, but detrimental for continuous tasks (e.g., Abe *et al.*, 2011).

Although the available data surrounding punishment and reward feedback is mixed, the possibility that punishment and reward feedback have dissociable effects on learning and retention, respectively, is interesting from both a theoretical and applied perspective. It highlights that the behavioural outcomes associated with punishment and reward feedback not only are likely mediated by separate neural substrates (Wachter *et al.*, 2009), but also preferentially engage fast and slow learning systems (Peterson and Seger, 2013; Wachter *et al.*, 2009). Practically, it raises the question of whether combining punishment and reward feedback during training would benefit both learning and retention; exceeding that found when provided in isolation. When designing such a practice environment, it is not entirely clear how to best combine punishment and reward feedback as neither motor learning theory (Adams, 1971; Schmidt, 1975; Smith *et al.*, 2006; Wolpert and Flanagan, 2016) nor prevailing perspectives on scheduling feedback (Salmoni *et al.*, 1984) have directly addressed this issue.

Within the coaching literature, however, the use of combined punishment and reward feedback schedules have actually been promoted as a method to improve an athlete's skill acquisition (Warren, 1983; Burton and Raedeke, 2008; Williams and Krane, 2015; Chen *et al.*, 2018). One approach for combining the delivery of both punishment and reward feedback is through transition schedules. Based on the coaching literature, reward feedback should be provided immediately and often in the early stages of practice to promote a bond between the reward and appropriate movement solutions. Punishment feedback should only be introduced later in practice as a means to eliminate less efficient and less effective movements. This order effect likely relates to wanting to prevent athlete's from developing resentment towards their coaches and dissatisfaction with their learning environment if punishment feedback is provided too early and too often (Warren, 1983; Williams and Krane, 2015). Alternatively, moving from punishment feedback to reward feedback could be the more effective transition schedule based on the laboratory-based motor learning literature (e.g., Galea *et al.*, 2015; Wachter *et al.*, 2009; Song *et al.*, 2020; Steel *et al.*, 2016). The faster learning rate found with punishment feedback results in participants reaching a steady-state of repeating successful actions earlier in practice (Huang *et al.*, 2011). This means a greater portion of training is spent repeating these actions, which are further reinforced when the switch to reward feedback occurs. This perspective is also consistent with a stage-dependent role of feedback during practice (e.g., Carter *et al.*, 2014, 2016).

The purpose of the present experiment was to test the utility of combining punishment and reward feedback through a transition scheduling approach. When determining when the transition should happen, we opted for the midpoint in training for two reasons. First, past work showed this to be an effective point to transition from constant to variable practice during motor skill learning (Lai *et al.*, 2000). Second, it ensured that independent of transition order, an equal number of practice trials had the potential to be rewarded and/or punished. By providing half of the participants with reward-to-punishment feedback and the other half with the reverse order, we were able to directly test competing predictions about feedback order between the coaching and motor learning literatures. Based on the coaching literature, reward-to-punishment feedback should be the most effective. However, in line with the motor learning literature, we predicted that it would be the punishment-to-reward feedback order that would be more beneficial for both learning and retention.

Chapter 3

Methods

We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study (Simmons *et al.*, 2012). The experimental design and analyses were preregistered using the Open Science Framework Registries and can be viewed here: https://doi.org/10.17605/osf.io/75byt.

3.1 Group sequential design

The present experiment was modeled off the design of Steel *et al.* (2016). As such, when defining our smallest effect size of interest we first calculated the effect size that Steel *et al.* (2016) had 33% power to detect (Simonsohn, 2015). This returned a Cohen's d of 0.62. We then took a more conservative value of d = 0.4 and set this as our smallest effect size of interest for when planning our sequential analysis (Lakens, 2021). Sequential analyses are a more efficient approach to hypothesis testing than only analyzing the data a single time once the entire planned sample size has been collected (Dodge and Romig, 1929; Wald, 1945; Lakens, 2014; Lakens *et al.*, 2021a).

The parameters of our sequential analysis were set at $\alpha = 0.05$, $\beta = 0.2$, $d = 0.4$, and we used the O'Brien-Fleming alpha spending function to control our Type 1 error rate across our two planned interim analyses (at 33 and 66% of total sample size) and one final (100% of sample size) analysis. This resulted in sample sizes of 68, 134, and 202 participants and corresponding alpha levels of 0.0002, 0.012, and 0.046 for interim analysis 1, interim analysis 2, and the final analysis, respectively.

3.2 Participants

To reach the necessary sample size of 68 participants for interim analysis one, 72 participants had to be recruited, with four having to be removed.[1] The 68 participants ($M_{age} = 21.78$ years, $SD = 1.79$, 30 females) included in the first interim analysis all self-reported being right-hand dominant and having normal or corrected-to-normal vision. The order of receiving punishment and reward feedback was counterbalanced across participants by randomly assigning half of the participants ($M_{age} = 21.62$ years, $SD = 1.61$, 14 females) to receive reward-to-punishment feedback and the other half of participants ($M_{age} = 21.94$ years, $SD = 1.97$, 16 females) the reverse order. Prior to the start of the online experiment, participants gave informed consent through LimeSurvey and the experiment was approved and conducted in accordance with the University's Research Ethics Board. Participants received entry into a lottery (see **Task** section below) to win one of six gift cards valued at $50 for their participation.

[1] One participant was removed because they did not complete the experiment and three participants had incomplete or no data collected due to server-related issues.

3.3 Task

The serial reaction time task was created using jsPsych (de Leeuw, 2015) and was administered online using Pavlovia (https://pavlovia.org/). The task closely resembled the version used by Steel *et al.* (2016), which was a variation of the original version (Nissen and Bullemer, 1987). Participants were presented with four white boxes arranged horizontally in the centre of their laptop or computer screen (see Figure 3.1A). Each white box was associated with one of four keys (H-J-K-L) on their keyboard. Participants were instructed to position their index, middle, ring, and pinky fingers of their right-hand on the H, J, K, and L keys, respectively. When one of the squares changed to black, participants were instructed to press the corresponding key on their keyboard as quickly and accurately as possible. A trial consisted of a single key press in response to a visual stimulus and a trial ended once a participant pressed a key or 5000 ms elapsed without a response (i.e., trial timed out). An inter-trial interval of 200 ms was used during which the four empty squares appeared on the screen.

The created jsPsych program controlled the presentation of all instructions and stimuli, the timing of the experimental protocol, and recorded and saved the data on the Pavlovia server for later retrieval and offline analysis.

3.4 Procedure

Participants completed two online data collection sessions on consecutive days. Session 1 consisted of four phases: familiarization (1 block), pre-test (3 blocks), training (6 blocks), and post-test (3 blocks). Blocks consisted of 96 trials with the stimuli presented in either a fixed sequence or a random sequence. During fixed sequence

