

3 Errors, rewards, and reinforcement in motor skill learning

*Keith Lobse, Matthew Miller, Mariane Bacelar,
and Olav Krigolson*

A brief introduction to reinforcement learning

Imagine a young skier in a tight turn on fast terrain. If they fall at the apex of the turn, how do they determine where the error occurred? Many factors could have contributed to this error (e.g., leaning too far back, not enough weight on the outside ski) and it can be difficult for the athlete to determine these errors on their own. Imagine then, that you are the coach of this athlete. What sort of feedback do you provide and when? Do you try to correct only the most egregious errors or reward the near-misses and successful performances?

Popularly, we believe that we learn from our mistakes. As it turns out, a more correct way to frame this belief is that we learn when outcomes deviate from expectations. Consider the above skiing example – learning occurs when the outcome (falling) is different from the expectation (not falling). These differences between expectations and outcomes are coined *prediction errors*. Prediction errors, if successfully identified, are a valuable signal that our behavior needs to change in order to increase the likelihood of being successful. This process of adjusting behavior based on errors to optimise rewards is referred to as *reinforcement learning* (Rescorla & Wagner, 1972; Sutton & Barto, 1998). Reinforcement learning, alongside unsupervised and supervised learning, is one of the dominant modes of learning and is a critical component of motor skill acquisition (Seidler, Kwak, Fling, & Bernard, 2013; Yarrow, Brown, & Krakauer, 2009). In this chapter, we will argue that reinforcement learning is probably the most important learning model for understanding skill acquisition.

Before explaining how reinforcement learning affects the acquisition of motor skills, and how coaches/teachers can use these reinforcement principles in practice, it is important to define some terms and models of motor skill learning. *Unsupervised learning* is when a series of inputs (e.g., sensory information) get mapped onto a series of outputs without any formal teaching “signal” (Dayan, 2009). In the brain, many neurons in the cortex operate on unsupervised learning principles (e.g., the neurons that fire together increase the strength of their connection; Hebb, 1961). For instance, many aspects of perception and multi-sensory integration rely on unsupervised learning, grouping inputs based on statistical regularities rather than explicit training. Indeed, this kind of learning

relates to why we are often subject to optical illusions (when statistical regularities are violated) or why an athlete might have strong tendencies toward certain behaviors (e.g., I might adopt my most habitual/comfortable movement patterns, even when they are inappropriate for the context).

Supervised learning is when a set of inputs produces an output that can be immediately contrasted against a known, “correct” output (Sutton & Barto, 1998). This process is analogous to being corrected by a teacher, hence “supervised” learning. In motor learning, supervised learning mechanisms are thought to work in the cortico-cerebellar networks (Doya, 2000; Haith & Krakauer, 2013), which are critical for motor adaptation. Actual sensory feedback arrives in the cerebellum from the ascending afferent pathways of the spinal cord. Around the same time, in theory, the *expected/anticipated* sensory feedback arrives in the cerebellum through descending pathways. This predicted sensory feedback is generated by a hypothetical forward model of the motor skill and can be compared against the actual outcome. This comparison creates a *prediction-error*. This prediction-error is the difference between actual and the predicted sensory consequences of the movement. If the error is small (or zero), this would mean that the movement is well calibrated, because the predicted sensory information matches the actual sensory information. If the error is large, this error serves as a signal for recalibrating our movement. An everyday example of this is picking up a jug of milk that we think is full but turns out to be empty. We initially generate forces appropriate for a full jug, which are too much, and the sensory feedback tells us we are moving the jug faster than predicted. This rapidly generated error is a signal to reduce the force and can lead to faster corrections than errors signaled by other longer feedback loops (e.g., correcting errors through visual feedback). These error signals can also be used to adapt subsequent movements. In sport, these error signals are critical for adaptation to different conditions such as an over-inflated versus a normal ball, a wet pitch or a dry pitch, or picking up a pool cue that is lighter/heavier than the one we typically use (e.g., our first attempts might be off, but we can quickly recalibrate these movements through cortico-cerebellar networks).

In reinforcement learning, an action (which could also be a long sequence of smaller actions) leads either to a success and thus reward, or a failure and thus no reward (or even a punishment; Rescorla & Wagner, 1972; Sutton & Barto, 1998). Unlike supervised learning, there is no teaching signal that tells us what the “correct” output should be, so the actor must make a change to the action, but it is often not clear how or what to change. Consider a discus thrower who is learning to improve her technique. When her throw comes up shorter than desired (i.e., failure), what is the correct change to make? The obvious, but unhelpful, answer is “throw farther,” but the real question is how can this be accomplished in a skill that unfolds over time with so many redundant degrees of freedom? Under supervised conditions, this question is aided through some sort of model or feedback specifying the “correct” technique. Another source of difficulty is that the feedback is delayed from where the error/s occurred; errors could arise anywhere from the initiation of the

throw to the moment of release. The actor must decide *where* the movement got off track and *how* it needs to be corrected. As such, reinforcement learning error signals can be employed for offline movement corrections. That is, following an error, new movements are explored and discovered, hopefully taking the learner closer to success.

