PUNISHMENT AND REWARD FEEDBACK DURING MOTOR LEARNING

THE IMPACT OF PUNISHMENT AND REWARD FEEDBACK ON SEQUENCE LEARNING

BY

MIRETTE SALIB MOUNIR, 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:	The impact of punishment and reward feedback on sequence learning
AUTHOR:	Mirette Salib Mounir B.Sc. (Kinesiology), McMaster University, Hamilton, Canada
SUPERVISOR:	Dr. Michael J. Carter, Ph.D.
NUMBER OF PAGES:	xii, 73

Lay Abstract

The information that you get from your senses, along with the comments and criticisms given to you by others, are all forms of feedback that may or may not be available in your environment. Feedback can often be given to you in the form of a punishment or a reward in an effort to facilitate your performance of a motor skill, such as learning to juggle a soccer ball. In this thesis, I explored whether punishment and reward feedback have dissociable effects on the way people learn and retain a new motor skill. Some individuals received punishment feedback by seeing a red box flash on their screen and losing points, while others received reward feedback by seeing a green box flash on their screen and gaining points. Although the participants learned the new motor skill in the experiment, the results showed that reward and punishment feedback did not differentially impact motor learning. These results suggest that either reward or punishment feedback may be a useful feedback strategy for promoting motor learning.

Abstract

Next to practice itself, feedback provided to a learner from an external source such as a coach or therapist is considered the most important factor facilitating skill acquisition. Past research has suggested that punishment and reward feedback have dissociable effects on learning and retention, respectively. However, other studies have suggested a more reliable effect of punishment feedback while failing to replicate the benefit of reward on retention. This discrepancy across experiments may be the result of seemingly innocuous methodological differences. Here, I ran a pre-registered online experiment to test the replicability of the supposed dissociable effects of punishment and reward on learning during training and retention, respectively. Participants were randomly assigned to receive either punishment feedback ($n = 34$) or reward feedback ($n = 34$) during the training period as they learned a repeating 12-element sequence in a serial reaction time task. Feedback consisted of participants either seeing a red (Punishment group) or green (Reward group) box flash on their computer screen and, unbeknownst to them, either a corresponding loss (Punishment group) or gain (Reward group) of points from their starting total. Participants were informed that a good final point score (i.e., the higher the better) could earn them extra entries into a gift card lottery. Contrary to what much of the literature has found, our results revealed no statistically significant differences between groups in either the training

or retention phases of the experiment. In conclusion, the findings of this experiment failed to replicate the previously found dissociable effects of punishment and reward feedback on learning and retention, respectively. The data instead suggests that providing participants with punishment or reward feedback may affect learning and retention in a similar manner.

Acknowledgements

I would like to begin by thanking my supervisor Mike for his guidance and support throughout the past two years. You've been not only a wonderful academic mentor, but a great lab social host, B.C. travel companion, and friend. I owe my masters experience to you, and for that I am extremely grateful.

I would also like to thank my supervisory committee Dr. Jim Lyons and Dr. Neil Drummond, and my external examiner Dr. Diane Ste-Marie. It has been a crazy year to say the least and I am greatly appreciative of your input and encouragement throughout my master's journey. You have bettered my work in every possible way.

To my parents, Mounir and Dalia, thank you for always pushing me to be the best version of myself and always supporting my aspirations. I have made it this far because of your love and support. Dad, I would not have pursued this degree without your encouragement. You inspire me with your work ethic, and I am proud to hold a master's degree just like you. Mom, I would not have pursued my next degree without your mentorship and guidance. I hope to someday follow in your footsteps and become an amazing diligent dentist. Joseph, as far as siblings go, I lucked out. Thank you for being my voice of reason every time I felt stressed and overwhelmed. John, thank you for taking on this degree as if it were your own. Your constant support, patience, and understanding are what made every late-night bearable.

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

Mirette Mounir; Brad McKay; Michael J. Carter

Contributions using CRediT (Contributor Roles Taxonomy)

M. Mounir: 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 *M. Tobis:* Investigation *M.J. Carter:* Conceptualization, Data curation, Funding acquisition, Methodology, Supervision, Writing - Review & Editing

Contents

List of Figures

List of Tables

Chapter 1

Literature Review

1.1 General introduction

Movement is all around us. Whether that be the movement of objects, people, animals and various other things that surround us, or the change in location of our own body, movement is vital to our everyday life. Can you imagine trying to complete basic fundamental tasks without moving? It would be impossible to reach forward and pick up the water bottle on your desk. Getting to school or work would be no easy task. How about playing your favourite sport or instrument? We often take for granted the importance of movement, something that on the surface appears to be so simple, and anything but complex. In reality, movement is present in multiple forms and can be divided into two main categories: those that are genetically defined and those that must be learned (Schmidt *et al.*, 2018). The nature versus nurture debate that was sparked long ago might have just popped into your mind. Following suit, a genetically defined movement is one you likely share with your conspecifics; you and other humans are likely less flexible than octopi. In this case, the way you move and the actions you can perform are genetically determined through growth and/or development. On the other hand, you can also learn certain movements or "skills". If you are currently still driving the same way you were when you just got your driver's licence, that may be slightly problematic. Learned movements normally require practice and take time to master. Our movement repertoire can be placed on a spectrum, ranging from very simple to very complex (Schmidt *et al.*, 2018). Repeatedly flexing your leg is very different from playing a piece by Mozart on the piano; one requires considerably more dexterity than the other. It is therefore important to not only examine how movements are controlled, an area of research called motor control, but also how motor skills are learned, an area of research known as motor learning. Although it is common for researchers to approach problems in motor control and motor learning separately—as will be done in subsequent sections of this thesis—it is important to remember that both disciplines are in fact highly interconnected (Schmidt *et al.*, 2018).

Wolpert *et al.* (2011) highlighted three main components of motor learning: information extraction, decisions and strategies, and classes of control. Sensory information must be effectively and efficiently gathered and processed. Our movements influence the information that our senses pick-up and provide us with. This incoming sensory information is filtered due to limited available resources; only a portion of our cognitive capacity can be allocated to the action at hand (Broadbent, 1958; Treisman, 1964). After information extraction, a series of decisions have to be made, and strategies have to be selected; which movement should happen next and when? In order to optimize what is about to happen, there are three classes of control that assist the process: predictive control, reactive control, and biomechanical control. Depending on the motor action being conducted, different combinations of these types of control are usually utilized to implement change (Flanagan *et al.*, 2006; Franklin *et al.*, 2008).

Krakauer *et al.* (2019) surveyed the major existing approaches, at both the behavioral and neural levels, in which motor learning is characterized. They also reviewed two long-standing paradigms (adaptation and sequence learning) used in motor learning research, which will be discussed in greater detail in the following section. They defined motor learning as any experience-dependent improvement in performance, wherein improvement refers to the production of more effective movements. Krakauer *et al.* (2019) further separated motor learning into *skill acquisition* and *skill maintenance*. Skill acquisition refers to "the process by which an individual acquires the ability to rapidly identify an appropriate movement goal given a particular task context, select the correct action given a sensory stimulus and/or the current state of the body and the world, and execute that action with accuracy and precision" (p. 615). Skill maintenance refers to "the ability to maintain performance levels of existing skills under changing conditions" (p. 615). Consider the task of juggling a soccer ball. A 15 year old girl acquiring the ability to properly control the ball and keep it off the ground for an extended period of time would fall into the motor learning category of skill acquisition. Now, fast forward 30 years. The now 45 year old lady adapting her behaviour by knocking the soccer ball higher to accommodate various age-related changes to her sensorimotor system would fall into the skill maintenance category of motor learning. To better understand the processes involved in motor learning, scientists have developed, refined, and relied on a variety of tasks and paradigms over the years.

1.2 Motor learning paradigms

In this section, I will focus on two paradigms that have been used extensively in the motor learning literature: adaptation learning and sequence learning (see (Figure 1.1). Although these paradigms differ in terms of what the participant is required to perform, both typically consist of three experimental phases: pre-training (or pre-test), training (or practice), and post-training (or post-test/retention).

1.2.1 Adaptation learning

When it comes to motor adaptation, the goal for the participant is to reduce their error to zero (or as close to zero as possible). As such, motor adaptation paradigms have long utilized a specific type of learning process known as error-based learning (Shadmehr and Mussa-Ivaldi, 1994; Martin *et al.*, 1996; Krakauer *et al.*, 2000). This type of learning, along with two others, will be discussed in a later section. To study motor adaptation in the laboratory, researchers have historically relied on upper limb reaching tasks using prism goggles, force-fields, or visuomotor rotations. As mentioned, all three tasks involve at least three key phases. Pre-training exposes the participant to the task under normal conditions and provides an assessment of their baseline performance. The training phase introduces a novel perturbation into the environment, which results in a high amount of error. The participant must learn to overcome this systematic perturbation by learning to adapt their performance. Lastly, during post-training the perturbation is removed and the participant performs the task under normal conditions, similar to pre-training. However, if the participant learned to adapt their movement during training, there is an initial high amount of error during post-training, but this time in the opposite direction. This is known

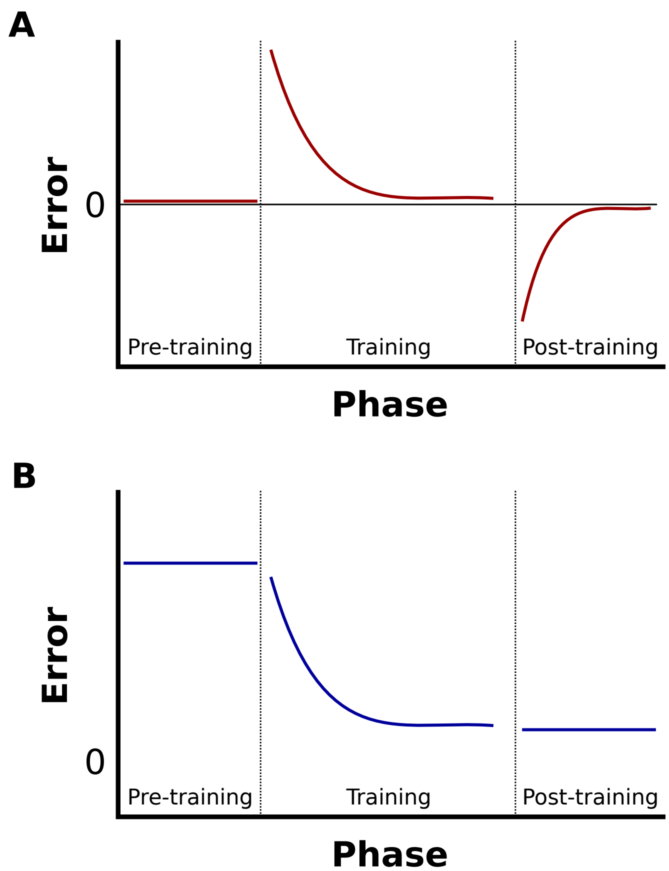

Figure 1.1: Representative performance curves found in adaptation learning tasks and sequence learning tasks. **(A) Adaptation learning.** Adaptation tasks are used to examine how individuals maintain their performance when a systematic perturbation is introduced into their learning environment. At the start of the training phase, the perturbation (e.g., force field, cursor rotation, etc) is introduced and there is an immediate spike in error. The error gradually declines to pre-training levels as the individual learns to change their behaviour to compensate for the perturbation. At the start of the post-training phase, there is also a spike in error of similar magnitude, but is now in the opposite direction to that seen in training. The error decays 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 over practice trials. 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.

as an aftereffect (Jeannerod, 1988; Fernández-Ruiz and Díaz, 1999). Performance quickly returns to their baseline levels during this post-training phase. Based on this overview, we can classify adaptation tasks as the skill maintenance type of motor learning (Krakauer *et al.*, 2019).