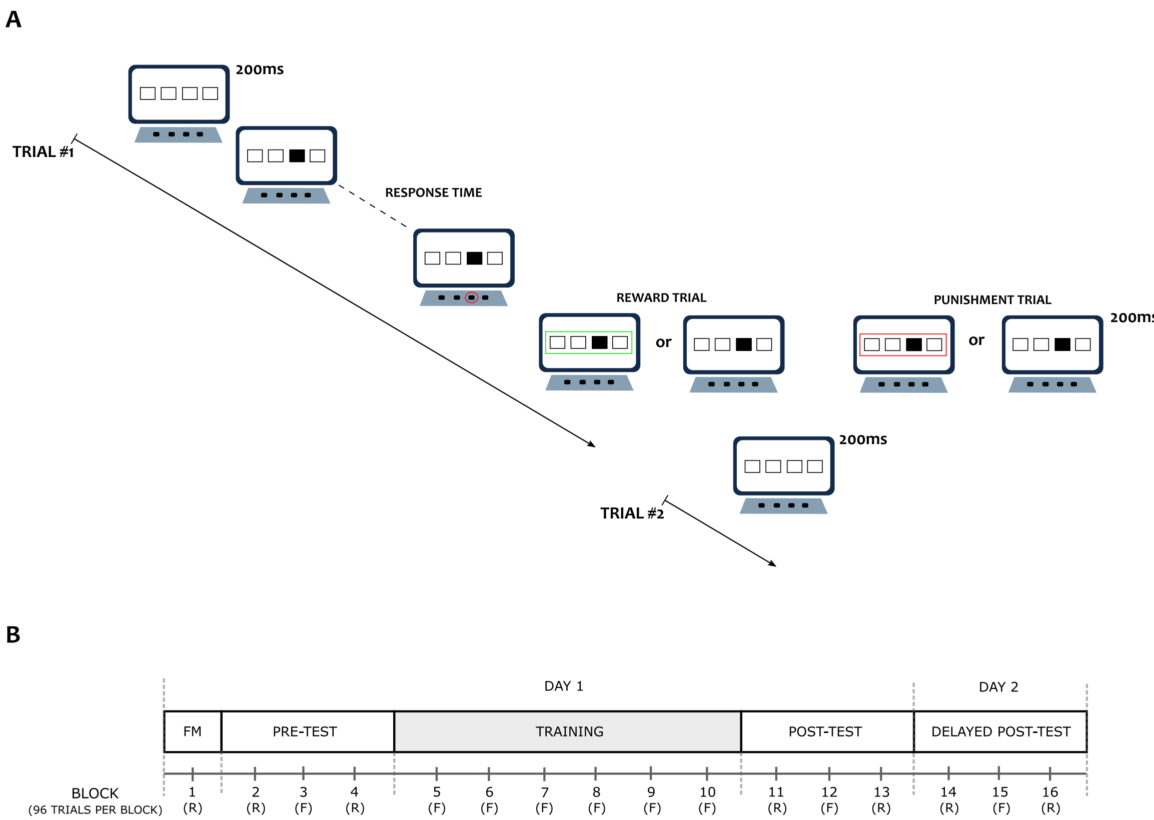

Figure 3.1: Overview of the serial reaction time task and procedure. **(A) Task setup.** Participants placed the index, middle, ring, and pinky fingers of their right hand on the H,J,K,L keys on their keyboard, respectively. Four white boxes were first shown to participants. After a fixed 200 ms interval, one box would change to a black box. This cued the participant to press the corresponding key as quickly and accurately as possible. On rewarded trials, a green frame only appeared if the participant's response was accurate and faster than their performance criterion. Otherwise, a blank frame appeared around the white boxes. On punished trials, a red frame only appeared if the participants' response was inaccurate or slower than their performance criterion. Otherwise, a blank frame was presented. **(B) Overview of experimental sessions.** Participants completed two testing sessions on consecutive days. Day 1 consisted of 4 phases. Familiarization (FM) consisted of 1 block, Pre-test consisted of 3 blocks, Training consisted of 6 blocks, and Post-test consisted of 3 blocks. All blocks had 96 trials and feedback was only available during the training blocks. The trials in a block followed either a random (R) sequence or a fixed (F) sequence. Day 2 had only one phase, which was the delayed Post-test (i.e., retention). This test consisted of 3 blocks and was identical to the Pre-test and Post-test phases.

blocks, the stimuli appeared according to a fixed 12 item sequence repeated eight times. Participants were randomly assigned to one of four possible patterns[2] and performed this same pattern for all fixed sequence blocks. Each fixed sequence block began at a different position within the repeating sequence to help reduce the development of explicit sequence knowledge (Schendan *et al.*, 2003). During random sequence blocks, the stimuli appeared based on a pseudorandomly generated pattern such that the same stimuli was never presented on consecutive trials.

The familiarization phase included one random sequence block. The pre-test and the post-test both consisted of three blocks: a random sequence block followed by a fixed sequence block followed by another random sequence block. No feedback was provided during familiarization, pre-test, or post-test. Before the first block in the training period, participants were informed that their performance on each trial would influence their total number of points and that their score at the end of training could earn extra entries in the gift card lottery. The training phase consisted of six fixed sequence blocks. Participants were given a 30 s break between all blocks in the experiment. During these breaks, the phrase "Nice job, take a breather" was displayed on the screen for 25 s. This message was then replaced with a black cross for the final 5 s. In the training phase, the participant's total number of points was also displayed on the screen during the 30 s break.

Feedback was only provided to participants during the training phase. Half of the participants received punishment feedback for the first three blocks followed by reward feedback for the last three blocks. The other half of participants received reward feedback for the first three blocks and punishment feedback for the last three blocks.

[2] Pattern 1: J-L-J-H-K-L-H-J-K-H-L-K. Pattern 2: K-L-K-H-J-L-H-K-J-H-L-J. Pattern 3: K-L-J-K-H-J-H-L-K-J-L-H. Pattern 4: K-L-H-J-L-K-H-L-J-H-K-J.

Feedback was provided on a trial-by-trial basis based on the participant's performance relative to a performance criterion. The initial criterion was computed based on each participant's median performance in the final pre-test block. This criterion was updated after each training block to encourage continuous improvement and keep participants engaged in the task. Punishment feedback consisted of a red frame around the four white boxes and was only provided when a response was incorrect or slower than the performance criterion. Trials that received punishment feedback were also penalized -0.10 points. Reward feedback consisted of a green frame around the four white boxes and was only provided when a response was correct and faster than the performance criterion. Rewarded trials earned participants +0.10 points. Points were only lost or gained during training and all participants began with 25 points. Participants were not made aware of their performance criterion or the point value lost or gained on punished or rewarded trials, respectively.

Session 2 had a single phase, which was the delayed retention test (3 blocks). The delayed retention phase consisted of three blocks: a random sequence block followed by a fixed sequence block followed by another random sequence block. No feedback was provided in the retention phase.

3.5 Data processing and analyses

Our primary measure of interest was response time, which was defined as the time between stimulus onset and the completion of the participant's key press.[3] The first key that was depressed after stimulus presentation was considered the participant's response. Consistent with Steel *et al.* (2016), data were first screened for any participants that were unresponsive or inaccurate on greater than 50% of the trials. No participants were removed after this screening process. For pre-test, post-test, and delayed retention, the difference in response time between the mean of the two random sequence blocks and the mean of the single fixed sequence block was used to measure sequence specific retention (Robertson, 2007; Steel *et al.*, 2016). For the training period, mean response time was calculated for each of the six fixed sequence blocks of 96 trials.

3.5.1 Primary statistical analyses

To test our prediction the punishment-to-reward feedback transition would be more effective for both learning and retention compared to reward-to-punishment feedback, we compared the distribution of response times on correct trials in training (Prediction 1) and in retention (Prediction 2). For both of these analyses, the 20% trimmed means of response times were calculated for each participant. Next, a shift function (e.g., Rousselet *et al.*, 2017) of the training or retention data was generated. The shift function compares the difference between two groups at each decile of their distribution via 95% confidence intervals and plots them as a function of one group. A family-wise error rate of 5% was maintained using Hochberg's method (Hochberg, 1988). This strategy ensures that the probability of at least one false positive will not exceed the nominal level as long as the nominal level is not exceeded for each quantile (Wilcox *et al.*, 2014). Shift functions are, overall, a more powerful and robust approach to understand whether groups of observations differ (Rousselet *et al.*, 2017;

[3]Past research has also defined their primary measure in this way but referred to it as reaction time. Reaction time, however, is the time between stimulus onset and the initiation of a response (Magill and Anderson, 2017). In fact, the inability to partition response time into its component parts, reaction time and movement time, has been argued as a limitation of the serial reaction time task (Krakauer *et al.*, 2019).

Rousselet and Wilcox, 2020).

For both analyses, the null hypothesis was rejected if the groups were significantly different at any decile. The alternative hypothesis that transitioning from punishment-to-reward feedback was more effective for learning (Prediction 1) and retention (Prediction 2) was accepted if those receiving this order had significantly shorter response times at any decile and no significantly longer response times at any decile.

Our design also allowed us to run a well-powered, within-subjects analysis of the typically found learning advantage of punishment feedback as compared to reward feedback during training. Due to the within-subjects nature of this test, we compared the distribution of response times from training during the Punishment condition to the Reward condition using a hierarchical shift function (e.g., Rousselet and Wilcox, 2020). This analysis was conducted using the following steps. First, sample deciles were computed for each participant and each condition using the Harrel-Davis quantile estimator (Harrell and Davis, 1982). Second, we subtracted the deciles for the Reward condition from the Punishment condition for each participant. Third, the distribution of differences at each decile was analyzed using a one sample t-test. Lastly, the resulting p-values were evaluated using the critical values from Hochberg's method (Hochberg, 1988) to control family-wise error rate. The null hypothesis was rejected if the conditions were significantly different at any decile using Hochberg's method. The alternative hypothesis that punishment feedback was more effective for learning during training than reward feedback was accepted if the punishment condition had significantly shorter response times at any decile and no significantly longer response times at any decile.