The example above serves to highlight some key components of reinforcement learning models. First, it is often not clear exactly *what* mistake was made or *when* (referred to as the “credit assignment” and “temporal credit assignment” problems respectively; Dam, Kording, & Wei, 2013). Second, this ambiguity means that an actor often must go through some trial-and-error search to find out which changes bring them closer to success. Third, the actor needs to learn from delayed reward. This delay means that the actor might have to go through many attempts, getting progressively closer to the goal. However, since there is a reward prediction error on each trial, attempts that come closer to achieving a goal are rewarded, even if the goal is not actually achieved. Delayed reward also means that the actor might have to tolerate a temporary reduction in reward to find a movement pattern that is even more successful. In reinforcement learning terms, this means leaving a “local” maxima to find the “global” maxima for rewards. In coaching terms, this means convincing an athlete that changing their technique will ultimately be worth it, despite a temporary reduction in performance. Importantly, the coach should reward an athlete for trials wherein proper technique is exhibited even if the athlete is far off target, rather than trials wherein the athlete comes close to the target but uses incorrect technique. Within a person’s brain, unsupervised, supervised and reinforcement learning are operating simultaneously in neural networks. All these ways of learning contribute to our ability to adjust/update our behavior, but we argue that reinforcement learning principles are the most important for coaches and practitioners to understand in skill acquisition.

Below, we briefly discuss mechanisms of reinforcement learning in the human brain and how they help us to adjust our performance over a short time-scale, which we will refer to as *adaptation*, see Figure 3.1A. We also discuss how these error signals help us make relatively permanent changes in performance over a long time-scale, which we will refer to as *learning* (Adams, 1971; Schmidt & Lee, 2011). It is important to separate adaptation during practice from learning over the long-term, because behavioral research shows that these changes are not necessarily correlated (sometimes even being negatively correlated; Katak & Winstein, 2012). Next, we will talk about the concept of “reward” and provide examples to show its breadth. In animal studies, rewards are often associated with obtaining sweet foods. In human studies, however, successful actions may be rewarded with money (extrinsic motivators) or more abstract concepts like feelings of competence (intrinsic motivators). In the latter half of the chapter, we will review applied research on how reinforcement learning principles have been used to improve motor learning and performance. Finally, we will comment on some exciting future directions for research and address how reinforcement learning principles can be used by coaches and practitioners.

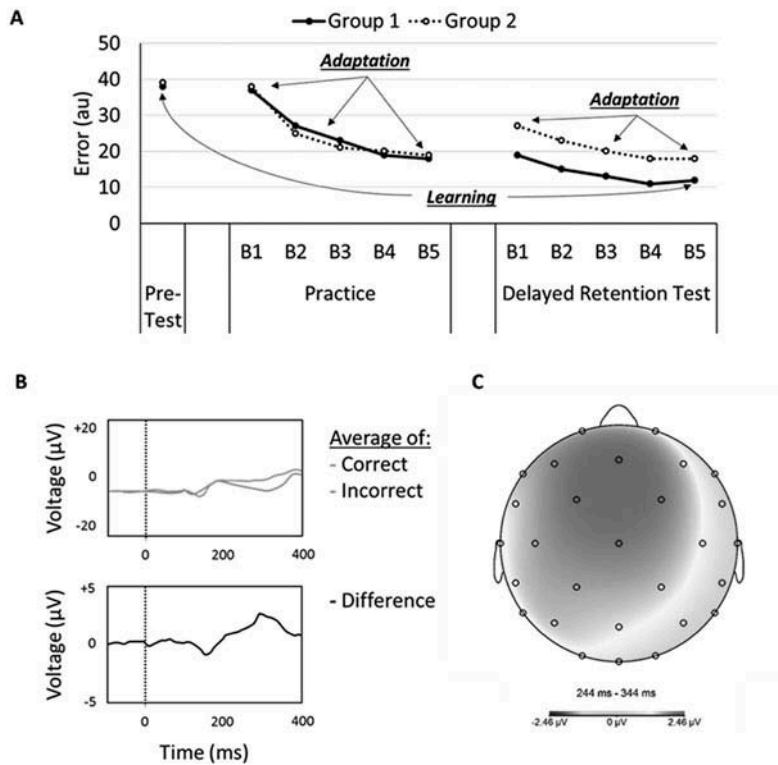


Figure 3.1 (A) Our conceptual model showing the difference between adaptation, the change in performance within a session, and learning, a relatively permanent change in performance across sessions. (B) Example event-related potentials for correct and incorrect trials (top panel) that can be subtracted to create a difference wave, bottom panel, in which the reward positivity component is visible and maximal at around 300 ms relative to onset of augmented feedback at 0 ms. (C) The scalp topography shows that the reward positivity component is maximal in the fronto-central region.

Reinforcement learning in the brain and behavior

One of the major drivers of reinforcement learning are *reward-prediction errors*. Reward-prediction error is the degree to which an actual reward differs from the anticipated/predicted reward (Holroyd & Coles, 2002). These errors can result from rewards being better (positive) or worse (negative) than predicted, and both positive and negative reward-prediction errors influence future behavior. For instance, when you are first learning a skill, you make many errors, so the brain predicts a low probability of reward. We phrase it as “the brain” rather than “the learner”, because these rapid assessments are non-

conscious, although correlated with a learner's conscious assessment of their ability (Holroyd & Krigolson, 2007). Thus, the first time you are successful (e.g., sink your first long putt; hit your first ace), there is a large positive reward-prediction error, because that movement produced an unexpectedly rewarding result. This positive reward-prediction error acts as a signal within the brain to "stamp in" that movement pattern/decision. Conversely, if you are performing well and make an error, this would lead to a large negative reward-prediction error, because the brain predicted a high probability of reward. This negative reward-prediction error can act as a signal to "stamp out" an erroneous movement pattern/decision. Importantly, success is not only defined by sinking a putt/hitting an ace, rather success (and failure) are relative to the performer, such that a near miss may elicit a positive reward-prediction error for a novice, but a negative reward-prediction error for an expert.