Prism adaptation

Prism adaptation tasks have been used by researchers long before computers were invented by implementing the use of prism glasses or prism goggles; a technique introduced by Hermann von Helmholtz in 1867. In this task, the goggles cause participants to experience a shift in their visual scene, which results in the performance of a reaching movement that misses the target (Held and Schlank, 1959; Held and Freedman, 1963; Redding *et al.*, 2005). Overtime, the individual learns to adjust their movement so it will result in task success (e.g., hitting a target). However, once the goggles are removed (i.e., in post-training), aftereffects are present in their reaching movements.

Visuomotor adaptation

With the advancement of technology, such as computers, a second motor adaptation paradigm emerged which provided a more modern approach to changing a participant's visual environment (Cunningham, 1989; Bock, 1992; Krakauer, 2009). Visuomotor adaptation tasks require participants to make reaching movements towards a target without being able to directly see their hand. Instead, the position and movement of their hand is mimicked by the location of a cursor on the screen in front of them, which is providing them with information regarding their performance. While they are completing the movement, a rotation is applied to the cursor (e.g., 30 degrees counterclockwise), and participants have to learn to counter the applied rotation by reaching in the opposite direction (e.g., 30 degrees clockwise) to successfully hit the target (Shadmehr *et al.*, 2010). Similar to prism adaptation, aftereffects in the participant's reaching movements are present once the rotation is removed.

Force field adaptation

Force-field adaptation paradigms are the final adaptation task to be discussed in this section. Participants are required to make reaching movements in the presence of an externally imposed force (e.g., dynamic perturbations in the leftward or rightward direction) produced by a robotic manipulandum that they are holding. From the starting position to the target, hand trajectories are originally grossly distorted displaying an angular convex path in the direction of the perturbation. Eventually, participants learn to compensate for these novel dynamics by applying opposing forces, which results in relatively straight hand trajectories by the end of training. However, when the force-field is removed, aftereffects in the participant's reaching behaviour are found (Shadmehr and Mussa-Ivaldi, 1994; Smith and Shadmehr, 2005).

Interim summary

Being able to quickly and flexibly adjust our behaviours to ensure task success in our ever-changing environment is an adaptive feature of our sensorimotor system and a crucial part of our everyday lives. For example, think back to when you were a child and would attempt to poke a pebble in a pond with a stick. At first, you would miss the pebble due to light refraction shifting where your stick appears to be. But quickly you are able to discover that if you make a slight adjustment in your reaching direction with the stick, you are able to successfully poke the pebble. This example illustrates how we can quickly modify our behaviours to meet task demands rather than having to learn a completely new skill.

1.2.2 Sequence learning

The second group of paradigms involve tasks that are composed of sequential actions. Oftentimes such actions must be completed in a specific order to achieve the task goal. For instance, when making peanut butter toast, the lid must be removed from the container before being able to scoop and spread the desired amount of peanut butter on the toast. Similar to adaptation, sequential actions are ubiquitous in day-to-day life—sentences, musical pieces, driving, and typing all involve sequences. "*Fun studying learning motor is!*" probably does not make much sense to you; however, if the words are arranged in the correct order things become comprehensible: "*studying motor learning is fun!*". In short, sequence learning paradigms examine how quickly and accurately individuals learn the correct order of a sequence of discrete actions (Krakauer *et al.*, 2019).

Simple sequences

As the name implies, the first group of sequence learning tasks are the most straightforward and involve simple sequences. The task requires participants to repeatedly press four to six pre-specified keys (using different fingers), with the goal of completing the sequence as fast and accurately as possible. Korman *et al.* (2003) had participants complete multiple blocks of one of two randomly assigned five-sequence key presses.

Their aim was to assess when learning-related changes in performance occur. They found that in one session, when the novel motor task was first assigned, there were large performance gains. They also noted that following a single training session, immediate and long-term gains were evident in both the trained and untrained hands. Their last finding indicated that after multisession training, the gains also appeared between-sessions; however, only in the trained hand this time. Similarly, Wiestler and Diedrichsen (2013) also used different sequences of five key presses and found that after multiple days of training, faster and more accurate sequential movements were produced.

Discrete sequence production tasks

A less commonly used sequence learning task is the discrete sequence production task, which is used to study chunking (Abrahamse *et al.*, 2013). Chunking refers to combining a series of actions into a cohesive whole. During this task, participants practice at least two short sequences (each roughly six key presses long). They are then presented with multiple sequences and are required to choose the correct one based on a cue (the first element of the correct sequence). Researchers aim to examine the ability of the subject to quickly select the correct sequential response (Krakauer *et al.*, 2019). In this task, the participant can produce multiple responses after the initial stimulus/cue appears.

Serial reaction time task

The most popular sequence learning task is the serial reaction time task (Nissen and Bullemer, 1987). This task has not only been used for multiple decades but has been used in psychology, neuroscience, and kinesiology research. During a serial reaction time task, participants are typically cued by stimuli (e.g., visual or auditory) and in response must perform the correct movement as quickly and accurately as possible. The movements are typically button presses, with stimuli appearing one after the other in a fixed sequence; for example, 1-3-4-2-3-1-2-4-3-2-3-4 (a sequence of 12 elements) on loop. In the context of the serial reaction time task, it is important to highlight implicit and explicit learning. Implicit learning refers to learning that happens outside of conscious awareness, whereas an individual is aware that learning is happening during explicit learning (Robertson, 2007). Participants are not made aware of the repeating sequence during the serial reaction time task and as such, it is thought that learning with this task is predominantly implicit in nature. When participants complete fixed sequence training/practice, researchers are then able to assess learning through analyzing two main outcome measures: response time and accuracy.

It is important to take note of the terminology being used. Most papers using this task often call this first measure reaction time, or use reaction time and response time interchangeably. However, reaction time and response time are distinct (yet related) measures of motor performance (Krakauer *et al.*, 2019; Schmidt *et al.*, 2018). Reaction time refers to the interval from stimulus presentation to the initiation of the response. Response time captures the time between stimulus onset and the end of the movement. In other words, response time is the sum of reaction time and movement time, which is a measure of the time from the beginning of the response to its completion. It has consistently been shown that very different processes are utilized and required to react quickly in comparison to moving quickly (Krakauer *et al.*, 2019; Schmidt *et al.*, 2018). Specifically, the former has been linked to action selection (and therefore knowledge of sequence order) while the latter is associated with action execution (Diedrichsen and Kornysheva, 2015; Chen *et al.*, 2018). Although performance of the serial reaction time task is likely dominated by reaction time (movement time is relatively constrained by the amount of key travel in a keyboard), it is difficult to distinguish between these two measures, even though attempts have been made (Moisello *et al.*, 2009; Chen *et al.*, 2018). As such, I will refer to this primary outcome measure of the serial reaction time task as response time.

Interim summary

The purpose of all sequence learning tasks is generally to examine how quickly and accurately individuals can learn the correct order of a sequence of discrete actions (Krakauer *et al.*, 2019). Speed and accuracy are both vital to the tasks we complete regularly. Not only is learning the correct order of the mitosis and meiosis cycles important, you must also recall the information quickly within the allotted test time to pass your biology test.

Now that the two main motor learning paradigms have been discussed, it is important to discuss the role of feedback in motor learning. Feedback is information that can be provided to us before, during, and/or after execution of a motor response.

1.3 Feedback and motor learning

Every minute, individuals receive a constant supply of sensory information; whether that be auditory (the voices surrounding you at the park), olfactory (the smell coming from the cookies your mom is baking in the kitchen), or visual (the constant supply of images that are rapidly changing before your eyes on television). It is not surprising that sensory information can be instrumental for moving efficiently and effectively. The general term for this information is feedback, as the information is "fed back" to the sensorimotor system. Feedback can be further categorized into intrinsic (or inherent) feedback and augmented (or extrinsic) feedback (Schmidt *et al.*, 2018).

1.3.1 Intrinsic feedback

Intrinsic feedback is almost always available whenever a movement is made and informs the individual about their performance. In other words, in response to making an action, this form of feedback is provided as a natural byproduct (Schmidt *et al.*, 2018). If you have completed the action multiple times before, you are likely to have learned what is about to happen before it even happens. For example, as you run towards the net with a soccer ball, the second your foot comes into contact with the ball (i.e., the force from your body has kicked the ball in the desired direction), you know whether you are on- or off-target based on feedback. This process of comparing a learned reference of correctness to current inherent feedback allows for error detection and correction processes (Schmidt *et al.*, 2018). However, inherent feedback is often even more obvious than what was just described and requires no prior learning, as the outcome can be directly observed. For example, the ball missed the net, and you therefore know an error was made.

1.3.2 Augmented feedback

As the name implies, augmented feedback supplements (or augments) the processing of inherent feedback. Schmidt *et al.* (2018) described augmented feedback as information from the measured performance outcome that is fed back to the learner by an external source. This could be verbal feedback from your coach after each shot you make, or visually through information displayed on a computer screen after an entire game. Therefore, as we saw with the two previous examples, augmented feedback can be presented to a person in a variety of ways differing in terms of timing, type, or amount. Augmented feedback can be further split into two main categories: knowledge of results and knowledge of performance. A simple way to distinguish between these is that knowledge of results contains information about the performance outcome relative to the task goal, whereas knowledge of performance provides information about the movement characteristics that led to a performance outcome (Wälchli *et al.*, 2016). For example, imagine a soccer player kicked the ball towards the net and missed. A coach could provide the player with knowledge of results by telling them "You were off target by 1 meter" or could provide knowledge of performance by telling them "Your knee was too bent".

Salmoni *et al.* (1984) examined the role of knowledge of results feedback for motor learning. Specifically, they outlined the transient and relatively permanent effects of nine feedback manipulations (e.g., absolute and relative frequency, timing of presentation). The authors highlighted a motivational and an associational role of knowledge of results, but suggested a more prominent informational role. Specifically, knowledge of results helps to guide the individual towards achieving the task goal. In this way, knowledge of results serves to enhance performance during practice when the feedback is available. However, if knowledge of results is presented too frequently, this can negatively impact long(er) term retention as individuals become dependent on the presentation of the augmented feedback at the expense of learning to interpret intrinsic feedback sources. These distinct effects of knowledge of results are captured within the guidance hypothesis (Salmoni *et al.*, 1984), which has received considerable support in the motor learning literature (e.g., Schmidt, 1991; Swinnen, 1996; Winstein and Schmidt, 1990).

1.3.3 Feedback characteristics

Motor learning has been argued to be supported by three main learning processes: use-dependent learning, error-based learning, and reinforcement learning (Wolpert *et al.*, 2011; Spampinato and Celnik, 2021). Each of these learning processes, which will be highlighted in the following section, are strongly influenced by feedback. Building upon the knowledge of results and knowledge of performance distinction, we can further classify the feedback provided to a learner in terms of performance information (i.e., categorical, graded, or finely graded), motivation value (i.e., reward, punishment, or neutral), or a mixture of both (Luft, 2014).

First, using a soccer related example, we will examine what it means for an individual to receive feedback in terms of their performance. If each time you kick the ball towards the net your coach tells you either "hit/correct" or "miss/incorrect", you are receiving categorical feedback. There are only two parts or options, therefore categorical feedback is often called binary feedback. If, however, your coach decides to tell you that you were "too short", "on target", or "too long", you are receiving graded feedback. The information you are receiving can be thought of as on a scale and generally provides the learner with only directional information. Alternatively, if your coach informs you that you were off target (in this case the target is the soccer net) by 2 m to the left, you are receiving finely graded feedback. The information given to you has a numerical value (error units) and provides information about both direction and magnitude.