3.5.2 Secondary statistical analyses

We also ran more traditional analyses of (co)variance to facilitate comparisons with past work. For ANOVA, univariate outliers were screened using the median absolute deviation technique with a pre-specified threshold of three (Leys *et al.*, 2013). For ANCOVA, both univariate and multivariate outliers were screened. Multivariate outliers were assessed using the minimum covariance determinant approach with a pre-specified alpha set to $p = 0.01$ (Leys *et al.*, 2019). Sensitivity analyses for both the training and retention data were run with all outliers removed. Results showed no significant changes with or without outliers included.

To test for a learning advantage of Punishment-to-Reward feedback, mean response times were analyzed in a mixed 2 (Order) x 6 (Block) ANOVA with repeated measures on Block. To test for a retention advantage of Punishment-to-Reward feedback, mean response times were analyzed in a mixed 2 (Order) x 2 (Test: Post-test, Delayed retention) ANCOVA controlling for pre-test. To test for a learning advantage of punishment feedback, mean response times were analyzed using a 2 (Condition) x 3 (Block) repeated measures ANOVA.

Chapter 4

Results

4.1 Training

4.1.1 Primary analysis

The shift function on training data (Figure 4.1) revealed no significant difference at any deciles between the Punishment-to-Reward and Reward-to-Punishment feedback orders. In addition, the hierarchical shift function (Figure 4.2) revealed no significant advantage of punishment feedback over reward feedback during training at any decile.

4.1.2 Secondary analyses

During training mean response times (Figure 4.3) decreased across training blocks for the Punishment-to-Reward and Reward-to-Punishment groups, $F(4.32, 285.30) = 27.100$, $p < .001$. Holm's post-hoc comparisons revealed that Block 1 was significantly longer than Blocks 2-6 and Blocks 2-4 was significantly longer than Block 6. However, there was no significant main effect of Order, $F(1, 66) = .328$, $p = .569$, or a significant

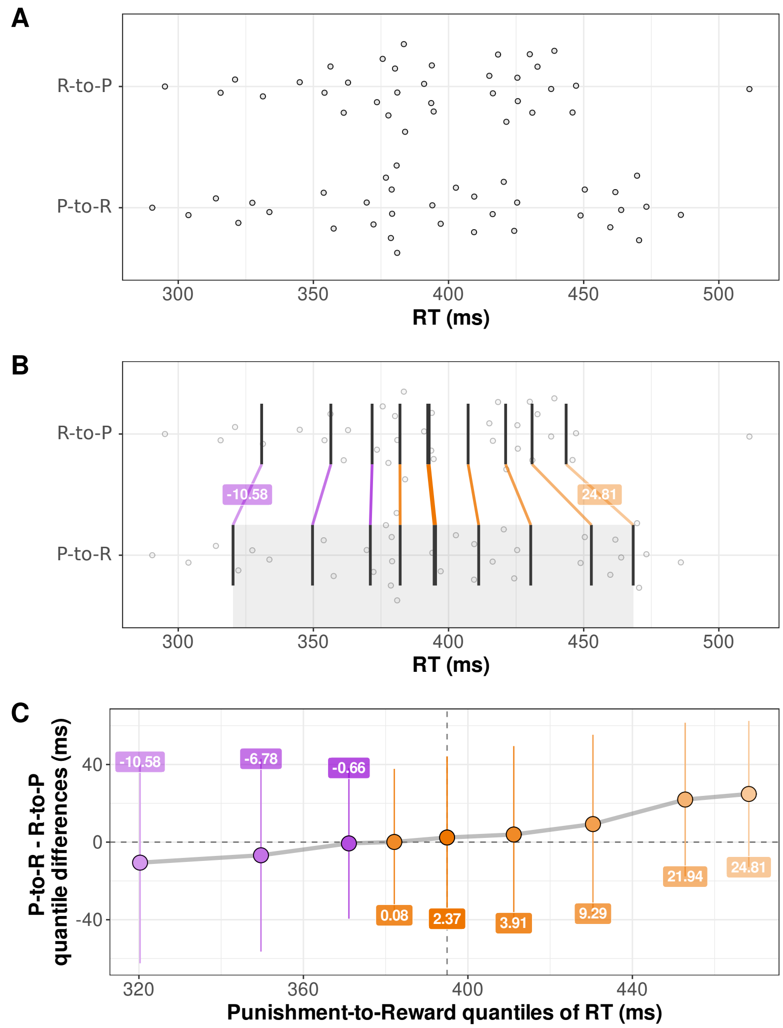

Figure 4.1: Feedback order distributions and the associated shift function during training. **(A) Scatterplot.** Response time is shown on the x-axis with Punishment-to-Reward feedback (P-to-R) and Reward-to-Punishment Feedback (R-to-P) on the y-axis. **(B) Scatterplot with deciles of the distribution.** Overlayed on the scatterplot from (A), are data for both groups which has been divided into deciles, shown through faded long vertical black lines. The thicker black line represents the median. The Reward-to-Punishment group's deciles were subtracted from the Punishment-to-Reward group's matching deciles and connected by coloured lines. If the difference is negative, indicating longer response times for the Reward-to-Punishment feedback, the connecting line is purple. If the difference is positive, indicating longer response times for the Punishment-to-Reward feedback, the connecting line is orange. The superimposed values indicate the difference for deciles 1 and 9. **(C) Shift function.** Punishment quantiles of RT in (ms) is shown on the x-axis, and the difference between Punishment-to-Reward and Reward-to-Punishment groups for each decile is plotted as a function of Punishment-to-Reward group's deciles. The superimposed values indicate the difference for each decile. The coloured vertical lines represent adjusted 95% confidence intervals.

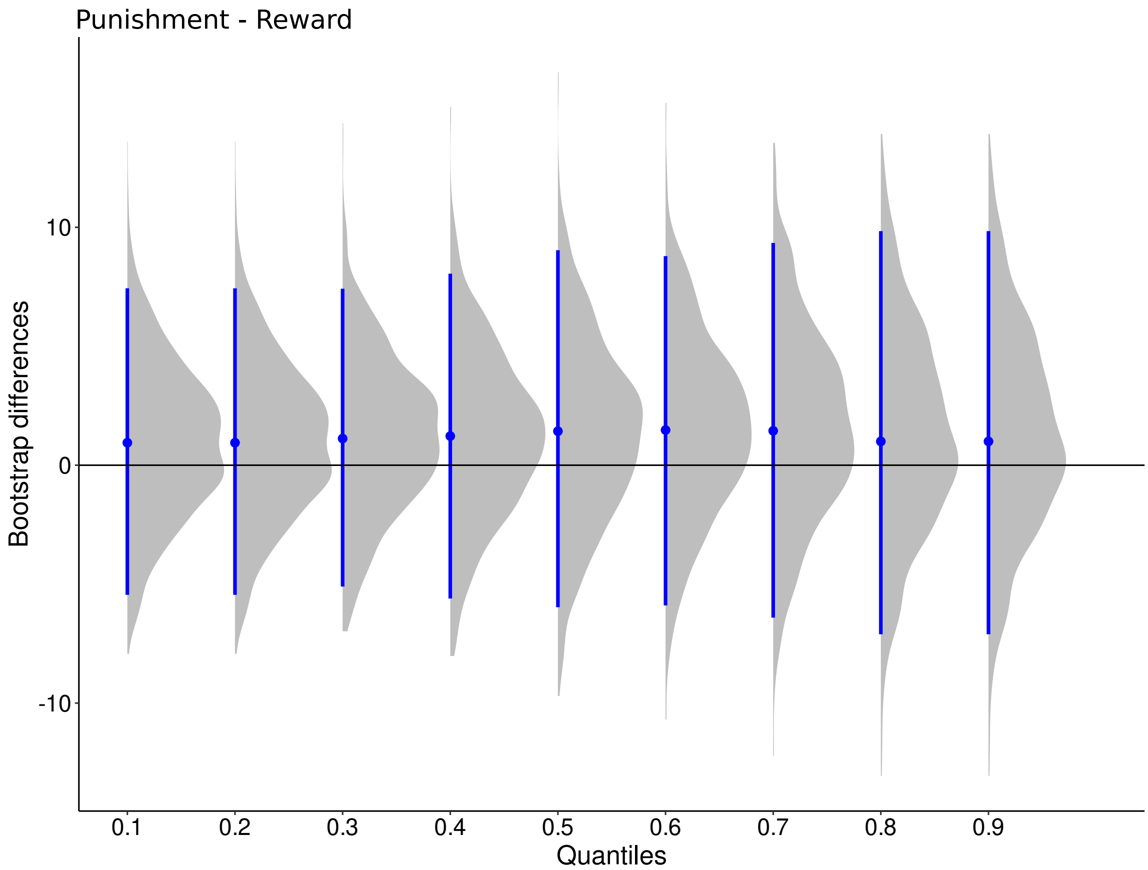

Figure 4.2: Hierarchical shift function for punishment versus reward feedback during training. Punishment and reward feedback were compared using a hierarchical shift function for each participant. For each decile, the 95% bootstrap confidence intervals and posterior distributions are illustrated in grey. The blue dots represent the 20% trimmed means and the vertical lines are 95% confidence intervals.