There are two important points to take away from the ideas of reward and associated reward-prediction errors. First, unexpected rewards are going to produce the largest positive reward-prediction errors that help solidify a behavior (involving a decision/movement pattern). Repetition of these positive signals helps consolidate a behavior, perhaps even an incidental one, from a transient experience into a stable memory. This is why gambling can be so addictive. The random rewards of gambling lead to large positive reward-prediction errors (Holroyd, Nieuwenhuis, Yeung, & Cohen, 2003). Second, these reward-prediction errors can assume three different "valences": we can have positive signals that help consolidate successful behaviors; we can have negative signals that help inhibit unsuccessful behaviors; *and* we can have neutral signals where predicted reward matches actual reward and therefore do not provide a strong impetus to change behavior. As alluded to in the prior paragraph, large reward prediction errors are common early in learning. These reward prediction errors tend to be positive (unexpectedly sinking a putt) and may explain the rapid improvements typically exhibited early in practice. Small prediction errors are common late in learning (when a performance is close to ones' expectations) and may explain the relatively slow improvements commonly demonstrated late in practice.

Much has been written about how reward-prediction errors work in the human brain (for detailed reviews, see Holroyd & Coles, 2002; Walsh & Anderson, 2012; Wise, 2004). Using electro-encephalography (EEG), we can measure voltages from the brain on the surface of the scalp. This powerful technique allows us to measure neural correlates of behavior. For instance, we can look at how the brain processes augmented feedback about successful versus unsuccessful trials of a golf putting task in a hypothetical experiment (Figure 3.1B). By time-locking the EEG signals to the onset of feedback (i.e., time = 0 is when feedback appears) we can see some characteristic peaks in the EEG waveform when we average all the successful (holed) trials together, average all the unsuccessful (missed) trials together, and then take the difference between the two (Luck, 2014). This positive going component in the difference wave (around 300 ms) is referred to as *reward positivity* (Holroyd, Pakzad-Vaezi, & Krigolson,

2008; Proudfit, 2015). The term “reward positivity” is new, and was previously known as “feedback-related negativity”. The change in terminology came about due to a change in how the difference wave was computed (now a positive value is observed based on subtraction of unsuccessful trials from successful trials; see Proudfit, 2015). Activity in this range will either tend to be positive (following successes) or negative (following errors). In most experiments, binary feedback is provided, but there are a few studies wherein graded or continuous feedback was given. In these studies, EEG activity may exhibit a negative deflection if the feedback is framed as error magnitude (how far *off* target the participant was), with greater errors eliciting larger negative deflections (Luft, Takase, & Bhattacharya, 2014). Conversely, EEG activity may exhibit a positive deflection if the feedback is framed as success magnitude (how far *on* target the participant was), with greater accuracy eliciting larger positive deflections (Frömer, Stürmer, & Sommer, 2016).

Positive rewards

The reward positivity is proposed to be an index of the reinforcement learning signal sent from the basal ganglia to the anterior cingulate cortex (Holroyd & Coles, 2002; Holroyd & Yeung, 2012). The basal ganglia uses the midbrain dopamine system to convey the reward-prediction error to the anterior cingulate cortex, where it is used to adaptively modify behavior. Thus, the reward positivity is thought to reflect the impact of a phasic dopaminergic signal conveying a reward-prediction error on the anterior cingulate cortex. This phasic increase in dopamine following successful trials is thought to be the mechanism behind consolidation of successful behaviors (Holroyd & Yeung, 2012), given dopamine’s role in augmenting neuroplasticity and learning (Wise, 2004).

This positive reward signal is modulated by feedback valence (positive/success versus negative/failure), reward magnitude (large rewards versus small rewards), and the probability/expectation of success, giving evidence that it reflects reward prediction error (Sambrook & Goslin, 2014). For example, unexpected positive feedback elicited a larger reward positivity than expected positive feedback in a time-estimation task (Holroyd & Krigolson, 2007). Amplitude of the reward positivity following successful feedback also correlated with the potential monetary reward (\$0–\$5) on that trial in a reaction time task (Meadows, Gable, Lohse, & Miller, 2016). In a computerised throwing task, where participants received feedback about the location of their throw relative to a target on a display monitor, reward positivity amplitude was largest for throws closest to the bull’s-eye (Frömer et al., 2016). This scaling was apparent, even though all throws were “on-target” in that feedback was only given for throws that hit the target.

Both reward-prediction error and the reward positivity are sensitive to the magnitude of a reward and the probability of that reward (see also Tobler, Fiorillo, & Shultz, 2014). The results of Meadows et al. (2016), based on monetary reward, and Frömer et al. (2016), based on accuracy, are consistent with

the idea that the reward positivity is sensitive to the magnitude of reward. Thus, actions with a low predicted probability of reward that lead to a large actual reward are going to create the largest reward-prediction error. This will lead to the largest reward positivity, the largest phasic release of dopamine, and thus theoretically lead to consolidation (i.e., long term retention) of that behavior.