Motivational feedback is most often provided to the learner as a reward signal or a punishment signal. Continuing with our soccer example, your coach could provide you reward feedback by showing you a green flag, playing a pleasant tone, and/or awarding you some points. Punishment feedback could be provided using a red flag, playing an unpleasant tone, and/or taking points away. This motivational feedback is often referred to as reinforcement feedback as it provides the learner with no direction or magnitude information; it only signals whether a response was successful or unsuccessful (Luft, 2014). The use of red and green when providing motivational feedback may be selected as they carry psychologically relevant meanings, with red associated with negative (e.g., alarms, danger, stop) and green with positive (e.g., achievement, growth, currency, go) (Moller *et al.*, 2009). Lastly, individuals can receive feedback through a mixed approach by combining performance information and motivation feedback.

1.4 Types of learning processes

1.4.1 Use-dependent learning

During use-dependent learning, no outcome information is given to the individual and instead, it is the mere repetition of movement that causes changes in the motor system (Bütefisch *et al.*, 2000; Diedrichsen *et al.*, 2010). Classen *et al.* (1998) administered transcranial magnetic stimulation to the motor cortex of participants, which evoked direction-specific movements of their thumb. Following this, participants were required to repeatedly move their thumb for 30 minutes in the opposite direction to that evoked by the transcranial magnetic stimulation. During post-training, transcranial magnetic stimulation was once again applied to the same area of the participant's motor cortex. However, the stimulation evoked thumb movements were now in the opposite direction from the pre-training phase, but in line with those made during training. This directional change lasted for approximately 15 to 20 minutes before returning to the pre-training direction.

1.4.2 Error-based learning

During error-based learning, our system utilizes a signed error signal and is guided by an internal estimate of the gradient in that direction (Wolpert *et al.*, 2011). Think back to when we discussed force-field adaptation. Shadmehr and Mussa-Ivaldi (1994) had participants sit facing a screen and grasp the handle of a robotic manipulandum. They were required to make reaching movements in the presence of externally imposed forces. They had multiple groups, some of which were allowed to see their hand position on the screen as they moved from the starting position to the target. This group was able to rely on error-based learning processes; they used the visual feedback of their hand trajectory to determine how to alter their next movement, in terms of direction and magnitude, to successfully hit the target. The provision of graded or finely graded feedback can promote the use of error-based learning processes as the error gradient is estimated by comparing the expected and actual movement outcome

(Luft, 2014; Wolpert *et al.*, 2011). Therefore, error-based learning is prominent in many motor adaptation paradigms. A drawback of this learning process is that once the average error is reduced to zero, no further improvements can be made.

1.4.3 Reinforcement learning

In reinforcement learning, the required directional change is not available to the learner; therefore, this form of learning is unsigned as only success or failure is indicated rather than the vector of errors (Sutton and Barto, 1998). Again, we are going to think back to when we discussed various paradigms, but this time we will use a simple sequence task example. Wiestler and Diedrichsen (2013) had participants rest their hand on a keyboard and complete sequences of five key presses. Each time participants pressed the right key, an asterisk on the screen turned green indicating that a correct response was made. However, if they pressed the wrong key, an asterisk on the screen turned red, indicating that an incorrect response was made. During reinforcement learning the individual learns which action(s) to perform by attempting to maximize reward or success through trial and error. Based on this, it is clear that receiving categorical or binary feedback would promote the use of reinforcement learning processes (Luft, 2014).

A limitation of reinforcement learning is that because the feedback you are provided with entails less specific information, learning generally occurs at a slower rate (Brown and Robertson, 2007). As such, error-based and reinforcement learning processes appear to influence the sensorimotor system on distinct timescales. This is important as these different learning processes can operate concurrently during motor learning (Wolpert *et al.*, 2011). For instance, Izawa and Shadmehr (2011) suggested that a decrease in error feedback quality is correlated with an increased reliance on reinforcement feedback. However, Cashaback *et al.* (2017) did not find this and instead found that when both error-based and reinforcement feedback are available, a greater reliance is placed on error feedback. As such, when studying reinforcement learning through punishment and reward feedback, it is critical to not introduce manipulations that may promote error-based learning processes.

1.5 Punishment and reward feedback during motor learning

The provision of punishment and reward feedback during motor learning has seen a renewed interest over the past decade. To better understand the impact of this feedback manipulation on motor learning and retention, researchers have predominantly used adaptation tasks and variations of the serial reaction time task. At this point, it is important to operationalize some key terms used in experiments employing these tasks. Within the kinesiology motor learning domain, learning and retention are often used interchangeably and refer to the relatively permanent changes inferred from performance after a period of no practice (Schmidt and Bjork, 1992; Kantak and Winstein, 2012). In the neuroscience motor learning domain, learning and retention are not synonymous. Specifically, learning refers to the changes in performance that occur during the training or practice period, and retention refers to the relatively permanent changes in performance (Diedrichsen and Kornysheva, 2015; Shmuelof *et al.*, 2012a,b). For the purpose of my thesis, learning and retention are operationally defined in line with the neuroscience motor learning domain as it is predominantly research from this area that has motivated my thesis work.

Wachter *et al.* (2009) had participants learn a serial reaction time task in either a reward feedback group, a punishment feedback group, or a control group. The authors differentiated between reward and punishment through the use of positive and negative monetary incentives. Subjects were required to use their right hand to press one of four keys on a key-press device depending on illumination of a visual stimulus. The reward group started at $0 and worked their way up, earning 4 cents for each press in which their response time was less than their cumulative response time (i.e., their median response time of the initial four random sequence blocks). The punishment group started at $38 and lost 4 cents for each press in which their response time was greater than their cumulative response time. Subjects received ongoing feedback in two forms: an increasing or decreasing counter displaying the current monetary amount, and as green (reward) or red (punishment) stimuli following each button press, depending on their group assignment and whether they were greater than or less than their cumulative response time. The results showed an association between punishment and faster response times in training, and an association between reward and faster response times in retention. However, it is important to note that the four retention blocks were added to the experiment late, and therefore 27 participants had already completed the study with only 11 blocks. In addition, retention was assessed once and on the same day as training. However, Abe *et al.* (2011) found a retention advantage of reward over punishment feedback 6 hours, 24 hours, and 30 days after training on an isometric pinch force tracking task.

Galea *et al.* (2015) had participants perform a visuomotor adaptation task that involved reaching towards visual targets on a screen without and with a cursor rotation of 30° counterclockwise applied. Online visual feedback was given in the form of a green cursor, while endpoint feedback was shown as a yellow circle. Participants completed this task in either a reward, punishment, or null feedback group. Money was either earned or lost depending on one's group and trial-by-trial endpoint angular error. After each trial, feedback appeared on the screen in the form of points (positive numbers for reward and negative numbers for punishment); the accumulation of points (whether positive or negative) signified earned or lost money. Participants first completed a baseline phase under normal reaching conditions. Following this, participants completed a training phase with the feedback manipulation present and the cursor rotation applied. Participants also completed a no-vision phase following training to assess retention. The results showed dissociable effects of punishment and reward feedback on learning and retention—learning was accelerated through punishment feedback and reward feedback increased retention.

It is important to think back to our discussion regarding the three learning processes. During reinforcement learning, the signal provided after performance is unsigned, meaning that the desired directional change is not provided. In this experiment however, through the implementation of what was termed online and endpoint feedback, the participant received signed information; participants could see the direction their cursor was travelling in relation to the target (i.e., to the left or to the right). In addition, the type of performance feedback that is given in reinforcement learning is typically binary. However, in this study the participants increased (or decreased) in different point values depending on their endpoint-to-target position, which is a form of graded feedback. Both of these manipulations would therefore engage error-based learning processes. Due to this mixing of performance and motivational feedback, it is difficult to isolate the effects of punishment and reward feedback alone on learning and retention in this experiment.

Steel *et al.* (2016) investigated whether the impact of punishment and reward feedback were task-dependent. In Experiment 1, participants learned a serial reaction time task in either a reward feedback group, a punishment feedback group, or a control group. In the reward group, participants started off at $0. Each time they were accurate and faster than their performance in the previous block, positive feedback was given in the form of a green flash, as well as an increase of 5 cents. In the punishment group, participants started off at $55. Each time they were inaccurate or slower than their performance on the previous block, negative feedback was given in the form of a red flash, and there was a decrease of 5 cents. Participants were not made aware of the meaning behind the color of the flash nor the associated 5 cent increase or decrease based on their performance. During the experiment, participants were presented with either a random sequence or a fixed sequence (a repeating 12-element sequence). On Day 1, participants completed a three block familiarization phase (all random), a three block pre-test (random-fixed-random), a six block training phase (all fixed), and finally a three block post-test (random-fixed-random). Delayed retention probes were completed 1 hour, 24 hours, and 30 days after the post-test. Feedback was only provided during the pre-test, training, and post-test phases on Day 1. While Steel et al. found the expected benefit of punishment feedback leading to better learning of a serial reaction time task during training, they did not replicate previously reported benefits of reward feedback on retention.

In Experiment 2 of Steel *et al.* (2016), a different set of participants learned an isometric pinch force tracking task in one of three feedback groups: reward, punishment, and control. Unlike with the serial reaction time task (Experiment 1), punishment feedback was found to negatively impact learning during the training phase. The reward group also did not show better retention. Taken together, these experiments suggest that punishment feedback may have more reliable effects but that these effects may be task-dependent. Additionally, these results question the effectiveness of reward feedback for retention as the data are inconsistent with previous studies (Abe *et al.*, 2011; Galea *et al.*, 2015; Wachter *et al.*, 2009).

Two additional studies examined the dissociable effects of punishment and reward feedback on visuomotor rotation. Song and Smiley-Oyen (2017) examined whether manipulating the amount of punishment and reward participants received resulted in distinct effects. They split participants into four groups: one group received punishments during 50% of the adaptation trials, another during 100%, while the last two groups were rewarded during 50% or 100% of adaptation trials, respectively. The trials in which participants received 50% punishments or 50% rewards were randomly selected. The results showed that punishments during all trials resulted in the fastest learning, whereas superior retention was found in the 100% reward and 50% punishment groups. As such, these findings suggest that the impact of punishment and reward feedback may interact with the relative frequency the feedback is provided.

Song *et al.* (2020) noted that a potential issue in the Galea *et al.* (2015) experiment was that participants had visual feedback of the cursor and thus, were able to make corrections based on this information. To address this, Song *et al.* (2020) removed visual feedback of the cursor during their visuomotor adaptation task and argued this resulted in participants having to rely only on their monetary feedback (increase or decrease in the reward and punishment groups, respectively) manipulation to make trial-by-trial performance adjustments. Consistent with Galea *et al.* (2015), punishment feedback resulted in faster learning. Contrary to Galea et al., reward feedback did not lead to enhanced retention. Thus, similar to Experiment 1 in Steel *et al.* (2016), Song *et al.* (2020) found a more reliable effect of punishment feedback on learning. However, the reward and punishment feedback provided in this experiment was not categorical (e.g., hit or miss) and was instead, graded as the amount of money (0 to 5 cents) gained or lost was based on performance error.

Interim summary

To date, the literature has revealed a fairly consistent effect of punishment feedback on facilitating the rate and amount of learning during training. In terms of reward feedback, although there is some inconsistency in its effects on retention, there is a reasonable support for the idea that reward feedback can enhance retention. Given these pattern of results, the primary aim of my thesis was to further examine the replicability of dissociable effects of punishment and reward feedback on learning and retention, respectively, using a serial reaction time task. An overview of the key experiments that motivated my thesis can be found in (see Table 1.1).

Table 1.1: Summary of main findings from relevant research examining dissociable effects of punishment and reward feedback.