Block X Order interaction, $F(4.32, 285.30) = 2.250$, $p = .059$. The equivalence test revealed that the estimated effect was not significantly smaller than our smallest effect size of interest, $t(62.16) = -1.120$, $p = 0.134$. The results were not affected after two univariate influential cases were removed.

During training mean response times decreased across the three training blocks for both types of feedback conditions, which was supported by a significant main effect of Block, $F(1.89, 126.38) = 46.146$, $p < .001$. Holm's post-hoc comparisons revealed that Block 1 was significantly slower than Blocks 2 and 3. Neither the main effect of Condition, $F(1, 67) = 1.12$, $p = .294$, or the Block X Condition interaction, $F(1.82, 121.82) = .461$, $p = .613$, were significant. The equivalence test was significant, revealing that the estimated effect was smaller than our smallest effect size of interest, $t(67) = 4.376$, $p < 0.001$. Again, the results were not affected after two univariate influential cases were removed.

4.2 Retention

4.2.1 Primary analysis

Similar to training, the shift function on the retention data (Figure 4.4) failed to reveal a significant difference between the Punishment-to-Reward (labelled P-to-R) and Reward-to-Punishment (R-to-P) feedback groups at any decile.

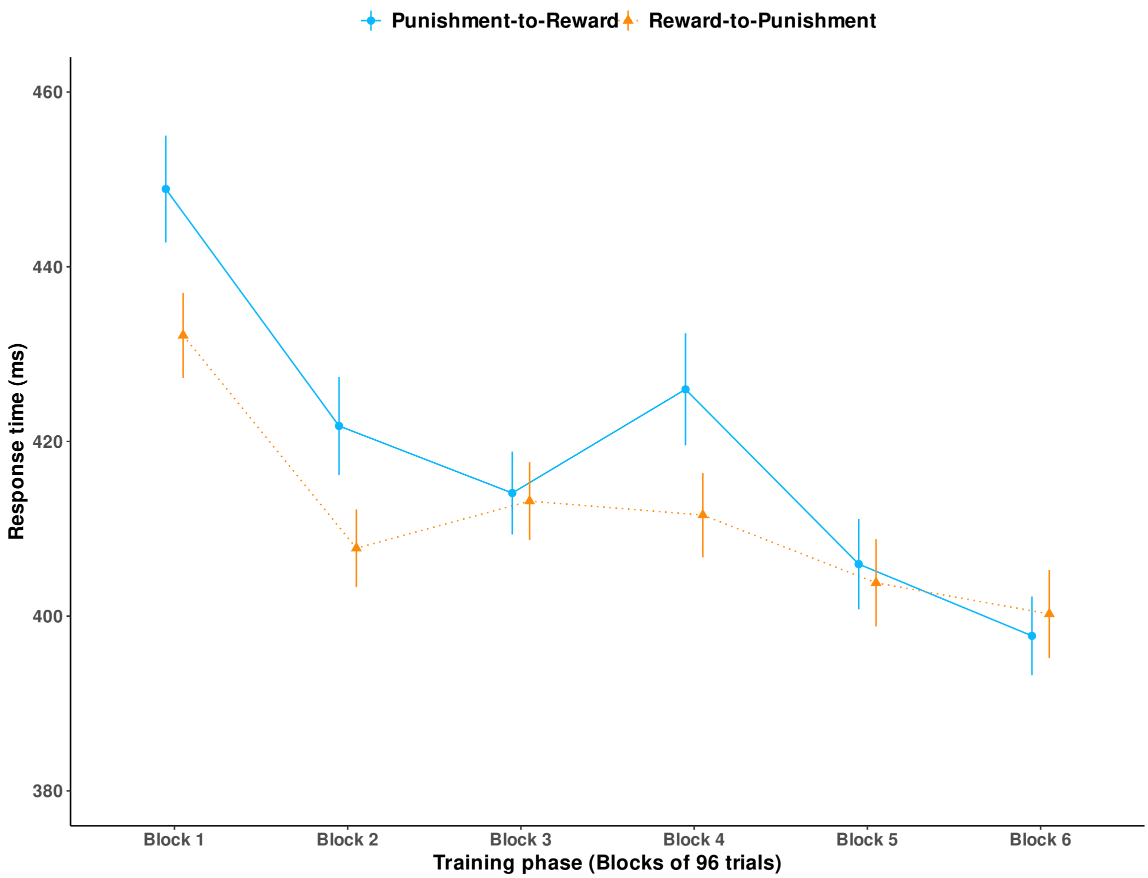

Figure 4.3: Mean response times during the training period. The mean response time (ms) across six training blocks of Punishment-to-Reward (blue) and Reward-to-Punishment (orange) feedback orders. Error bars indicate 95% confidence intervals.

4.2.2 Secondary analysis

Response times (Figure 4.5) were numerically faster in post-training and retention compared to pre-test. At all time points, the differences between the Punishment-to-Reward and Reward-to-Punishment feedback groups were minimal The main effects of Test, $F(1, 65) = .002$, $p = .962$, and Order, $F(1, 65) = .050$, $p = .824$, as well as the Test X Order interaction, $F(1, 65) = .031$, $p = .860$, were not significant. The equivalence test revealed that the estimates of the effect was not significantly smaller than our smallest effect size of interest, $t(66) = 1.510$, $p = 0.068$. These results were not affected after two univariate and five multivariate influential cases were removed.