One way to think of this process with respect to reward-prediction error is as an action “landscape,” as shown in Figure 3.2. Initially, this landscape is flat; all actions are equally preferable because the effect of choosing a specific action, a , is unknown for the given state, s . Actions that lead to rewards get elevated and therefore are more likely to be used again. Actions that lead to failures get depressed and thus are less likely to be used again. This model is mostly correct, but has a problem. The problem is that rewards and punishments are not two sides of the same coin. Non-human primate research shows dissociated effects of rewards and punishments on different brain structures. For instance, midbrain dopaminergic neurons only increase their activity for positive reward-prediction errors (Bayer & Glimcher, 2005), whereas negative reward-prediction errors might be reflected in serotonergic activity (Daw, Kakade, & Dayan, 2002). Human research also shows that punishments are not merely the opposite of rewards when it comes to learning (Frank, Seeberger, & Reilly, 2004).

Rewards versus punishments

We can see the difference between rewards and punishments in learning paradigms that use financial incentives. These extrinsic rewards are relatively easy to manipulate and allow researchers to create balanced experiments. For instance, in a sequential tracking task where participants learned to regulate their grip force to match a complicated pattern that unfolded over time, rewards and punishments were shown to affect learning differently (Abe et al., 2011). Three groups were compared: a neutral group (who were paid a flat rate for participation), a reward group (who received money proportional to their accuracy), and a punishment group (who were given money up front and then lost money proportional to their *in*accuracy). The exact payment rates were controlled so that all groups of participants left with approximately the same amount of money at the end of practice. Participants in all groups successfully adapted their behavior during practice, although the reward group was significantly more accurate than the neutral group, which was significantly more accurate than the punishment group. Nonetheless, the groups did not significantly differ on an immediate post-test. The critical finding, however, is that participants who practiced with financial rewards showed superior learning compared to the other two groups when tested at a 6-h, 24-h, and 30-day post-test. Notably, rewarded participants showed performance gains from the immediate post-test to 24-h and 30-day post-test (referred to as offline learning, see Figure 3.1A in Kantak & Winstein, 2012). Conversely, the punishment and neutral groups exhibited performance loss from immediate post-test to 6-h and 30-day post-test (as well as from 24-h to 30-day post-test). Thus, rewards, but not punishments, during practice led to better

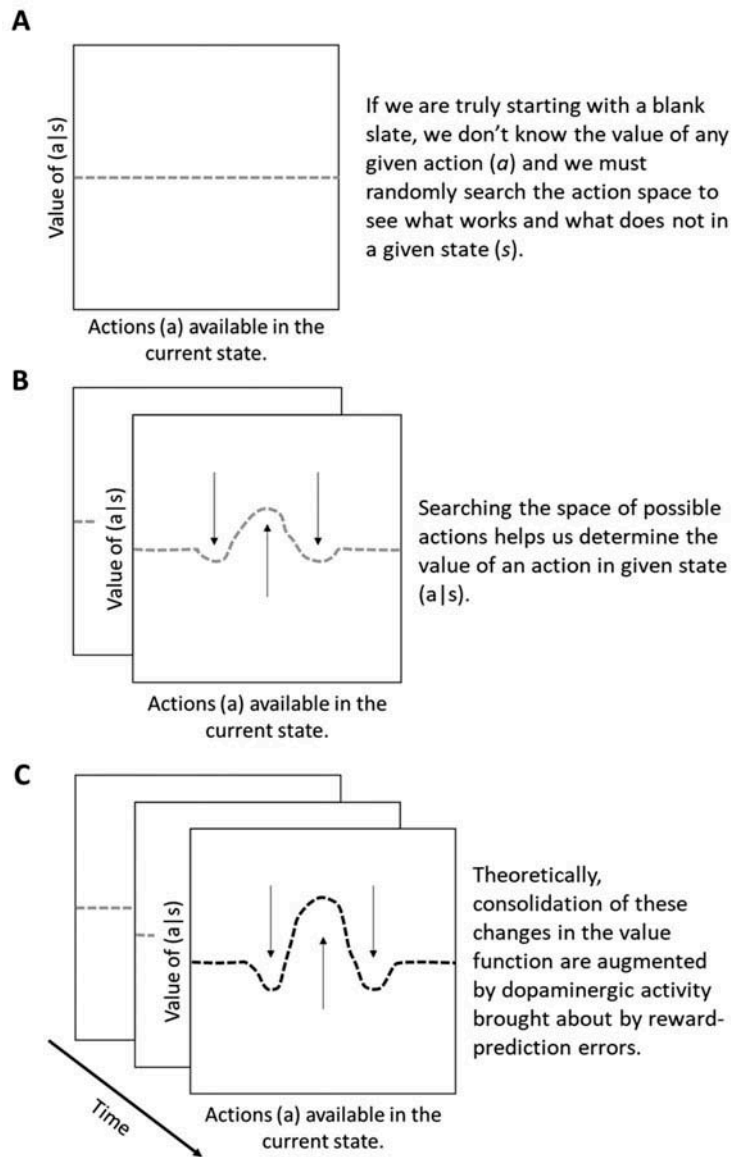


Figure 3.2 An illustration showing the hypothetical relationship (A) between value (expected total reward) of an action, a , in a given state, s . As the action-space is explored (B), the brain updates the value of taking different actions in that state. In order to select the appropriate action again in the future, these changes need to be consolidated (C).

consolidation of the motor skill. Moreover, there was no evidence that the punishment group showed worse learning than the neutral control group. Improved consolidation would explain why the skill was more resistant to decay and, indeed, exhibited offline gains, with the assumption being that dopaminergic signaling was responsible (although dopamine was not measured in this study).