Experiment	Learning effect	Retention effect
Wachter *et al.* (2009)	Yes	Yes
Abe *et al.* (2011)	No	Yes
Galea *et al.* (2015)	Yes	Yes
Steel *et al.* (2016)	Yes	No
Song *et al.* (2020)	Yes	No

1.6 Neural correlates of punishment and reward feedback

Cognition is a vital component of numerous fields, including motor learning (see Krakauer *et al.*, 2019; Lee *et al.*, 1994, for discussions). Hardwick *et al.* (2013) conducted a meta-analysis and reviewed the brain areas that contributed to human motor learning. They found consistent associations across tasks with the following regions: dorsal premotor cortex, supplementary motor cortex, primary motor cortex, primary somatosensory cortex, superior parietal lobule, thalamus, putamen (part of the basal ganglia), and cerebellum. The left dorsal premotor cortex was uniquely highlighted as playing a critical role in the motor learning network. In addition, Krakauer *et al.* (2019) also noted an association with five more areas: prefrontal cortex, presupplementary motor area, ventral premotor cortex, posterior parietal cortex, and the hippocampus. Furthermore, certain brain regions have shown differential activation patterns during the goal selection, action selection, and/or action execution stages that support motor learning processes (Krakauer *et al.*, 2019).

Through the description of various tasks in the Motor Learning Paradigms section above, it was shown that researchers generally rely on "simple" motor tasks in their experiments, with the assumption that such tasks can be used to better understand more complex skills. It is important to highlight and distinguish simple skills, such as those used in laboratory settings, from those that are more demanding, as the results from the former have been argued to not generalize to complex skill learning (Wulf and Shea, 2002). Hardwick *et al.* (2013) highlighted the brain areas that showed greater activation as a function of the motor learning paradigms used by researchers. During motor adaptation tasks where participants must learn new kinematics (e.g., visuomotor rotation) or dynamics (e.g., force-field), strong activation profiles were found in the basal ganglia and cerebellum. In contrast, during sequence learning tasks (e.g., serial reaction time task), preferential activation in cortical structures and the thalamus was found (Hardwick *et al.*, 2013).

1.6.1 Cognitive tasks

Elliott *et al.* (2000) narrowed in on the neural responses associated with rewards and penalties. Participants completed a gambling task which consisted of 12 trials within each of the 24 blocks; two cards were placed in front of them, and they were required to pick either the black card or the red card. Each time they guessed correctly they were rewarded with 1 pound, and when an incorrect guess was made, they lost 1 pound. Throughout task performance, participants were scanned using functional magnetic resonance imaging. Dissociable neural responses to rewards and penalties were observed; the midbrain and ventral striatum (part of the basal ganglia) showed neural sensitivity to financial rewards, while hippocampus sensitivity was linked to financial penalties. Similarly, Delgado *et al.* (2000) also used a card game associated with a monetary incentive. They found that both the dorsal and ventral striatum were activated by the paradigm. Differential responses to reward and punishment were observed, with reward feedback induced activation lasting longer than punishment. Various other studies have supported the above findings showing that reward and punishment show partially different brain system activation (Daw *et al.*, 2002). In particular, reward has been shown to be associated with frontostriatal circuits such as the ventral striatum (Daw *et al.*, 2002), while punishment has been linked to the striatum and insula (O'Doherty *et al.*, 2004).

Out of the many brain regions that have been mentioned so far, the striatum seems to be a prominently recurring structure. This is not surprising as the striatum has been implicated in multiple facets of cognition, including motor functions, decision-making, motivation, reinforcement, and reward-related processing (Delgado, 2007). It is thus worth taking a closer look at the striatum, starting off with what it is and then examining its relationship with reward-related responses. The striatum is the input structure of the basal ganglia and can be further subdivided into a dorsal portion (caudate nucleus) and a ventral portion (nucleus accumbens). Both subdivisions have been found to play a role in reward processing, with the former likely involved in the learning and updating of actions that lead to reward (Delgado, 2007).

While the previously described experiments have focused on patterns of brain activation and functional connectivity measures, other researchers have examined the relationship of punishment and reward feedback on neurotransmitters. For instance, den Ouden *et al.* (2013) found that dopamine was associated with reward and serotonin with punishment feedback. It was also suggested that reward feedback involves the recruitment of slow learning systems, such as the caudate via dopaminergic signalling, while punishment leads to the recruitment of fast learning systems, such as the medial temporal lobe (Peterson and Seger, 2013; Wachter *et al.*, 2009). Given that the caudate is part of the dorsal striatum and that the medial temporal lobe has several structures, including the hippocampus, these findings are not only relatively consistent with those mentioned earlier, but also with data reported by others (e.g., Murty *et al.*, 2012, 2016; Shigemune *et al.*, 2014).

1.6.2 Motor tasks

Two previously described experiments (see **Punishment and reward feedback during motor learning** section) used functional magnetic imaging resonance to examine the brain activity while receiving punishment or reward feedback during learning of the serial reaction time task (Wachter *et al.*, 2009; Steel *et al.*, 2016) and the force-tracking task (Steel *et al.*, 2016). Wachter *et al.* (2009) found increased activation in the dorsal striatum in their reward feedback group whereas the punishment group showed greater insula activation. These results suggest that reward and punishment feedback may be associated with different motivational systems that have distinct behavioral effects and neural substrates.

Steel *et al.* (2016) reported their neuroimaging data in a separate paper (Steel *et al.*, 2019) and found that reward and punishment feedback influenced premotor cortex functional connectivity in different ways. For the serial reaction time task, the reward group showed increased functional connectivity between the premotor cortex and the cerebellum and striatum, while the punishment group showed greater connectivity between the premotor cortex and the medial temporal lobe. For the force-tracking task, the reward group showed increased functional connectivity between the premotor cortex and the parietal and temporal cortices, while the punishment group showed the increase between the premotor cortex and the ventral striatum. Steel *et al.* (2019) concluded that these regions showed diverging patterns of results across the serial reaction time task and the force-tracking task for reward and punishment feedback. More generally, it was concluded that both feedback and motor task strongly influence spontaneous brain activity after training.

It is clear that there are numerous brain regions and circuits involved in motor learning, even when focusing on punishment and reward. There does however appear to be dissociable effects between the two in terms of recruited brain regions, and this might partially explain the differences often observed at the behavioural level.

Chapter 2

Introduction

Feedback is all around us. Consider the abundant supply of information you receive from your senses, and the various comments or criticisms your friends, family, professors, and coaches regularly share with you. It is well established that feedback has a vital role in the learning process when it is provided before, during, or after the completion of a task as compared to when that same feedback is withheld (Newell, 1977; Salmoni *et al.*, 1984; Sigrist *et al.*, 2013; Wulf and Shea, 2004). As such, motor learning scientists have sought to identify effective ways to provide feedback not only to facilitate the learning process during practice, but also long(er)-term retention. Recently, the effects of punishment and reward feedback have received considerable attention in the motor learning literature (for a review see Chen *et al.*, 2018), possibly due to their motivational effects on human behaviour (Luft, 2014). A critical feature of reward and punishment feedback that distinguishes it from other forms of feedback, such as graded (e.g., "too short") or finely-graded (e.g., -8.47 cm), is that it consists of an unsigned, categorical (e.g., "hit" or "miss") signal (Luft, 2014).

Galea *et al.* (2015) investigated the impact of reward and punishment feedback during visuomotor adaptation. During training trials, a rotation was applied to a cursor, representing the location of the participant's unseen hand, and participants had to learn to compensate for this perturbation by updating their reaching direction. Participants in the reward and punishment groups earned or lost points, respectively, which corresponded with a monetary value. Points were based on endpoint angular error and both online and endpoint feedback of the cursor were provided. Galea *et al.* (2015) found that learning was accelerated during training with punishment feedback while reward feedback resulted in greater retention during the no-vision phase. This dissociation was found for both multiple- and single-target task variations. However, the use of continuous cursor feedback in their experiments introduced error-based learning processes during visuomotor adaptation and thus, it is difficult to isolate such processes from reinforcement (i.e., reward and punishment) processes. Song *et al.* (2020) addressed this issue by removing continuous visual feedback of the cursor during training and replicated the effect of punishment feedback inducing faster learning. A retention advantage for reward feedback was not replicated. However, in both Galea et al. and Song et al. participants actually received performance-based scalar punishment and reward feedback. Providing reinforcement feedback this way makes it more similar to graded feedback, which engages error-based learning processes (Luft, 2014). Thus, it remains unclear whether previously reported dissociable effects of punishment and reward feedback are actually attributable to error-based rather than reinforcement-based learning given the sensorimotor system heavily weighs error feedback over reinforcement feedback (Cashaback *et al.*, 2017).

The effects of reward and punishment feedback—unconfounded by error-based learning processes—have been explored in the sequence learning literature. Wachter *et al.* (2009) had participants learn a serial reaction time task while receiving either punishment or reward feedback based on whether their reaction time was slower or faster than a non-updating criterion response time, respectively. Punishment feedback consisted of a visual red stimulus and a loss of 4 cents. Reward feedback included a gain of 4 cents and a visual green stimulus. They found that punishment feedback led to faster learning during training, but short-term retention was enhanced by reward feedback. Steel *et al.* (2016) recently extended this work to examine the impact on longer-term retention. Punishment and reward feedback was administered in a similar manner to Wachter *et al.* (2009) except the money lost or gained was 5 cents and the authors used an updating criterion response time. The authors found punishment feedback had the expected effect on learning in the training phase, but no advantage of reward feedback was found in the 1 hour, 24 hour, or 30 day retention tests.

While there appears to be reasonable evidence to suggest that punishment feedback can accelerate learning (Galea *et al.*, 2015; Song and Smiley-Oyen, 2017; Song *et al.*, 2020; Steel *et al.*, 2016, 2020; Wachter *et al.*, 2009), the retention benefits from reward feedback seem more tenuous given the mix of support (e.g., Abe *et al.*, 2011; Galea *et al.*, 2015; Wachter *et al.*, 2009) and non-support (e.g., Song *et al.*, 2020; Steel *et al.*, 2016, 2020). Overall, this heterogeneity of results across experiments may in part arise from variations in experimental design, how researchers compute their selected performance metrics, and/or differences in retention timescales. This in turn has made it challenging to establish best practices for providing punishment and reward feedback during motor learning.

The purpose of the present experiment was to further test whether there are dissociable effects of punishment and reward feedback on learning and retention, respectively. We designed our experiment to closely match that of Steel *et al.* (2016), with some exceptions (see Table 2.1). As such, our experiment is more closely aligned with a conceptual replication rather than a direct replication (Makel *et al.*, 2012; Nosek and Errington, 2017; Schmidt, 2009). Participants learned a serial reaction time task and received unsigned, categorical punishment or reward feedback during training. Retention was assessed one-day following the completion of the training phase. Despite the equivocal results to date, we predicted that punishment feedback would result in faster and superior learning during the training phase (Prediction 1) and that reward feedback would lead to greater retention (Prediction 2).

Table 2.1: Overview of methodological changes from Steel *et al.* (2016) and the rationale for each modification.

Steel *et al.* (2016)	**Modification**	**Rationale**
In-person data collection using a handheld device.	Online data collection using keyboard of participant's laptop or desktop computer.	The impact of COVID-19 on in-person human participant data collection.
Monetary incentive.	Entry into gift card lottery incentive, with possibility of earning more entries based on points earned.	Ease of compensating participants and permitted larger dollar amount gift cards to motivate participants.
Visual stimuli consisted of four 0s and cued stimulus was switching to an X.	Visual stimuli consisted of four white boxes and cued stimulus was switching to a black box.	Ease of implementation and a more salient stimulus change (based on comments from pilot testing).
Max response time window of 800 ms due to functional magnetic resonance imaging collection.	Max response time window of 5000 ms.	No neuroimaging was used in our experiment. As a max response time window is typically not specified with the serial reaction time task, we set a large window in case participants became distracted given data collection was online.
Black cross that was displayed during the 30 s break between blocks changed from black (25 s) to blue (5 s).	No cross for the first 25 s of the 30 s break between blocks. Black cross appeared with 5 s remaining in the block.	A cross being displayed throughout the break was not necessary for our purposes. We had the cross appear with 5 s left in an effort to grab the participant's attention and cue them to get ready as the next block would start shortly.
Punishment and reward feedback provided during pre-test, training, and post-test.	Punishment and reward feedback provided only during training.	To allow a fair comparison between the post-training and the retention test, and to use pre-test as a covariate in this analysis.