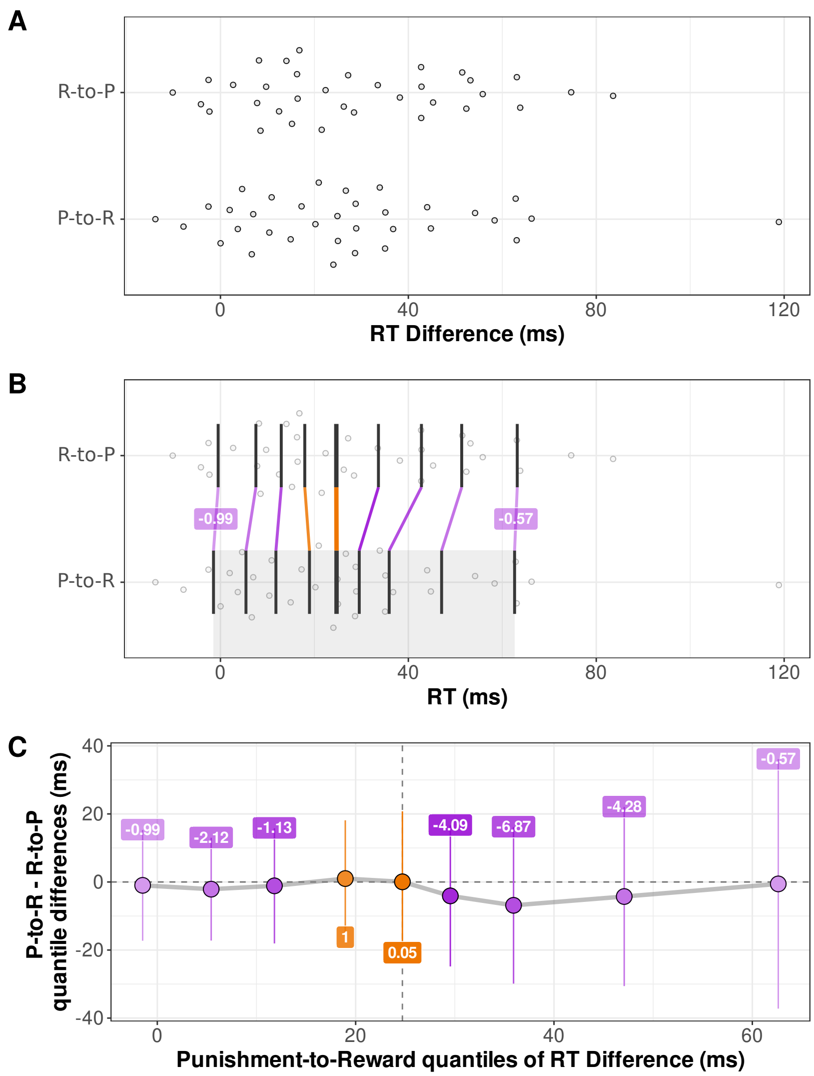

Figure 4.4: Feedback order distributions and the associated shift function during retention. **(A) Scatterplot.** Response time is shown on the x-axis with Punishment-to-Reward feedback (P-to-R) and Reward-to-Punishment feedback (R-to-P) on the y-axis. **(B) Scatterplot with deciles of the distribution.** Refer to Figure 4.1 for in depth explanation. **(C) Shift function.** Refer to Figure 4.1 for in depth explanation.

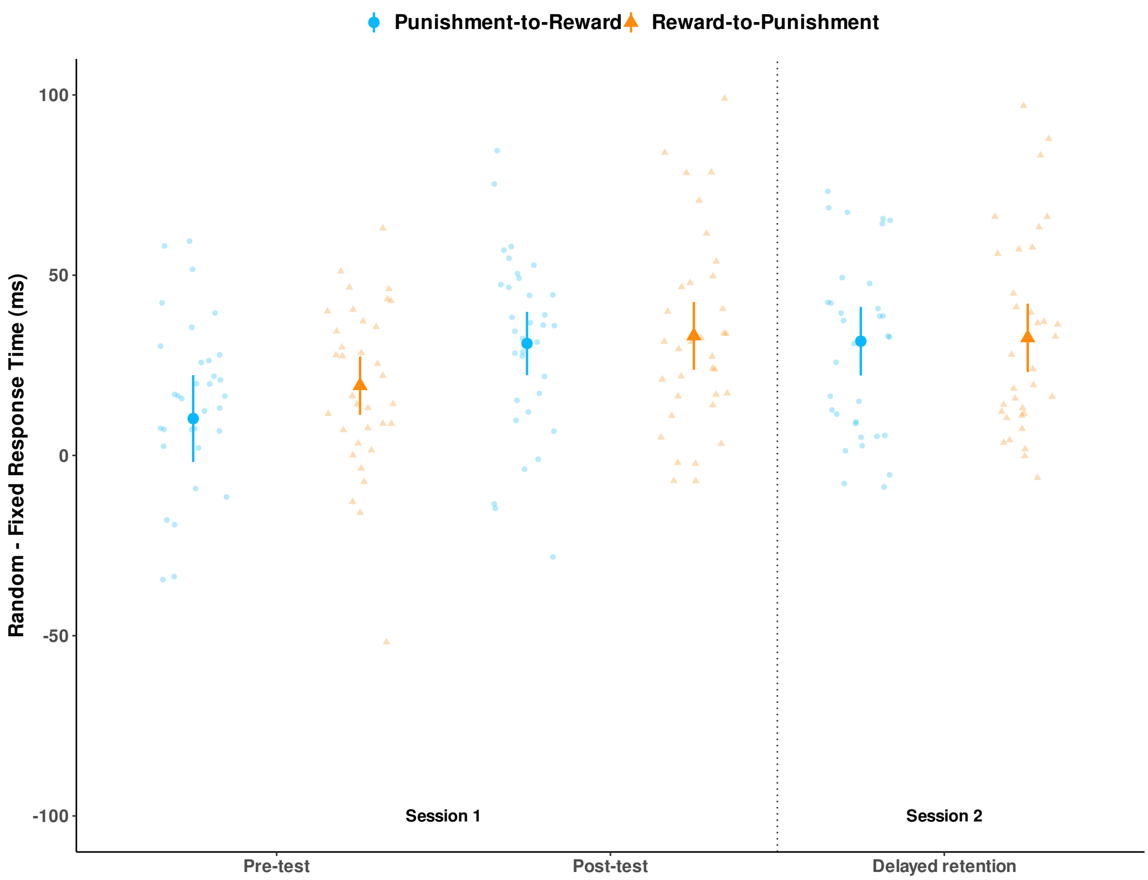

Figure 4.5: Mean response times in pre-test, post-test, and retention. The mean random - fixed response time (ms) of the Punishment-to-Reward (large blue circle) and Reward-to-Punishment (large orange triangle) feedback order during pre-test, post-test, and delayed retention. The small dots represent the mean random - fixed response times for individual participants in the Punishment-to-Reward (small blue circle) and Reward-to-Punishment (small orange triangle) feedback order. Mean random - fixed response times were derived from the difference in response time between the mean of the two random sequence blocks and the single fixed sequence block. Vertical dotted line separates Session 1 and 2. Error bars indicated 95% confidence intervals.

Chapter 5

Discussion

The main goal of the present experiment was to test whether the order of receiving a combined punishment and reward feedback schedule differentially impacted sequence learning and retention. To this end, we designed our pre-registered experiment based on Steel *et al.* (2016) as they used a variation of the serial reaction time task (Nissen and Bullemer, 1987), which was amenable to online data collection due to the COVID-19 global pandemic. This research question was motivated by distinct predictions from the coaching and motor learning literatures. Based on the motor learning literature, we predicted that a punishment-to-reward feedback order would be more effective learning and retention. The results of the first interim analysis in our sequential design did not support this prediction. The data was also not conclusive in terms of stopping data collection for futility (Lakens *et al.*, 2021b); thus, data collection will proceed to the second interim analysis. A discussion of our interim one analysis results follows.

5.1 No punishment-to-reward feedback advantage

Both the primary (i.e., shift functions) and secondary (i.e., general linear models) analyses did not support a learning and retention advantage of receiving punishment-to-reward feedback over reward-to-punishment feedback at this stage in our sequential design. Based on the motor learning literature (e.g., Galea *et al.*, 2015; Wachter *et al.*, 2009), we anticipated that the previously found faster learning rate with punishment feedback would have resulted in participants reaching a steady-state of repeating successful actions earlier in practice, and such actions would then be reinforced once the transition to reward feedback occurred. However, this does not seem to be the case. Instead, those in the Reward-to-Punishment group had numerically faster response times in multiple training blocks than those in the Punishment-to-Reward group. Although such trends were not significant based on the parameters of our statistical tests, they are in line with the predictions of the coaching literature. Specifically, it has been argued that rewards should be used early and often in training to form and strengthen its relationship with the desired movement (e.g., Warren, 1983; Burton and Raedeke, 2008). Following this, punishment feedback can be introduced in the later stages of training to reduce undesired movements and also prevent athlete's from developing negative feelings towards their coaches had punishment been used too early in the learning process (Williams and Krane, 2015). Thus, at this point if an order effect truly exists, we did not have the statistical power to detect an effect of this size. Further, it seems plausible that effect could be in the opposite direction to our predictions based on current trends in the data. Alternatively, it is possible that there is no effect to be detected and instead, just receiving both punishment and reward feedback, irrespective of order, during training is beneficial. Unfortunately, we cannot rule out such a possibility at this time as the results of our equivalence test were inconclusive. That is, the current estimate was not statistically equivalent to zero based on our smallest effect size of interest.

5.2 Punishment feedback did not benefit the learning process

Past motor learning research has shown that punishment feedback not only leads to a faster learning rate (e.g., Galea *et al.*, 2015), but also better overall learning during training (e.g., Steel *et al.*, 2016; Song *et al.*, 2020). The results from our hierarchical shift function (Rousselet and Wilcox, 2020) did not replicate this training-related advantage of punishment feedback. This failed replication (see also Abe *et al.*, 2011) was surprising given our choice of a more powerful statistical test (Rousselet and Wilcox, 2020) and that other researchers have actually argued that the training-related benefits of punishment feedback are more robust than those associated with reward feedback (e.g., Steel *et al.*, 2016; Song and Smiley-Oyen, 2017; Song *et al.*, 2020; Steel *et al.*, 2019).