How broad is the context of reward?

The experiment by Abe et al. (2011) is just one example of how rewards have been shown to affect skill acquisition, in terms of either adaptation or long-term learning (Cashaback, McGregor, Mohatarem, & Gribble, 2017; Izawa & Shadmehr, 2011; Nikooyan & Ahmed, 2015; Wachter, Lungu, Liu, Willingham, & Ashe, 2009). In Table 3.1, we present a summary of studies that have investigated the effects of financial rewards/punishments during practice in the retention of motor skills. Although the results are somewhat mixed, overall these data support the benefit of rewards for long-term retention (which is further supported by empirical work on reward in other domains and reinforcement learning theories; Schultz, 2015; Suri & Schultz, 1998; Thorndike, 1927). However, it is important to consider that “reward” is a relatively broad construct in reinforcement learning models (Sutton & Barto, 1998). Mathematically, reward-functions constrain the way an actor ought to behave in a given state, but these rewards are abstract. That is, rewards have specific numerical values, but these values are abstract representations. If we think about reinforcement learning at the level of a person, these could be rewards that satisfy basic needs (like food or water) or have symbolic value (e.g., financial rewards as in Abe et al., 2011). However, rewards can also be immaterial things like social status or the feeling of being successful. Introspectively, succeeding at a task feels rewarding, even if there is no tangible, external reward associated with the success. This feeling would suggest that for something to function like a reward, it just needs to have some *utility* (usefulness, value, or benefit) in the brain of the actor.

Debate over rewards and their functions is a very rich and detailed area of research (Niv, 2009; Schultz, 2015). In fact, it is difficult to nail down a precise definition of a reward, because rewards are usually defined by the behavioral and neural responses they elicit, and not by their physical characteristics. For our purposes, it will suffice to say that physical items that meet basic needs are rewarding and symbolic items (like money) are also rewarding. However, even abstract variables can be rewarding and formally modelled in reinforcement learning (Singh, Barto, & Chentanez, 2004). For instance, electrophysiological studies of the reward signals (reward positivity) have shown that it can be augmented by increasing a person’s perceived autonomy (Meng & Ma, 2015), or by the effort a person expends in order to achieve that reward (Ma, Meng, Wang, & Shen, 2013).

Finally, it is important to reiterate that the same reward can have different values for different people, or even the same person at different times (e.g. a classic example is that food rewards are less effective when rats enter the task

Table 3.1 Summary of reinforcement learning studies using long-term retention/transfer tests

| Study | Task | Manipulation | Training Session | Tests | Main Significant Results |
|---|---|---|-------------------------------|--|--|
| Abe et al., 2011 | Tracking isometric pinch force task along a repeated sequence | <p><i>Groups:</i></p> <p>Reward Group: earned money for time on target (starting point: \$0)</p> <p>Punishment Group: lost money for time off target (starting point: \$72)</p> <p>Neutral Group: Received \$40 at end of training unrelated to performance</p> | Four blocks of 20 trials/each | <p><i>Retention Tests:</i></p> <p>Immediate post-test</p> <p>6-h post-test</p> <p>24-h post-test</p> <p>30-day post-test</p> | <p>Reward group was better during training than neutral group, which was better than punishment group;</p> <p>Reward group was better at 6-h, 24-h, and 30-day post-test than other groups;</p> <p>Reward group exhibited performance improvement from immediate post-test to 24-h and 30-day post-test;</p> <p>Punishment and neutral group exhibited performance decrement from immediate post-test to both 6-h and 30-day post-test</p> |
| E. Dayan, Averbeck, Richmond, & Cohen, 2014 | Sequential pinch force task along repeated sequence | <p><i>Groups¹:</i></p> <p>Low Stochasticity Group: earned \$0.06/successful trial either 25% or 75% of the time²</p> <p>High Stochasticity Group: earned \$0.06/successful trial 50% of the time</p> | Five blocks of 20 trials/each | <p><i>Retention Tests:</i></p> <p>Immediate post-test</p> <p>24-h post-test</p> <p>7-day post-test</p> | <p>High stochasticity group was significantly better at immediate and 7-day post-test than the other two groups; High stochasticity group exhibited greater improvement from pre-test</p> |

| | | | |
|--|---|--|---|
| <p>Fixed Reward Group: earned \$0.06/successful trial 100% of the time</p> | | | <p>to immediate post-test than the other two groups. Further, the high stochasticity group exhibited greater improvement from pre-test to 7-day post-test than the other two groups</p> |
| <p>Hasson, Manczurovsky, & Yen, 2015 (Experiment 2)</p> | <p>New gait pattern</p> | <p><i>Groups</i>²: Supervised Group: received directional error feedback after each trial + fictitious money Reinforcement Group: received categorical feedback after each trial + fictitious money</p> | <p><i>Retention Tests</i>: Immediate post-test 24-h post-test <i>Transfer Tests</i>³: Immediate post-test 24-h post-test</p> <p>The reinforcement group exhibited a smaller performance decrease from immediate retention test to 24-h retention test than the supervised group</p> |
| <p>Steel, Silson, Stagg, & Baker, 2016</p> | <p>Serial Reaction Time Task (SRTT) and Force Tracking Task (FTT) with both repeated and random sequences</p> | <p><i>Groups</i>: Reward Group: earned \$0.05 each time they met the criterion Punishment Group: lost \$0.05 each time they did not meet the criterion</p> | <p><i>Retention Tests</i>: Immediate post-test 24-h post-test 30-day post-test</p> <p><i>SSRT</i>: Punishment and reward groups exhibited better sequence learning⁶ than control at early-training probe; control showed better sequence learning improvement from early-training to late-training</p> |