Money (5 cents) gained or lost on rewarded and punished trials, respectively.	Points (0.10 points) gained or lost on rewarded and punished trials, respectively.	As we had half the number of blocks where feedback could be provided, we doubled the value associated with punished and reward trials as we had participants in our groups start at the same score as in the original experiment.
Three familiarization blocks.	One familiarization block.	To reduce the number of overall trials and mitigate becoming bored with the task given familiarization trials were not analyzed.
Performance criterion updated after each block only if it was faster than the previous criterion.	Performance criterion was always updated based on previous block.	To ensure the performance criterion was based on current performance levels.
Three retention probes were used: 1 hour, 24 hours, and 30 days.	One retention probe was used: 24 hours.	To reduce the likelihood of participant dropout given online collection format.

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/ep5fj.

3.1 Group sequential design

We determined sample size according to Simonsohn's 2015 suggestion of setting the smallest effect size of interest to what the original experiment had 33% power to detect. Steel *et al.* (2016) had 33% power to detect an effect of $d = 0.65$; however, a more conservative estimate of $d = 0.4$ was used as our smallest effect size of interest (Lakens, 2021). We adopted a sequential analysis design in the present experiment as these are more efficient for hypothesis testing than only analyzing the data once the entire planned sample size has been collected (Dodge and Romig, 1929; Lakens, 2014; Lakens *et al.*, 2021; Wald, 1945). 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 (O'Brien and Fleming, 1979) when planning our two interim analyses (at 33 and 66% of total sample size) and one final (100% of sample size) analysis. The alpha spending function conserves the Type I error rate to 5% across all analyses. These parameters resulted in a sample size of 68, 134, and 202 total participants and corresponding alpha levels of 0.0002, 0.012, and 0.046 for each analysis, respectively. Based on the results of the first interim analysis, a decision to not continue with data collection was made (Lakens *et al.*, 2021).

3.2 Participants

To reach the required sample size of 68 participants for the first interim analysis, 80 participants were recruited as twelve participants had to be removed.[1] The 68 participants ($M_{age} = 22.57$, $SD = 2.49$ years; 35 Female, 33 Male) included in interim analysis one all self-reported being right-hand dominant and having normal or corrected-to-normal vision. Participants were randomly assigned to either the Punishment Feedback group ($M_{age} = 22.53$, $SD = 2.42$ years; 17 Female, 17 Male) or the Reward Feedback group ($M_{age} = 22.62$, $SD = 2.61$ years; 18 Female, 16 Male). Prior to beginning data collection, all participants provided their informed consent through LimeSurvey and in accordance with and approved by 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 in the experiment.

[1] One participant was removed because their hand shifted off the correct keys; two participants did not complete all of Session 1 and/or Session 2; two participants self-reported being left-handed; and seven participants had incomplete or no data collected due to server-related issues).

3.3 Task

The serial reaction time task (see Figure 3.1A) was similar to the one used in Steel *et al.* (2016), which was a variation of the original (Nissen and Bullemer, 1987). Participants were presented with four visual square stimuli arranged horizontally in the centre of their laptop or computer screen. 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. Each key was mapped to one of the four square stimuli. 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. Stimuli were presented in blocks of 96 trials, and a trial consisted of a single key press. Trials ended after either a participant pressed a key or 5000 ms elapsed (i.e., trial timed out). An inter-trial interval of 200 ms was used during which the four empty squares appeared on the screen.

Blocks consisted of either fixed sequence trials or random sequence trials. In the fixed sequence blocks, the stimuli appeared according to a fixed 12 element sequence repeated eight times. Participants were randomly assigned to one of four possible patterns[2] and performed this same pattern in all their 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). In the random sequence blocks, the stimuli appeared based on a pseudorandomly generated pattern such that the same stimuli was never presented on consecutive trials.

[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.

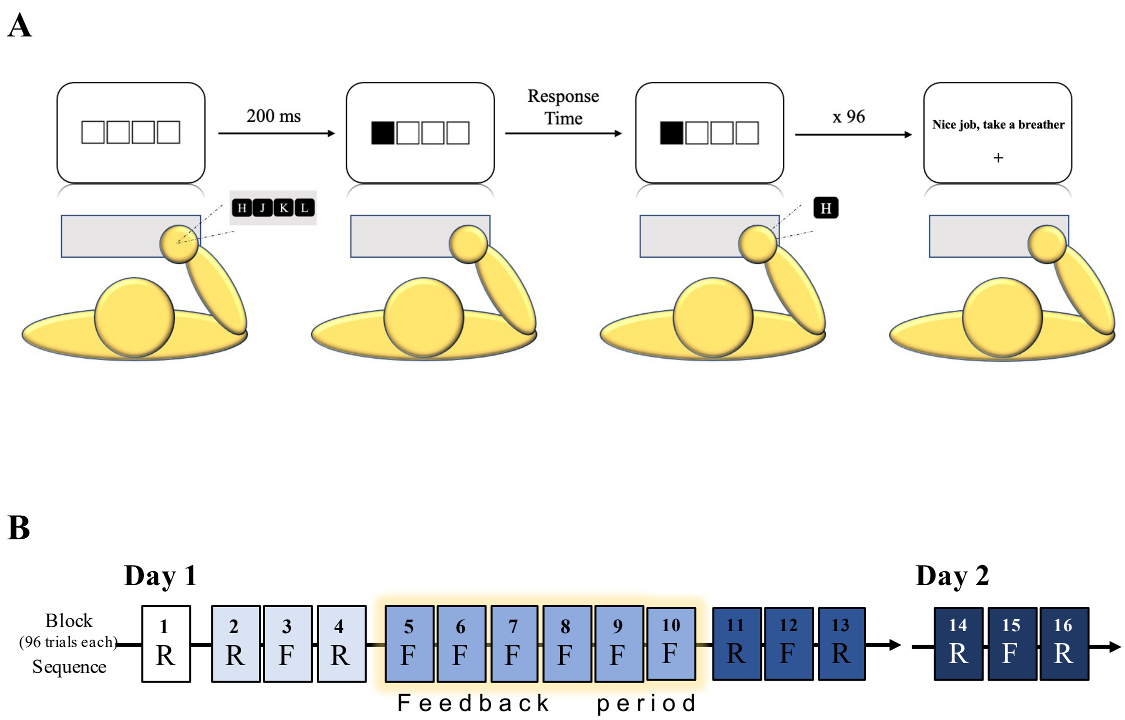

Figure 3.1: Overview of the serial reaction time task and procedure. **(A) Task setup.** Participants placed the index, middle, ring, and pinky fingers of their right hand on the H,J,K,L keys on their keyboard, respectively. Four white boxes were first shown to participants. After a fixed 200 ms interval, one box would change to a black box. In the figure, the far left box turned black which would require the participant to press the H key as quickly and accurately as possible. This sequence of events would repeat for 96 trials in each block. Blocks were separated by a 30 s rest period. At the start of the break, the phrase "Nice job, take a breather" was displayed. With 5 s remaining in the break, a black cross appeared on the screen to prompt the participant that the next block was about to begin. **(B) Overview of experimental sessions.** Participants completed two testing sessions on back-to-back days. Day 1 consisted of 4 phases which are color coded. Familiarization consisted of 1 block, Pre-training consisted of 3 blocks, Training consisted of 6 blocks, and Post-Training 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 retention test. The retention test consisted of 3 blocks and was identical to the Pre-Training and Post-Training phases.

3.4 Procedure

Participants completed two online data collection sessions on consecutive days (see Figure 3.1B). Session 1 consisted of four phases: familiarization (96 trials), pre-test (288 trials), training (576 trials), and post-test (288 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. Prior to the start of the training period, which had six fixed sequence blocks, 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. All blocks, independent of phase and session, were separated by a 30 s break. The phrase “Nice job, take a breather” was displayed on the screen for 25 s after which a black cross for the remaining 5 s. During the 30 s breaks in the training phase, the participant’s total number of points was also displayed on the screen.

Throughout the training phase, participants received either punishment or reward feedback based on their performance relative to a performance criterion on a trial-by-trial basis (Figure 3.2). 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 engage participants and encourage continuous improvement. The Punishment Feedback group received feedback as a red frame around the four visual square stimuli if their response was incorrect or slower than their performance criterion. Punished trials carried a loss of 0.10 points and each participant in the Punishment Feedback group started with 55 points. The Reward Feedback group

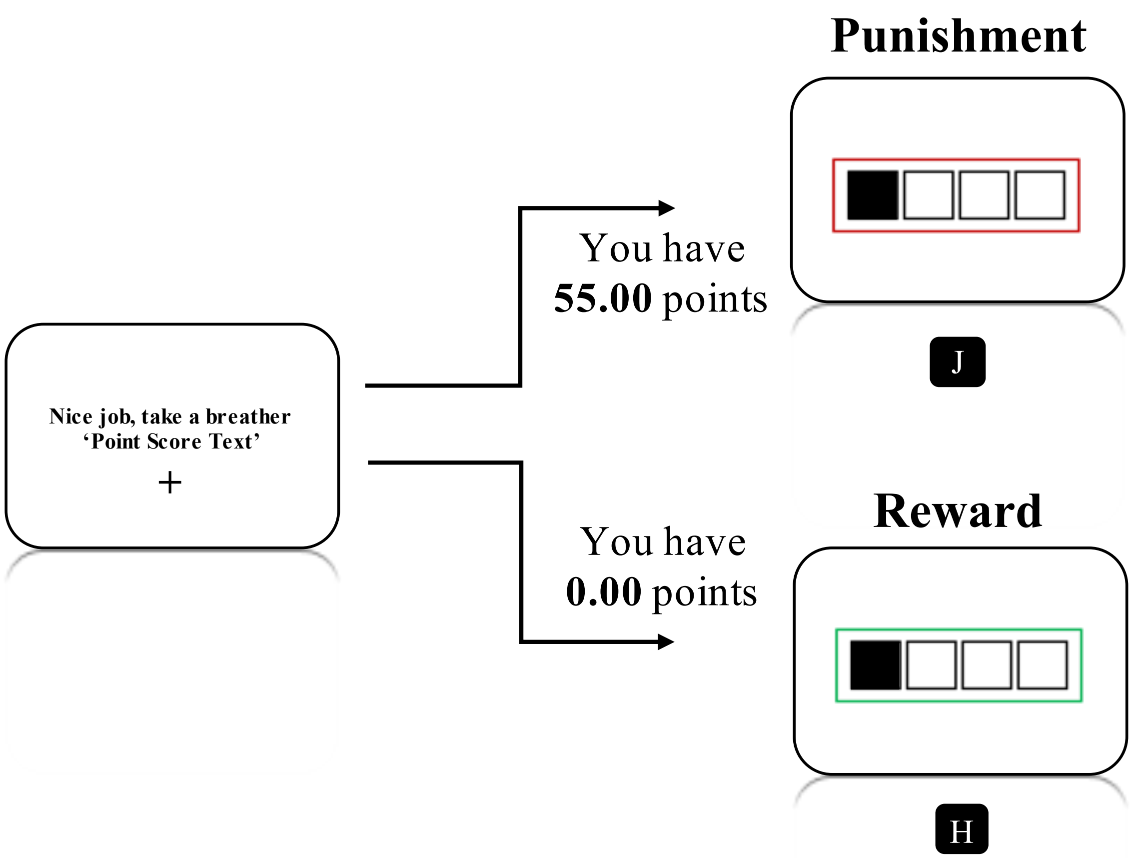

Figure 3.2: Punishment and reward feedback during the training period. **Illustrative punishment and reward feedback displays.** Following the last block of pre-training, participants completed 6 training blocks where feedback was given. Participants were told "For the upcoming trials, your performance will influence your point score shown below. A good point score at the end of the experiment will allow you to earn extra entries into the gift card draw. You have [*amount of points shown*]". The Punishment Feedback group started with 55.00 points. Each time the participant responded by pressing the incorrect key (as shown) or responded slower than their performance criterion a red frame appeared on their screen. The Reward Feedback group started with 0.00 points. Each time the participant responded by pressing the correct key (as shown) and responded faster than their performance criterion a green frame appeared on their screen. The performance criterion for both groups was always their median response time from the previous block.

received feedback as a green frame around the four visual square stimuli if their response was correct and faster than their performance criterion. Rewarded trials resulted in a gain of 0.10 points and each participant in the Reward Feedback group started with 0 points. Points were only lost or gained during training and participants were not made aware of the point value lost or gained on punished or rewarded trials.