The learning benefits of punishment feedback have been linked to loss aversion (Kahneman and Tversky, 1979, 1984). Loss aversion describes the behavioural phenomenon of avoiding choices that might result in losses, even when equal or larger gains are available (De Martino *et al.*, 2010; Kahneman and Tversky, 1979; Tversky and Kahneman, 1981). Past research has used monetary (e.g., 5 cents) gains and losses on feedback trials, whereas in the present experiment we used ± 0.10 points on feedback trials. Although this decision was largely a logistical one, it is possible that losing points does not have the same motivational salience as losing cents, and our participants therefore did not experience the same aversion to punished trials as those in previous experiments (e.g., Galea *et al.*, 2015; Steel *et al.*, 2016; Wachter *et al.*, 2009). To increase the motivational salience of our administration of feedback, we informed participants that a good point score at the end of the training period would earn them more entries into the gift card lottery, thereby increasing their chances of winning one of the six gift cards. Whether this incentive is as effective as using monetary incentives is unclear; however, we encourage researchers to use monetary gains and losses rather than points in future research as this will ensure greater consistency in methods when drawing conclusions relative to previous research.

5.3 Statistical power and design issues

While any of the above mentioned reasons may have contributed to the outcomes found with our first interim look, a more general explanation may be that the effects associated with reward and punishment feedback during motor learning are much smaller than previously estimated. Although we only collected 33% of our intended sample size, and thus our design is underpowered relative to our smallest effect size of interest ($d = 0.4$), our sample of 34 participants per group is much larger than the median ($n = 21$) of the research that motivated and informed our experiment. Our sample size was 2.8 times that of Steel *et al.* (2016), which had one of the smaller sample sizes per group ($n = 12$) of previous research. Interestingly, our 33% of our total planned sample size for our group sequential design exceeds the recommendation that replication experiments should aim for 2.5 times the sample size of the original experiment (Simonsohn, 2015). Small sample sizes can result in underpowered designs and therefore in comparison to larger sample sizes, significant results are more likely to be a Type 1 error (Lakens *et al.*, 2021a; Simmons *et al.*, 2011a). In other words, researchers may conclude that an effect is present when in reality there is not an effect.

Experimental designs with low statistical power (see Lohse *et al.*, 2016, for a discussion specific to motor learning) are a challenge when interpreting the available research for at least three other reasons. First, there is greater variability around the effect sizes that are estimated. This contributes to the magnitude of an effect being overestimated (Lohse *et al.*, 2016; Gelman and Carlin, 2014),[1] which might have been the case in one or more of the experiments in this area of research. Some support for this notion is that when we calculated our smallest effect size of interest based on what Steel *et al.* (2016) had 33% power to detect, this resulted in a d of 0.62. This would suggest that the effect of punishment and reward feedback in motor learning is larger than the effect that men weigh more than women ($d = 0.56$; Simmons *et al.*, 2013). Second, underpowered designs are susceptible to making a signed error (i.e., Type S error, where the significant finding of an experiment is estimated in the incorrect direction (Gelman and Carlin, 2014). Lastly, it can lead to low reproducibility of results (Button *et al.*, 2013; Lohse *et al.*, 2016; Open Science Collaboration, 2015). To address this issue, researchers are encourage to disclose the rationale behind their sample size, analysis plan, data exclusion criteria, and data manipulations (e.g., Munafò *et al.*, 2017; Lakens, 2021; Button *et al.*, 2013; Simmons *et al.*, 2011b). By doing so, it will allow other researchers to gain a better understanding of the reported effects and increase the transparency regarding researcher degrees of freedom.

[1] Gelman and Carlin (2014) termed this a Type M error.

Chapter 6

Conclusion

Overall, the current study failed to find conclusive evidence that the order of receiving a combined punishment and reward feedback schedule matters for learning and retention. However, contrary to previous studies, the true difference between the punishment and reward feedback condition during training may be smaller than previously thought.

6.1 Limitations

A limitation in the current study is that the retention interval, the time between the end of training and the retention test, is not uniform across all the participants. In fact, some participants may have had a retention interval greater or less than 24 hours depending on when they completed the next-day post test. However, previous studies have demonstrated that retention of a motor skill is likely sleep dependent (Kami *et al.*, 1995; Fischer *et al.*, 2002; Walker *et al.*, 2002). Specifically, performance during a retention test benefits from a period of sleep, independent of whether participants slept during the day or night (Fischer *et al.*, 2002). Therefore, it is unlikely that a non-uniform retention interval impacted our findings.

Another limitation is that we failed to separate response time into its individual parts, reaction time and movement time. Accordingly, a change in response time may have been associated with a change in reaction time, movement time, or a combination of both. Furthermore, a change in reaction time or movement time might represent a change in action selection (i.e., choosing the appropriating movement) or action execution (i.e., the ability to execute the movement), respectively (Diedrichsen and Kornysheva, 2015; Krakauer *et al.*, 2019). However, the current study was unable to address these questions because we did not distinguish between the individual measures.

6.2 Future directions

Future work should consider investigating whether the order of a transition feedback schedule has trivial effects on specific elements of response time. According to (Chen *et al.*, 2018), punishment and reward feedback could be affecting only reaction time, movement time, or a combination of both. As a result, designing a task that can isolate these specific components would allow researchers to gain a greater understanding of the impact of a combined punishment and reward feedback schedule on motor learning.

References

Abe, M., Schambra, H., Wassermann, E. M., Luckenbaugh, D., Schweighofer, N., and Cohen, L. G. (2011). Reward Improves Long-Term Retention of a Motor Memory through Induction of Offline Memory Gains. *Current Biology*, **21**(7), 557–562.

Adams, J. A. (1971). A closed-loop theory of motor learning. *Journal of Motor Behavior*, **3**(2), 111–149.

Anderson, D. I., Magill, R. A., Mayo, A. M., and Steel, K. A. (2019). Enhancing motor skill acquisition with augmented feedback. In A. M. Williams and N. J. Hodges, editors, *Skill acquisition in sport: Research, theory and practice*. Routledge.

Bennett, D. M. and Simmons, R. W. (1984). Effects of Precision of Knowledge of Results on Acquisition and Retention of a Simple Motor Skill. *Perceptual and Motor Skills*, **58**(3), 785–786.

Bilodeau, E. A. and Bilodeau, I. M. (1958). Variable frequency of knowledge of results and the learning of a simple skill. *Journal of Experimental Psychology*, **55**(4), 379–383.

Bromberg-Martin, E. S., Matsumoto, M., and Hikosaka, O. (2010). Dopamine in

Motivational Control: Rewarding, Aversive, and Alerting. *Neuron*, **68**(5), 815–834.

Burton, D. and Raedeke, T. D. (2008). *Sport psychology for coaches*. Human Kinetics, Champaign, IL.

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., and Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, **14**(5), 365–376.

Carter, M. J., Carlsen, A. N., and Ste-Marie, D. M. (2014). Self-controlled feedback is effective if it is based on the learner's performance: a replication and extension of chiviacowsky and wulf (2005). *Frontiers in Psychology*, **5**, 1325.

Carter, M. J., Rathwell, S., and Ste-Marie, D. (2016). Motor skill retention is modulated by strategy choice during self-controlled knowledge of results schedules. *Journal of Motor Learning and Development*, **4**(1), 100–115.

Cashaback, J. G. A., McGregor, H. R., Mohatarem, A., and Gribble, P. L. (2017). Dissociating error-based and reinforcement-based loss functions during sensorimotor learning. *PLOS Computational Biology*, **13**(7), e1005623.

Cashaback, J. G. A., Lao, C. K., Palidis, D. J., Coltman, S. K., McGregor, H. R., and Gribble, P. L. (2019). The gradient of the reinforcement landscape influences sensorimotor learning. *PLoS Computational Biology*, **15**(3), e1006839.

Chen, X., Holland, P., and Galea, J. M. (2018). The effects of reward and punishment on motor skill learning. *Current Opinion in Behavioral Sciences*, **20**, 83–88.

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, **47**(1), 1–12.