(Continued)

Table 3.1 (Cont.)

| Study | Task | Manipulation | Training Session | Tests | Main Significant Results |
|-------|------|---|--|-------|---|
| | | Control Group ⁴ : received feedback unrelated to their performance | <p><i>FTT</i>:</p> <p>Pre-training probe: three blocks of 8 trials/each</p> <p>Training probe: six blocks of 8 trials/each</p> <p>Post-training probe: six blocks of 8 trials/each</p> | | <p>probe than reward group; punishment group exhibited better sequence learning than control during training</p> <p><i>FTT</i>: Reward group demonstrated better improvement from early- to late-training probe than punishment group</p> |

*Notes:

- 1 Stochasticity refers to the level of uncertainty as to whether participants would be rewarded. Low stochasticity therefore means rewards were more predictable (either predictably low as in the 25% group or predictability high as in the 75% group) than in high stochasticity, where rewards occurred 50% of the time.
- 2 Directional error feedback consisted of guiding information toward the goal (e.g., desired ankle position and subject's ankle position), whereas categorical feedback had a non-directional nature as it informed the participants how far/close they were from the goal (e.g., very close to the goal). Fictitious monetary reward was added for motivational purposes. The amount of reward delivered was proportional to the errors, where small errors yielded high rewards, and rewards accumulated throughout the training session. No additional information was provided regarding the monetary reward.
- 3 The transfer test consisted of over-ground walking.
- 4 Control group was subdivided into two other groups: the reward-control group that received feedback at a frequency that matched the reward group, and the punishment-control group that received feedback at a frequency that matched the punishment group. (Note, then, that relative frequency of feedback increased during training for the reward group, but decreased during training for the punishment group. This follows because participants made fewer errors throughout practice.)
- 5 For both tasks, pre- and post-training probes were performed under random and fixed sequences such as the first and third blocks were random-sequence blocks and the second block was a fixed-sequence block (random-fixed-random). For the *FTT* task, one trial lasted for 14 seconds.
- 6 Sequence learning refers to difference in reaction time for repeated and random sequences for SSRT and difference in mean squared error for repeated and random sequences *FTT*. (For SSRT, accuracy data differed as a function of group in ways not entirely consistent with group differences in reaction time.)

having recently eaten; Tolman & Hoznik, 1930). This is important not only in experimental studies, but also in practice. The same system of rewards is not likely to have the same effect for everyone, nor the same effect over time.

Reinforcement from intrinsic rewards

Succeeding at a task is rewarding, but subjective success is more important than objective success, since feeling successful is what defines a rewarding experience in the brain of an actor. One way to make learners feel as though they are succeeding is by providing positive feedback. For example, one might ask whether an instructor should focus on providing feedback following good attempts and near misses, or following the most egregious failures. Scientists have shown that providing feedback about relatively well-performed attempts increases the learner's perceptions of competence relative to providing feedback about poorly-performed attempts (Chiviawosky, Wulf, & Lewthwaite, 2012; Saemi, Porter, Ghotbi-varzaneh, Zarghami, & Maleki, 2012). Hypothetically, increased perceptions of competence could result in more positive reward-prediction errors during practice, due to perceived competence's relationship to intrinsic motivation. Perceived competence is positively related to intrinsic motivation, and intrinsic motivation can be thought of as a multiplier for the reward-prediction error. That is, the default reward-prediction error is a relatively objective assessment by the motor system (e.g., how different was what we accomplished from what we wanted?). However, the reward-side of this calculation can be augmented by motivation (i.e., the more motivated I am, the more subjective reward I experience for a given objective success). This hypothesis is consistent with behavioral experiments where participants who receive feedback about their best practice trials exhibit superior learning relative to counterparts who receive feedback about poor trials, or a neutral mix of good and poor trials (Wulf & Lewthwaite, 2016). For example, Abbas and North (2017) had participants practice golf putting with vision of the putt's outcome occluded. Participants completed five blocks of six trials during practice. At the end of each block, participants received (positive) feedback about the outcome of their three most successful putts on the block, (neutral) feedback about the outcome of three random putts, or (negative) feedback about the outcome of their three least successful putts. The positive feedback group exhibited superior putting accuracy at 1-day and 7-day post-tests, indicating greater learning. These participants also generally reported higher levels of intrinsic motivation and perceived competence.

This hypothesis has been tested at a behavioral level by manipulating the psychological determinants of intrinsic motivation during motor skill learning. According to Self-Determination Theory (SDT), perceived competence in the task, autonomy, and/or relatedness are positively associated with intrinsic motivation (Ryan & Deci, 2000). If a learner achieves a successful outcome while experiencing competence, autonomy, and/or relatedness, then the learner could have a larger reward-prediction error and dopaminergic signal, thus promoting

superior learning relative to learners who are not experiencing those correlates of intrinsic motivation. A number of experiments have revealed that participants who practice under conditions designed to enhance perceived competence and/or autonomy exhibit superior motor skill learning (see Wulf & Lewthwaite, 2016). For example, participants who are given positive feedback not only experience the basic reward of success, but may be more intrinsically motivated toward the task, which effectively augments the reward and the resulting reward-prediction error.