Session 2 had a single phase, which was the delayed retention test (288 trials) that occurred approximately 24 hours after Session 1 (Kantak and Winstein, 2012). 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.

The experiment was created using jsPsych (de Leeuw, 2015) and was deployed using Pavlovia (https://pavlovia.org/). The created 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.

3.5 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 pre-registered statistical analyses

To test our predictions regarding punishment feedback and reward feedback, we compared the distribution of response times on correct trials for each group 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. The family-wise error was controlled 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). Overall, shift functions are a more powerful and robust approach to understand whether groups of observations differ (Rousselet *et al.*, 2017; Rousselet and Wilcox, 2020).

For prediction 1, the null hypothesis will be rejected if the groups are significantly

[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 (Schmidt *et al.*, 2018). 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).

different at any decile. The alternative hypothesis that punishment feedback is more effective for learning will be accepted if the Punishment Feedback group has significantly shorter response times at any decile and does not have any significantly longer response times at any decile.

For prediction 2, the null hypothesis will be rejected if the groups are significantly different at any decile. The alternative hypothesis that reward feedback is more effective for retention will be accepted if the Reward Feedback group has significantly shorter response times at any decile and does not have any significantly longer response times at any decile.

3.5.2 Secondary pre-registered statistical analyses

We also ran more traditional analyses of (co)variance to facilitate comparisons with past work and in particular, Steel *et al.* (2016). For ANOVA, univariate outliers were screened using the median absolute deviation technique with a pre-specified threshold of three (Leys *et al.*, 2019). Fourteen univariate outliers were revealed. 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). Eight univariate and eight multivariate outliers were revealed. 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 feedback during training, mean response times were analyzed in a mixed 2 (Group) x 6 (Block) ANOVA with repeated measures on Block. To test for a retention advantage of reward feedback, mean response times were analyzed in a mixed 2 (Group) x 2 (Test: Post-test, Delayed retention) ANCOVA controlling for pre-test.

Chapter 4

Results

4.1 Training

4.1.1 Primary analysis

The results from the shift function (Figure 4.1) failed to reveal significant differences between the Punishment Feedback and the Reward Feedback groups.

4.1.2 Secondary analyses

As can be seen in (Figure 4.2), both groups of participants decreased their response times over the training period, which was supported by a significant Block main effect, $F(3.02, 199.37) = 21.297$, $p < 0.001$. Tukey's post hoc comparisons revealed that Block 1 response times were significantly slower than all other Blocks. Neither the Group main effect, $F(1, 66) = 0.007$, $p = 0.932$, nor Group x Block interaction were significant, $F(3.02, 199.37) = 0.584$, $p = 0.627$.

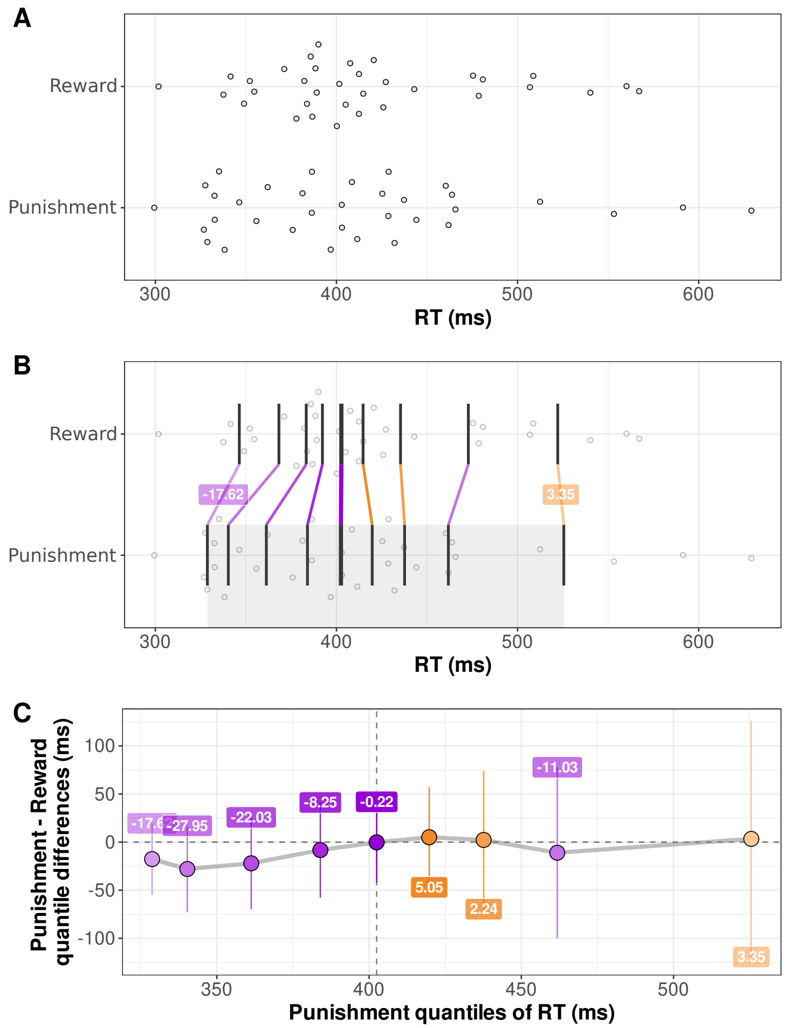

Figure 4.1: Shift function reveals no significant differences between groups at any decile during training. **(A) Scatterplot.** Response time is shown on the x-axis with Punishment Feedback and Reward Feedback on the y-axis. Data appears skewed for both groups with the center appearing to be around 400 ms. **(B) Scatterplot with deciles overlayed.** Overlayed on the scatterplot from (A) are data for both groups which has been divided into deciles, shown through faded long vertical black lines. Each group's decile median is shown through reinforced short vertical black lines. A purple connecting line indicates that the Punishment Feedback group had a shorter response time in that decile, while an orange connecting line indicates that the Reward Feedback group had a shorter response time. For example, -17.62 shown in purple indicates the difference between groups for decile 1, while 3.35 shown in orange indicates the difference between groups for decile 9. **(C) Shift function.** Punishment quantiles of RT in (ms) is shown on the x-axis, and punishment minus reward quintile differences in (ms) is shown on the y-axis. The colors and values have identical meanings to those shown in (B). Adjusted 95% confidence intervals are shown for each decile median. All confidence intervals are crossing the midline (0) and thus, there are no significant differences between groups at any decile.

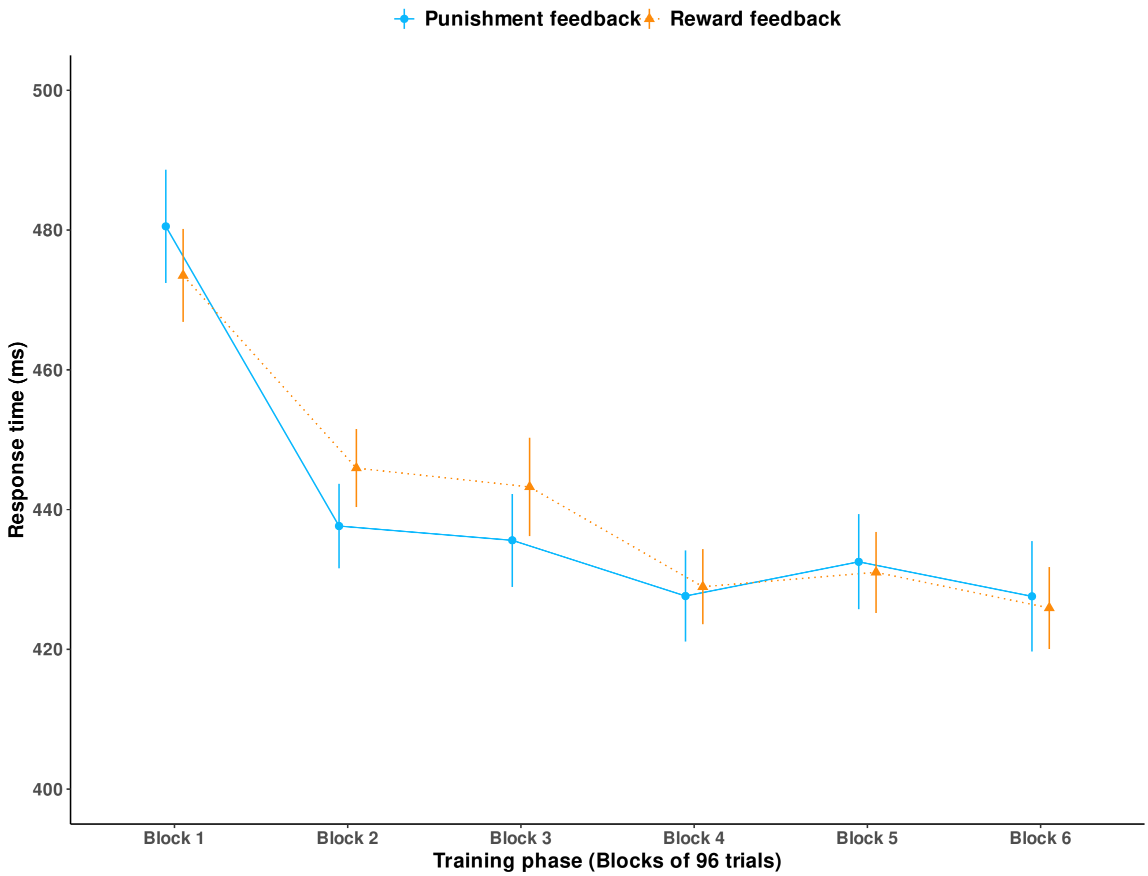

Figure 4.2: No significant differences detected between groups with secondary statistical analysis of training. Training phase is shown on the x-axis with response time in (ms) on the y-axis. The Punishment Feedback group is shown using a blue solid line and circular points. The Reward Feedback group is shown using an orange dotted line and triangular points. Error bars are 95% confidence intervals. No significant main effect of Group or Block X Group interaction; both lines appear to be following similar trends. A significant effect of Block is evident as response time (ms) declines for both groups as training progresses.

4.2 Retention

4.2.1 Primary analysis

The shift function (Figure 4.3) failed to reveal any significant differences between Punishment Feedback group and the Reward Feedback group at any decile during retention.

4.2.2 Secondary analysis

Response times (Figure 4.4) were numerically faster in post-training and retention compared to pre-test. However, at all time points the differences between the Punishment Feedback and Reward Feedback groups were minimal. The main effects of Group, $F(1, 65) = 2.644$, $p = 0.109$, and Test, $F(1, 65) = 0.066$, $p = 0.798$, were not significant. The Group x Test interaction was also not significant, $F(1, 65) = 1.239$, $p = 0.270$.