De Martino, B., Camerer, C. F., and Adolphs, R. (2010). Amygdala damage eliminates monetary loss aversion. *Proceedings of the National Academy of Sciences of the United States of America*, **107**(8), 3788–3792.

Diedrichsen, J. (2005). Neural Correlates of Reach Errors. *Journal of Neuroscience*, **25**(43), 9919–9931.

Diedrichsen, J. and Kornysheva, K. (2015). Motor skill learning between selection and execution. *Trends in Cognitive Sciences*, **19**(4), 227–233.

Dodge, H. F. and Romig, H. G. (1929). A Method of Sampling Inspection. *Bell System Technical Journal*, **8**(4), 613–631.

Doyon, J., Owen, A. M., Petrides, M., Sziklas, V., and Evans, A. C. (1996). Functional Anatomy of Visuomotor Skill Learning in Human Subjects Examined with Positron Emission Tomography. *European Journal of Neuroscience*, **8**(4), 637–648.

Doyon, J., Song, A. W., Karni, A., Lalonde, F., Adams, M. M., and Ungerleider, L. G. (2002). Experience-dependent changes in cerebellar contributions to motor sequence learning. *Proceedings of the National Academy of Sciences*, **99**(2), 1017–1022.

Fischer, S., Hallschmid, M., Elsner, A. L., and Born, J. (2002). Sleep forms memory for finger skills. *Proceedings of the National Academy of Sciences*, **99**(18), 11987–11991.

Galea, J. M., Mallia, E., Rothwell, J., and Diedrichsen, J. (2015). The dissociable effects of punishment and reward on motor learning. *Nature Neuroscience*, **18**(4), 597–602.

Gelman, A. and Carlin, J. (2014). Beyond Power Calculations: Assessing Type S (Sign) and Type M (Magnitude) Errors. *Perspectives on Psychological Science*, **9**(6), 641–651.

Hadjiosif, A. M., Krakauer, J. W., and Haith, A. M. (2021). Did We Get Sensorimotor Adaptation Wrong? Implicit Adaptation as Direct Policy Updating Rather than Forward-Model-Based Learning. *The Journal of Neuroscience*, **41**(12), 2747–2761.

Harrell, F. E. and Davis, C. E. (1982). A new distribution-free quantile estimator. *Biometrika*, **69**(3), 635–640.

Harris, C. S. (1963). Adaptation to Displaced Vision: Visual, Motor, or Proprioceptive Change? *Science*, **140**(3568), 812–813.

Helmholtz, H. v. (2013). *Treatise on Physiological Optics*. Courier Corporation.

Hochberg, Y. (1988). A sharper Bonferroni procedure for multiple tests of significance. *Biometrika*, **75**(4), 800–802.

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

Izawa, J. and Shadmehr, R. (2011). Learning from Sensory and Reward Prediction Errors during Motor Adaptation. *PLoS Computational Biology*, **7**(3), e1002012.

Izawa, J., Criscimagna-Hemminger, S. E., and Shadmehr, R. (2012). Cerebellar Contributions to Reach Adaptation and Learning Sensory Consequences of Action. *Journal of Neuroscience*, **32**(12), 4230–4239.

Jean-Richard-Dit-Bressel, P., Killcross, S., and McNally, G. P. (2018). Behavioral and neurobiological mechanisms of punishment: implications for psychiatric disorders. *Neuropsychopharmacology*, **43**(8), 1639–1650.

Kaelbling, L. P., Littman, M. L., and Moore, A. W. (1996). Reinforcement Learning: A Survey. *Journal of Artificial Intelligence Research*, **4**, 237–285.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, **47**(2), 263–291.

Kahneman, D. and Tversky, A. (1984). Choices, values, and frames. *American Psychologist*, **39**(4), 341–350.

Kami, A., Meyer, G., Jezzard, P., Adams, M. M., Turner, R., and Ungerleider, L. G. (1995). Functional MRI evidence for adult motor cortex plasticity during motor skill learning. *Nature*, **377**(6545), 155–158.

Kantak, S. S. and Winstein, C. J. (2012). Learning–performance distinction and memory processes for motor skills: A focused review and perspective. *Behavioural Brain Research*, **228**(1), 219–231.

Krakauer, J., Hadjiosif, A., Xu, J., Wong, A., and Haith, A. (2019). Motor Learning. *Comprehensive Physiology*, **9**, 613–663.

Krakauer, J. W. (2009). Motor Learning and Consolidation: The Case of Visuomotor

Rotation. In D. Sternad, editor, *Progress in Motor Control*, volume 629, pages 405–421. Springer US, Boston, MA.

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

Lai, Q., Shea, C. H., Wulf, G., and Wright, D. L. (2000). Optimizing Generalized Motor Program and Parameter Learning. *Research Quarterly for Exercise and Sport*, **71**(1), 10–24.

Lakens, D. (2014). Performing high-powered studies efficiently with sequential analyses: Sequential analyses. *European Journal of Social Psychology*, **44**(7), 701–710.

Lakens, D. (2021). Sample size justification. *Psyarxiv*.

Lakens, D., Pahlke, F., and Wassmer, G. (2021a). Group sequential designs: A tutorial. *Psyarxiv*.

Lakens, D., Pahlke, F., and Wassmer, G. (2021b). Group Sequential Designs: A Tutorial. preprint, PsyArXiv.

Leys, C., Ley, C., Klein, O., Bernard, P., and Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, **49**(4), 764–766.

Leys, C., Delacre, M., Mora, Y. L., Lakens, D., and Ley, C. (2019). How to classify, detect, and manage univariate and multivariate outliers, with emphasis on Pre-Registration. *International Review of Social Psychology*, **32**(1), 270.

Lindahl, L. G. (1945). Movement analysis as an industrial training method. *Journal of Applied Psychology*, **29**(6), 420–436.

Lohse, K., Buchanan, T., and Miller, M. (2016). Underpowered and overworked: Problems with data analysis in motor learning studies. *Journal of Motor Learning and Development*, **4**(1), 37–58.

Lohse, K., Miller, M., Bacelar, M., and Krigolson, O. (2019). Errors, rewards, and reinforcement in motor skill learning. In A. M. Williams and N. J. Hodges, editors, *Skill acquisition in sport: Research, theory and practice*. Routledge.

Luft, C. D. B. (2014). Learning from feedback: The neural mechanisms of feedback processing facilitating better performance. *Behavioural Brain Research*, **261**, 356–368.

Magill, R. A. and Anderson, D. (2017). *Motor learning and control: concepts and applications*. McGraw-Hill Education, New York, NY, eleventh edition edition.

Magill, R. A., Chamberlin, C. J., and Hall, K. G. (1991). Verbal knowledge of results as redundant information for learning an anticipation timing skill. *Human Movement Science*, **10**(4), 485–507.

Martin, T. A., Keating, J. G., Goodkin, H. P., Bastian, A. J., and Thach, W. T. (1996). Throwing while looking through prisms. I. Focal olivocerebellar lesions impair adaptation. *Brain: A Journal of Neurology*, **119 (Pt 4)**, 1183–1198.

McDougle, S. D., Ivry, R. B., and Taylor, J. A. (2016). Taking aim at the cognitive side of learning in sensorimotor adaptation tasks. *Trends in Cognitive Sciences*, **20**(7), 535–544.

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., and Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, **1**(1), 0021.

Newell, K., Quinn, J., Sparrow, W., and Walter, C. (1983). Kinematic information feedback for learning a rapid arm movement. *Human Movement Science*, **2**(4), 255–269.

Nissen, M. J. and Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, **19**(1), 1–32.

Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science*, **349**(6251), aac4716.

Penhune, V. B. and Steele, C. J. (2012). Parallel contributions of cerebellar, striatal and M1 mechanisms to motor sequence learning. *Behavioural Brain Research*, **226**(2), 579–591.

Peterson, E. J. and Seger, C. A. (2013). Many hats: intratrial and reward level-dependent BOLD activity in the striatum and premotor cortex. *Journal of Neurophysiology*, **110**(7), 1689–1702.

Pine, Z. M., Krakauer, J. W., Gordon, J., and Ghez, C. (1996). Learning of scaling factors and reference axes for reaching movements:. *NeuroReport*, **7**(14), 2357–2362.

Redding, G. M. and Wallace, B. (1993). Adaptive Coordination and Alignment of Eye and Hand. *Journal of Motor Behavior*, **25**(2), 75–88.