Besides enhancing intrinsic motivation by increasing perceived competence, Self-Determination Theory posits that increasing feelings of autonomy and/or relatedness also enhance intrinsic motivation (Ryan & Deci, 2000). Autonomy and social-relatedness could, therefore, have a similar effect on reward-prediction errors to augment learning. Indeed, at the behavioral level, increasing autonomy or relatedness during practice has been shown to increase learning, but the underlying mechanisms for these effects are not clear. Autonomy support has been manipulated by allowing participants some degree of control over their practice environment. For example, Leiker et al. (2016) had participants in a self-control of practice group play a motion-controlled videogame while controlling the level of game difficulty. This group was compared to a yoked-group, which played the game with the level of difficulty matched to a self-controlled group counterpart (i.e., no control). The self-controlled group exhibited superior performance on 1-week delayed post-tests.

In a study designed to increase or reduce social-relatedness, a heightened sense of relatedness was similarly found to improve learning (assessed 1-day later) in comparison to a reduced sense of relatedness (Gonzalez & Chiviawosky, 2016). In this between-subjects study, participants practiced swimming at 50% maximal speed in one of three groups; (a) relatedness support, wherein the experimenter's instructions emphasised acknowledgement, caring, and interest in participants' experiences, (b) under conditions of relatedness reduction, wherein the instructions emphasised disinterest in the participants, or (c) under neutral conditions, wherein no attempt to manipulate relatedness was made. Thus, positive feedback, autonomy and relatedness support, all appear to be means to increasing learning, ostensibly by enhancing intrinsic motivation and, consequently, reward-prediction errors.

Theoretically, the benefits of these three determinants of intrinsic motivation are attributable (at least in part) to their effects on reward-processing and subsequent consolidation of successful behaviors. However, there is little direct evidence linking the promotion of intrinsic motivation to reward-prediction errors. Exceptions to this are from researchers who have been studying EEG, and reward positivity in particular, in relatively simple tasks without a delayed post-test. For example, Meng and Ma (2015) had participants perform a time estimation task in a within-subjects design involving two conditions. In the choice condition, participants were free to choose one of two time estimation tasks to perform for a trial. In the no choice condition, participants were randomly assigned a task for a trial. In both conditions, participants received augmented

feedback about success/failure on the task trial. Results showed an enhanced reward positivity for feedback about choice trials in comparison to no choice trials. This result suggests autonomy enhances reward-prediction errors, as indexed by reward positivity amplitude, presumably due to elevated intrinsic motivation. However, there is no single study showing that the effects of intrinsic motivation on reward-processing actually explain the effects of intrinsic motivation on learning.

Grand, Daou, Lohse, and Miller (2017) studied the relationship between intrinsic motivation, reward positivity magnitude, and motor learning, with a manipulation designed to increase autonomy. Participants in a task-irrelevant “choice” condition were allowed to choose the color of a beanbag used during practice of a throwing-task, whereas participants in a “yoked” condition received a color-schedule matched to a choice group counterpart. Feedback about the accuracy of throws was provided in practice and the reward positivity amplitude elicited by that feedback was recorded. The groups did not differ in reward positivity amplitude during practice, nor did they differ in throwing accuracy when tested in a 1-week post-test, see Figure 3.3. Additionally, reward positivity amplitude did not predict post-test performance. However, a composite of self-reported intrinsic motivation, internalised motivation, and effort predicted the magnitude of reward positivity, suggesting that increased motivation may indeed make successful outcomes more rewarding. Reward positivity amplitude predicted changes during the acquisition phase, such that participants exhibiting the largest amplitude showed the greatest improvement from the first block of practice trials to the last block. This result is consistent with other studies demonstrating a relationship between reward positivity amplitude and changes in performance within a practice session (e.g., Holroyd & Krigolson, 2007; Reinhart & Woodman, 2014; van der Helden, Boksem, & Blom, 2010). However, the task-irrelevant choice of bean-bag color did not significantly affect intrinsic motivation. Unfortunately, this lack of an effect makes it difficult to draw strong conclusions about the relations between these three variables (motivation, reward positivity, and motor learning).

In summation, we think that integrating the rich history of reinforcement learning with motor skill learning is a fruitful area for future research. Although reinforcement learning models have certainly been discussed in the context of motor skill learning (Doya, 2000; Seidler et al., 2013), relatively few researchers have investigated reward-processing effects on long-term learning at a mechanistic level. There has been a significant amount of research on reward-prediction errors and their neurophysiological basis, but almost all of those studies focused on adaptation within a session and not long-term learning (e.g., Cohen & Ranganath, 2007; Frank, Woroch, & Curran, 2005; Miltner, Braun, & Coles, 1997). It is important to dissociate adaptation from learning as we saw in the study by Grand et al. (2017), where reward positivity magnitude was related to the degree of adaptation, but not learning. Although reward-prediction errors and adjustments in behavior during practice are important for learning, there are significant