4.3 Equivalence and inferiority tests

Equivalence and inferiority tests were conducted comparing results to our smallest effect size of interest ($d = 0.4$, $\alpha = 0.05$). For the training period, the equivalence test was not significant, $t(63.8) = 1.569$, $p = 0.0608$, and the null hypothesis test was also not significant, $t(63.8) =$ -0.0801, $p = 0.936$. For delayed retention, the equivalence test was not significant, $t(61.51) =$ -0.156, $p = 0.562$, and the null hypothesis test was also not significant, $t(61.51) =$ -1.805, $p = 0.0759$.

An inferiority test was conducted on our training data based on Steel et al. (2016)

33% power (d = 0.65, α = 0.05). The equivalence test was significant, $t(63.8)$ = 2.476, p = 0.00797, and the null hypothesis test was not significant, $t(63.8)$ = -0.0801, p = 0.936. As such, the observed effect was statistically not different from zero and statistically equivalent to zero based on the combined equivalence test and null-hypothesis test.

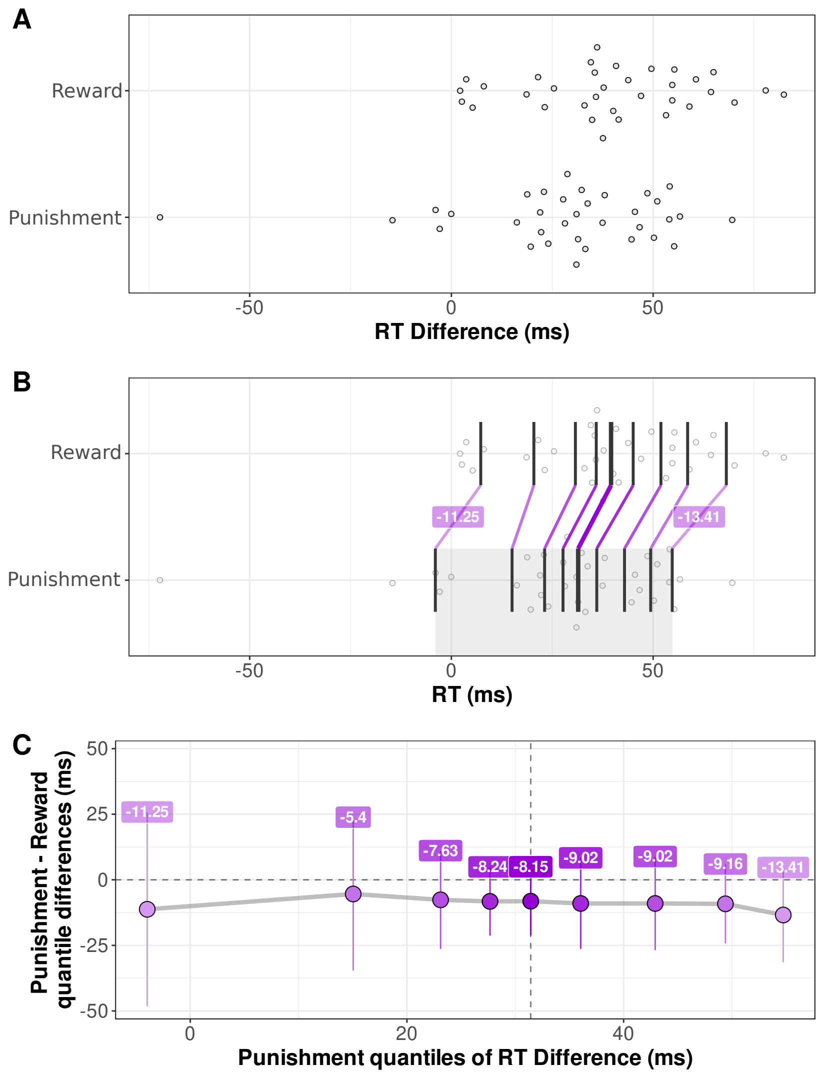

Figure 4.3: Shift function reveals no significant differences between groups at any decile during retention. **(A) Scatterplot.** Response time (RT) difference is shown on the x-axis with Punishment Feedback and Reward Feedback on the y-axis. Data again appears skewed for both groups. **(B) Scatterplot with deciles overlayed.** Refer to Figure 4.1 for in depth figure explanation. Here, -11.25 shown in purple indicates the difference between groups for decile 1, while -13.41 indicates the difference between groups for decile 9. The Punishment Feedback group had a shorter response time for both. **(C) Shift function.** Refer to Figure 4.1 for in depth figure explanation. All confidence intervals are crossing the midline (0) and thus, there are no significant differences between groups at any decile.

Figure 4.4: No significant differences detected between groups during secondary statistical analysis of retention. Pre-test, post-test, and delayed retention are shown on the x-axis with Random minus Fixed response time in (ms) on the y-axis. The vertical dotted line separates Session 1 and Session 2 data. The Punishment Feedback group is shown using blue error bars (95% confidence intervals) and circular data points. The Reward Feedback group is shown using orange error bars (95% confidence intervals) and triangular data points. The smaller points represent the individual score of each participant while the larger points represent the group mean. Both groups had a similar level of performance at all time points. No significant effects were detected.

Chapter 5

Discussion

The purpose of the present experiment was to test the replicability of punishment and reward feedback having dissociable effects on learning and retention, respectively (e.g., Galea *et al.*, 2015; Wachter *et al.*, 2009). To this end, we designed our pre-registered experiment to closely match the methods used by Steel *et al.* (2016) as they used a sequence learning paradigm,[1] which was amenable to online data collection due to COVID-19. Contrary to our predictions, we did not find an advantage of receiving punishment feedback during training, nor did we find a retention advantage of reward feedback. Based on the outcomes of the first interim analysis of our sequential design, data collection will continue to the second interim analysis (for a discussion see Lakens *et al.*, 2021). Our results after interim analysis one are not in line with the simple heuristic that punishment feedback enhances learning during practice and reward feedback benefits retention. Instead, the data suggests that differences between these two types of feedback on the serial reaction time task might be negligible.

[1]see Table 2.1 for an overview of methodological differences between experiments.

5.1 No learning benefit of punishment feedback

Both the primary (i.e., shift function) and secondary (i.e., mixed-design ANOVA) analyses converged on no learning advantage of punishment feedback during training. This finding is inconsistent with the extant literature that has reported faster and overall better learning during training with punishment feedback (e.g., Galea *et al.*, 2015; Wachter *et al.*, 2009). This failure to replicate the training benefits of punishment feedback (but see Abe *et al.*, 2011, for a similar finding) was surprising given other researchers have actually argued that punishment-related feedback effects are more reliable than those associated with reward feedback (e.g., Song and Smiley-Oyen, 2017; Song *et al.*, 2020; Steel *et al.*, 2016, 2020).

The training 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 carry 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 with previous research.

5.2 No retention advantage of reward feedback

Similar to our analyses on the training data, the primary (i.e., shift function) and secondary (i.e., ANCOVA controlling for pre-test) analyses showed no retention advantage for reward feedback. Although this finding was not in line with our predictions based on past work (e.g., Abe *et al.*, 2011; Galea *et al.*, 2015; Wachter *et al.*, 2009), it is consistent with those reported by other researchers (e.g., Song and Smiley-Oyen, 2017; Song *et al.*, 2020; Steel *et al.*, 2016, 2020). This raises the question as to why the impact of reward feedback on retention is so mixed. One possible explanation for this is the variations in how feedback has been provided to participants across experiments. Galea *et al.* (2015) provided participants with both online visual feedback and endpoint feedback of the cursor in addition to scalar punishment and reward feedback. Feedback was therefore a mix of binary, graded, and finely-graded performance information, with finely-graded information (i.e., cursor feedback) available on every trial during training. As such, the found retention advantage in Galea *et al.* (2015) may in part arise from participants engaging both error-based and reinforcement learning processes (Luft, 2014) during training. This may have had a greater contribution to the retention benefits than previously thought as the retention advantage has been shown to disappear when finely-graded cursor information—but not the scalar (i.e., graded) punishment and reward feedback—is removed (Song *et al.*, 2020). It is important to not mix error-based and reinforcement learning processes through feedback provision as under such conditions, there is a greater reliance on error-based learning processes (Cashaback *et al.*, 2017).

However, even when punishment and reward feedback are provided as unsigned categorical feedback, the impact on retention is mixed. Wachter *et al.* (2009), for instance, found a short-term retention advantage of reward feedback, whereas neither a short- or longer-term retention advantage was found by Steel and colleagues (Steel *et al.*, 2016, 2020). Given our design was modeled off Steel *et al.* (2016), it is not surprising that we found similar results to them; albeit, our feedback period was half the number of their feedback trials. A difference between Wachter *et al.* (2009) and Steel *et al.* (2016), as well as our experiment is the frequency that the participant's current money or points total was available to them. In our experiment and Steel *et al.* (2016), the total score was only shown in between blocks during the break, whereas Wachter *et al.* (2009) made the score available at all times. Although in each of these experiments, participants were not explicitly told their feedback manipulation or associated value of punished and reward trials, the real-time updating of the total score likely allowed participants to better associate their performance with changes in their total score. This may have altered the motivational salience of the feedback compared to only seeing your score in between blocks where the cause of changes to the total score would not have been as obvious to the participants. Future research in this area will need to consider the availability of the participant's money bank when designing their experiments as this appears to influence the effectiveness of reward feedback on short-term retention.

5.3 Statistical power and design issues

While any of the above mentioned reasons may have contributed in some way to our failure to replicate the dissociable effects of punishment and reward feedback in motor learning, we contend that a more plausible explanation is that both of these effects are likely 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.*, 2021; Simmons *et al.*, 2011). 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 (Button *et al.*, 2013; Gelman and Carlin, 2014), 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 0.65. 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 ($d = 0.52$; Simmons *et al.*, 2013). Additionally, the outcome of our inferiority test relative to Steel *et al.* (2016) suggests that these effects, if they in fact exist, are too small to have been reliably detected by their design. Second, it can increase the probability of making a signed error—an error in which the results of an experiment are estimated in the wrong 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).

Taken together, we believe issues surrounding replicability and low statistical power have contributed to the notion of dissociable effects of punishment and reward feedback on learning and retention, respectively. Although we did not collect to our fully planned sample size and are thus, underpowered relative to our smallest effect size of interest, our design and analyses allowed for greater statistical power compared to past work even at our first interim analysis. Consequently, these effects of punishment and reward feedback are either likely much smaller than previously estimated or there may in fact be no effect at all.

Chapter 6

Conclusion

Overall, the results from this study failed to find dissociable effects of punishment and reward feedback on learning and retention, respectively. This has led us to conclude that the impact of punishment and reward feedback on sequence learning may be smaller than previously expected or may possibly be absent altogether.

6.1 Limitations

The findings of this study have some possible limitations. As already noted above, this study utilized an online method of data collection, which leads to no possible way of preventing environmental distractions when compared to a quiet laboratory setting. Additionally, at home settings do not have researchers present; participants were unable to conveniently ask questions, sometimes resulting in misunderstandings and execution errors. To give an example, a participant had their fingers placed on G, H, J, K instead of H, J, K, L. Only during Day 2 did the participant realize that a shift had taken place; the trials were now passing quickly with each press resulting in the commencement of the next trial, instead of waiting 5000 ms each time the first box was filled as G was not programmed to mount a response. Had data collection been in person, the participant would have likely asked why in some cases trial transition was not immediate or the researcher would have realized and corrected the shift thus avoiding having to collect an additional participant to replace the current one due to error. This, in conjunction with internet connection and third-party online mediator issues led to the data removal of twelve participants from the study.