Robertson, E. M. (2007). The Serial Reaction Time Task: Implicit Motor Skill Learning? *Journal of Neuroscience*, **27**(38), 10073–10075.

Rousselet, G. A. and Wilcox, R. R. (2020). Reaction Times and other Skewed Distributions. *Meta-Psychology*, **4**.

Rousselet, G. A., Pernet, C. R., and Wilcox, R. R. (2017). Beyond differences in means: robust graphical methods to compare two groups in neuroscience. preprint, Neuroscience.

Salmoni, A. W., Schmidt, R. A., and Walter, C. B. (1984). Knowledge of results and motor learning: A review and critical reappraisal. *Psychological Bulletin*, **95**(3), 355–386.

Schendan, H. E., Searl, M. M., Melrose, R. J., and Stern, C. E. (2003). An fMRI Study of the Role of the Medial Temporal Lobe in Implicit and Explicit Sequence Learning. *Neuron*, **37**(6), 1013–1025.

Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological Review*, **82**(4), 225–260.

Schmidt, R. A. and Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science*, **3**(4), 207–218.

Schmidt, R. A., Lee, T. D., Winstein, C. J., Wulf, G., and Zelaznik, H. N. (2019). *Motor control and learning: a behavioral emphasis*. Human Kinetics, Champaign, IL, sixth edition edition.

Shadmehr, R. and Mussa-Ivaldi, F. (1994). Adaptive representation of dynamics during learning of a motor task. *The Journal of Neuroscience*, **14**(5), 3208–3224.

Shadmehr, R., Smith, M. A., and Krakauer, J. W. (2010). Error Correction, Sensory Prediction, and Adaptation in Motor Control. *Annual Review of Neuroscience*, **33**(1), 89–108.

Shmuelof, L., Krakauer, J. W., and Mazzoni, P. (2012). How is a motor skill learned? Change and invariance at the levels of task success and trajectory control. *Journal of Neurophysiology*, **108**(2), 578–594.

Sigrist, R., Rauter, G., Riener, R., and Wolf, P. (2013). Augmented visual, auditory, haptic, and multimodal feedback in motor learning: a review. *Psychonomic Bulletin & Review*, **20**(1), 21–53.

Simmons, J. P., Nelson, L. D., and Simonsohn, U. (2011a). False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, **22**(11), 1359–1366.

Simmons, J. P., Nelson, L. D., and Simonsohn, U. (2011b). False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. *Psychological Science*, **22**(11), 1359–1366.

Simmons, J. P., Nelson, L. D., and Simonsohn, U. (2012). A 21 word solution. *Available at SSRN: https://ssrn.com/abstract=2160588.*

Simmons, J. P., Nelson, L. D., and Simonsohn, U. (2013). Life after P-Hacking. *Available at SSRN: https://ssrn.com/abstract=2160588.*

Simonsohn, U. (2015). Small Telescopes: Detectability and the Evaluation of Replication Results. *Psychological Science*, **26**(5), 559–569.

Smith, M. A., Ghazizadeh, A., and Shadmehr, R. (2006). Interacting Adaptive Processes with Different Timescales Underlie Short-Term Motor Learning. *PLoS Biology*, **4**(6), e179.

Song, Y. and Smiley-Oyen, A. L. (2017). Probability differently modulating the effects of reward and punishment on visuomotor adaptation. *Experimental Brain Research*, **235**(12), 3605–3618.

Song, Y., Lu, S., and Smiley-Oyen, A. L. (2020). Differential motor learning via reward and punishment. *Quarterly Journal of Experimental Psychology*, **73**(2), 249–259.

Sparrow, W. A. and Summers, J. J. (1992). Performance on Trials without Knowledge of Results (KR) in Reduced Relative Frequency Presentations of KR. *Journal of Motor Behavior*, **24**(2), 197–209.

Steel, A., Silson, E. H., Stagg, C. J., and Baker, C. I. (2016). The impact of reward and punishment on skill learning depends on task demands. *Scientific Reports*, **6**(1), 36056.

Steel, A., Silson, E. H., Stagg, C. J., and Baker, C. I. (2019). Differential impact of reward and punishment on functional connectivity after skill learning. *NeuroImage*, **189**, 95–105.

Sutton, R. S. and Barto, A. G. (2018). *Reinforcement learning: an introduction.*

Adaptive computation and machine learning series. The MIT Press, Cambridge, Massachusetts, second edition edition.

Swinnen, S. P. (1996). Information feedback for motor skill learning: A review. In H. N. Zelaznik, editor, *Advances in motor learning and control*, pages 37–66. Human Kinetics.

Taylor, J. A. and Ivry, R. B. (2014). Cerebellar and prefrontal cortex contributions to adaptation, strategies, and reinforcement learning. *Progress in Brain Research*, **210**, 217–253.

Thorndike, E. L. (1927). The Law of Effect. *The American Journal of Psychology*, **39**(1/4), 212.

Tresilian, J. (2012). *Sensorimotor control and learning: An introduction to the behavioral neuroscience of action.* Palgrave Macmillan.

Tseng, Y.-w., Diedrichsen, J., Krakauer, J. W., Shadmehr, R., and Bastian, A. J. (2007). Sensory Prediction Errors Drive Cerebellum-Dependent Adaptation of Reaching. *Journal of Neurophysiology*, **98**(1), 54–62.

Tversky, A. and Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, **211**(4481), 453–458.

Ungerleider, L. (2002). Imaging Brain Plasticity during Motor Skill Learning. *Neurobiology of Learning and Memory*, **78**(3), 553–564.

Verwey, W. B. (1999). Evidence for a multistage model of practice in a sequential movement task. *Journal of Experimental Psychology: Human Perception and Performance*, **25**(6), 1693–1708.

Wachter, T., Lungu, O. V., Liu, T., Willingham, D. T., and Ashe, J. (2009). Differential Effect of Reward and Punishment on Procedural Learning. *Journal of Neuroscience*, **29**(2), 436–443.

Wald, A. (1945). Sequential tests of statistical hypotheses. *The Annals of Mathematical Statistics*, **16**(2), 117–186.

Walker, M. P., Brakefield, T., Morgan, A., Hobson, J., and Stickgold, R. (2002). Practice with Sleep Makes Perfect. *Neuron*, **35**(1), 205–211.

Warren, W. E. (1983). *Coaching and motivation: a practical guide to maximum athletic performance*. Prentice-Hall, Englewood Cliffs, N.J.

Weeks, D. L. and Kordus, R. N. (1998). Relative Frequency of Knowledge of Performance and Motor Skill Learning. *Research Quarterly for Exercise and Sport*, **69**(3), 224–230.

Wickens, J. R., Reynolds, J. N., and Hyland, B. I. (2003). Neural mechanisms of reward-related motor learning. *Current Opinion in Neurobiology*, **13**(6), 685–690.

Wilcox, R. R., Erceg-Hurn, D. M., Clark, F., and Carlson, M. (2014). Comparing two independent groups via the lower and upper quantiles. *Journal of Statistical Computation and Simulation*, **84**(7), 1543–1551.

Williams, J. M. and Krane, V. (2015). *Applied sport psychology: personal growth to peak performance*. McGraw-Hill Education, New York, seventh edition edition.

Winstein, C. J. and Schmidt, R. A. (1990). Reduced frequency of knowledge of results enhances motor skill learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **16**(4), 677–691.

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

Wolpert, D. M. and Flanagan, J. R. (2016). Computations underlying sensorimotor learning. *Current Opinion in Neurobiology*, **37**, 7–11.

Wolpert, D. M., Diedrichsen, J., and Flanagan, J. R. (2011). Principles of sensorimotor learning. *Nature Reviews Neuroscience*, **12**(12), 739–751.

Wong, A. L., Lindquist, M. A., Haith, A. M., and Krakauer, J. W. (2015). Explicit knowledge enhances motor vigor and performance: motivation versus practice in sequence tasks. *Journal of Neurophysiology*, **114**(1), 219–232.

Wulf, G. and Shea, C. H. (2002). Principles derived from the study of simple skills do not generalize to complex skill learning. *Psychonomic Bulletin & Review*, **9**(2), 185–211.

Young, D. E. and Schmidt, R. A. (1992). Augmented Kinematic Feedback for Motor Learning. *Journal of Motor Behavior*, **24**(3), 261–273.