Related to the limitation above, online data collection led to no scheduled in-person laboratory appointments. Thus, there was variability between participants in terms of retention collection times. For example, some participants completed Day 1 at night and Day 2 the next morning, while others completed Day 1 in the morning and Day 2 at night, resulting in fairly different time gaps. While this variability was equally likely to be present in both groups, and is thus unlikely to have affected results, it may have been a contributing factor to our retention findings. However, it is important to note that there are still many benefits to online data collection and that this method should not be discredited. If the above limitations can be addressed, for instance through live online assistance or a tighter scheduling window, then the advantages of online collection may outweigh the challenges as online data collection allows sampling from a more diverse population, a shorter collection period, greater flexibly to cater to participants' schedules, and closely mimic the real world in which supervision may not always be an option.

6.2 Future directions

Should researchers wish to continue to examine why there are inconsistent results regarding punishment and reward feedback in motor learning, we strongly encourage them to justify their sample size decisions (Lakens, 2021), use unsigned categorical feedback manipulations to not confound different learning processes (Cashaback *et al.*, 2017), and to either pre-register their experimental design and analysis plan (Munafò *et al.*, 2017) or use the Registered Report format (Scheel *et al.*, 2021). We believe these will address issues that have contributed to issues with replicability and will also contribute to more transparent reporting in motor learning research.

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. *Curr. Biol.*, **21**(7), 557–562.

Abrahamse, E. L., Ruitenberg, M. F. L., de Kleine, E., and Verwey, W. B. (2013). Control of automated behavior: insights from the discrete sequence production task. *Frontiers in Human Neuroscience*, **7**, 82.

Bock, O. (1992). Adaptation of aimed arm movements to sensorimotor discordance: evidence for direction-independent gain control. *Behavioral Brain Research*, **51**(1), 41–50.

Broadbent, D. E. (1958). *Perception and communication*. Pergamon Press.

Brown, R. M. and Robertson, E. M. (2007). Inducing motor skill improvements with a declarative task. *Nat. Neurosci.*, **10**(2), 148–149.

Bütefisch, C. M., Davis, B. C., Wise, S. P., Sawaki, L., Kopylev, L., Classen, J., and Cohen, L. G. (2000). Mechanisms of use-dependent plasticity in the human motor cortex. *Proc. Natl. Acad. Sci. U. S. A.*, **97**(7), 3661–3665.

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.

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 Comput. Biol.*, **13**(7), e1005623.

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.

Classen, J., Liepert, J., Wise, S. P., Hallett, M., and Cohen, L. G. (1998). Rapid plasticity of human cortical movement representation induced by practice. *J. Neurophysiol.*, **79**(2), 1117–1123.

Cunningham, H. A. (1989). Aiming error under transformed spatial mappings suggests a structure for visual-motor maps. *Journal of Experimental Psychology: Human, Perception, & Performance*, **15**(3), 493–506.

Daw, N. D., Kakade, S., and Dayan, P. (2002). Opponent interactions between serotonin and dopamine. *Neural Netw.*, **15**(4-6), 603–616.

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.

Delgado, M. R. (2007). Reward-related responses in the human striatum. *Ann. N. Y. Acad. Sci.*, **1104**(1), 70–88.

Delgado, M. R., Nystrom, L. E., Fissell, C., Noll, D. C., and Fiez, J. A. (2000). Tracking the hemodynamic responses to reward and punishment in the striatum. *J. Neurophysiol.*, **84**(6), 3072–3077.

den Ouden, H. E. M., den Ouden, H. E. M., Daw, N. D., Fernandez, G., Elshout, J. A., Rijpkema, M., Hoogman, M., Franke, B., and Cools, R. (2013). Dissociable effects of dopamine and serotonin on reversal learning. *Neuron*, **80**(6), 1572.

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

Diedrichsen, J., White, O., Newman, D., and Lally, N. (2010). Use-dependent and error-based learning of motor behaviors. *J. Neurosci.*, **30**(15), 5159–5166.

Dodge, H. F. and Romig, H. G. (1929). A method of sampling inspection. *The Bell System Technical Journal*, **8**(4), 613–631.

Elliott, R., Friston, K. J., and Dolan, R. J. (2000). Dissociable neural responses in human reward systems. *J. Neurosci.*, **20**(16), 6159–6165.

Fernández-Ruiz, J. and Díaz, R. (1999). Prism adaptation and aftereffect: specifying the properties of a procedural memory system. *Learning & Memory*, **6**(1), 47–53.

Flanagan, J. R., Bowman, M. C., and Johansson, R. S. (2006). Control strategies in object manipulation tasks. *Current Opinion in Neurobiology*, **16**(6), 650–659.

Franklin, D. W., Burdet, E., Tee, K. P., Osu, R., Chew, C.-M., Milner, T. E., and Kawato, M. (2008). CNS learns stable, accurate, and efficient movements using a simple algorithm. *Journal of Neuroscience*, **28**(44), 11165–11173.

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

Gelman, A. and Carlin, J. (2014). Beyond power calculations: Assessing type S (sign) and type M (magnitude) errors. *Perspectives in Psychological Science*, **9**(6), 641–651.

Hardwick, R. M., Rottschy, C., Miall, R. C., and Eickhoff, S. B. (2013). A quantitative meta-analysis and review of motor learning in the human brain. *Neuroimage*, **67**, 283–297.

Held, R. and Freedman, S. J. (1963). Plasticity in human sensorimotor control. *Science*, **142**(3591), 455–462.

Held, R. and Schlank, M. (1959). Adaptation to disarranged Eye-Hand coördination in the Distance-Dimension. *American Journal of Psychology*, **72**(4), 603–605.

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

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

Jeannerod, M. (1988). The neural and behavioural organization of goal-directed movements. *Oxford psychology series, No. 15.*, **283**.

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.

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

Korman, M., Raz, N., Flash, T., and Karni, A. (2003). Multiple shifts in the representation of a motor sequence during the acquisition of skilled performance. *Proc. Natl. Acad. Sci. U. S. A.*, **100**(21), 12492–12497.

Krakauer, J. W. (2009). Motor learning and consolidation: the case of visuomotor rotation. *Adv. Exp. Med. Biol.*, **629**, 405–421.

Krakauer, J. W., Pine, Z. M., Ghilardi, M.-F., and Ghez, C. (2000). Learning of visuomotor transformations for vectorial planning of reaching trajectories. *The Journal of Neuroscience*, **20**(23), 8916–8924.

Krakauer, J. W., Hadjiosif, A. M., Xu, J., Wong, A. L., and Haith, A. M. (2019). Motor learning. *Comprehensive Physiology*, pages 613–663.

Lakens, D. (2014). Performing high-powered studies efficiently with sequential analyses. *Eur. J. Soc. Psychol.*, **44**(7), 701–710.

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

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

Lee, T. D., Swinnen, S. P., and Serrien, D. J. (1994). Cognitive effort and motor learning. *Quest*, **46**(3), 328–344.

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.

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.

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

Makel, M. C., Plucker, J. A., and Hegarty, B. (2012). Replications in psychology research: How often do they really occur? *Perspectives on Psychological Science*, **7**(6), 537–542.

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*, **119 (Pt 4)**, 1183–1198.

Moisello, C., Crupi, D., Tunik, E., Quartarone, A., Bove, M., Tononi, G., and Ghilardi, M. F. (2009). The serial reaction time task revisited: a study on motor sequence learning with an arm-reaching task. *Exp. Brain Res.*, **194**(1), 143–155.

Moller, A. C., Elliot, A. J., and Maier, M. A. (2009). Basic hue-meaning associations. *Emotion*, **9**(6), 898.

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.

Murty, V. P., Labar, K. S., and Adcock, R. A. (2012). Threat of punishment motivates memory encoding via amygdala, not midbrain, interactions with the medial temporal lobe. *J. Neurosci.*, **32**(26), 8969–8976.

Murty, V. P., LaBar, K. S., and Adcock, R. A. (2016). Distinct medial temporal networks encode surprise during motivation by reward versus punishment. *Neurobiol. Learn. Mem.*, **134 Pt A**, 55–64.

Newell, K. M. (1977). Knowledge of results and motor learning. *Exercise and Sport Sciences Reviews*, **4**, 195–228.

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

Nosek, B. A. and Errington, T. M. (2017). Making sense of replications. *eLife*, **6**.

O'Brien, P. C. and Fleming, T. R. (1979). A multiple testing procedure for clinical trials. *Biometrics*, **35**(3), 549–556.

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., and Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science*, **304**(5669), 452–454.

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

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

Redding, G. M., Rossetti, Y., and Wallace, B. (2005). Applications of prism adaptation: a tutorial in theory and method. *Neurosci. Biobehav. Rev.*, **29**(3), 431–444.

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: problems with the mean and the median. *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. *Eur. J. Neurosci.*, **46**(2), 1738–1748.

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.

Scheel, A. M., Schijen, M. R. M. J., and Lakens, D. (2021). An excess of positive results: Comparing the standard psychology literature with registered reports. *Advances in Methods and Practices in Psychological Science*, **4**(2), 25152459211007467.

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. (1991). Frequent augmented feedback can degrade learning: Evidence and interpretations. *Tutorials in Motor Neuroscience*, pages 59–75.

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., Winstein, C., Wulf, G., and Zelaznik, H. (2018). *Motor control and learning, 6th Edition*. Human Kinetics.

Schmidt, S. (2009). Shall we really do it again? the powerful concept of replication is neglected in the social sciences. *Review of General Psychology*, **13**(2), 90–100.

Shadmehr, R. and Mussa-Ivaldi, F. A. (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.

Shigemune, Y., Tsukiura, T., Kambara, T., and Kawashima, R. (2014). Remembering with gains and losses: Effects of monetary reward and punishment on successful encoding activation of source memories. *Cerebral Cortex*, **24**(5), 1319–1331.

Shmuelof, L., Krakauer, J. W., and Mazzoni, P. (2012a). 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.

Shmuelof, L., Huang, V. S., Haith, A. M., Delnicki, R. J., Mazzoni, P., and Krakauer, J. W. (2012b). Overcoming motor "forgetting" through reinforcement of learned actions. *Journal of Neuroscience*, **32**(42), 14617–14621a.

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. (2011). 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. *Psychological Science*, **26**(5), 559–569.

Smith, M. A. and Shadmehr, R. (2005). Intact ability to learn internal models of arm dynamics in huntington's disease but not cerebellar degeneration. *Journal of Neurophysiology*, **93**(5), 2809–2821.

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.

Spampinato, D. and Celnik, P. (2021). Multiple motor learning processes in humans: Defining their neurophysiological bases. *Neuroscientist*, **27**(3), 246–267.

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).

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.

Steel, A., Baker, C. I., and Stagg, C. J. (2020). Intention to learn modulates the impact of reward and punishment on sequence learning. *Scientific Reports*, **10**(1), 8906.

Sutton, R. S. and Barto, A. G. (1998). *Introduction to reinforcement learning*, volume 135. MIT Press Cambridge.

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.

Treisman, A. M. (1964). Selective attention in man. *British Medical Bulletin*, **20**(1), 12–16.

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

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.

Wälchli, M., Ruffieux, J., Bourquin, Y., Keller, M., and Taube, W. (2016). Maximizing performance. *Medicine & Science in Sports & Exercise*, **48**(4), 714–719.

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

Wiestler, T. and Diedrichsen, J. (2013). Skill learning strengthens cortical representations of motor sequences. *eLife*, **2**.

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 andSimulation*, **84**(7), 1543–1551.

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

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

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.

Wulf, G. and Shea, C. H. (2004). Understanding the role of augmented feedback: The good, the bad, and the ugly. In A. M. Williams and N. J. Hodges, editors, *Skill acquisition in sport: Research, theory and practice*, volume 121, page 144. Routledge.