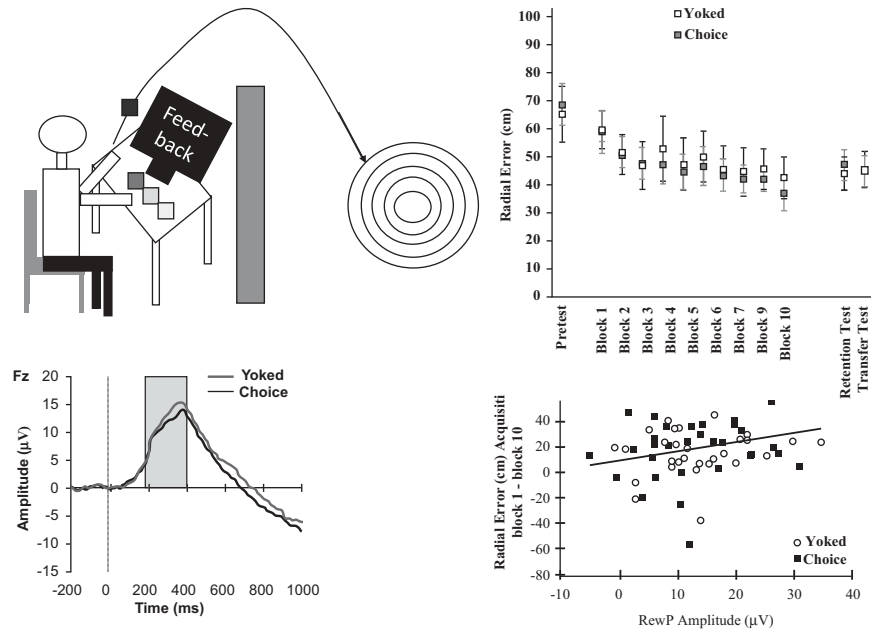


Figure 3.3 (Upper-left) Experiment set-up: Participants sat a table and completed 10 blocks of 10 non-dominant arm beanbag tosses to a target, which was occluded from vision by a board. Participants received augmented feedback (knowledge of results about toss accuracy) after each trial from a computer on the table. Participants in the choice group chose the color beanbag they tossed on each block, whereas participants in the yoked group were given a color schedule matched to a choice group counterpart. (Upper-right) Accuracy (radial error) as a function of group and experimental phase (Day 1: Pretest, Acquisition Blocks; Day 2: 7-day retention and transfer post-test). Both groups improved their performance throughout acquisition (Block 1 – Block 10), and both groups learned the skill, as evidenced by superior performance on post-test relative to pre-test. There were no group differences. (Lower-left) EEG-derived event-related potentials (ERPs) for the choice and yoked groups time-locked to the onset of augmented feedback, which occurred at 0 ms. The ERPs are averaged across all trials for each participant and then averaged across participants within groups. The time window in which the reward positivity was quantified is highlighted. Reward positivity amplitude did not differ between groups. (Lower-right) Reward positivity amplitude and acquisition improvement from Block 1–10 were significantly correlated (controlling for group), with higher reward positivity amplitude predicting greater acquisition improvement ($\beta_{\text{unstandardized}} = 0.720 \text{ cm}$, $p = .032$).

intervening psychological processes that make this relationship more complex (e.g., consolidation and retrieval; Kantak & Winstein, 2012). As such, we think there is considerable “low-hanging fruit” for researchers studying motor learning to

illuminate how reward-processing during practice affects the long-term retention of motor skills.

Recommendations for practitioners

We think that more research into reinforcement learning principles in motor skill acquisition will improve our ability to make specific and personalised recommendations for practitioners. As research in this area grows, we expect to be able to make detailed prescriptions based on interactions of the individual, task, and environment. In the interim, however, there are some general recommendations that can be made from burgeoning research in this area and decades of reinforcement learning studies in other domains.

Reinforcement learning describes the process by which we adjust our behavior based on prediction errors in order to maximise rewards and minimise punishments. In this chapter, we offer empirical evidence and strong theoretical arguments to support the idea that this learning model is probably the most important for understanding skill acquisition. Reinforcement learning principles can be used to explain performance adaptation and long-term retention of motor skills and should, therefore, inform how instructors and practitioners design practice conditions. Specifically, instructors should attempt to maximise reward-prediction errors by increasing the reward associated with successful outcomes.

- If reward-prediction errors facilitate learning, then these errors should be maximised. To that end, we do not want success (or failure) to be too predictable. Therefore, the difficulty of practice should be individualised. As success becomes too predictable, the difficulty of the task should increase (see also Guadagnoli & Lee, 2004).

Providing positive feedback after relatively good trials can increase a learner's perceptions of competence, thereby enhancing intrinsic motivation and learning (potentially via augmenting reward-prediction errors). Perceived competence is especially important for rehabilitation. Since low levels of motivation and engagement are common issues faced by therapists, they may want to implement strategies that make even the smallest improvement rewarding for their patients. This is an important consideration for coaches as well as therapists when progressively increasing difficulty. We encourage instructors and learners to embrace difficulty during practice, but emphasise the need to support and maintain learners' intrinsic motivation. When properly supported, difficult practice can create an engaging environment and produce prediction errors that provide the impetus for the nervous system to adapt.

- It is not only objective success in a task or skill that matters, but also subjective perceptions of success that matter for motor learning (e.g., Abbas & North, 2017). These successes are both psychologically and physiologically rewarding. Thus, instructors should provide augmented feedback to enhance learners'

56 *Keith Lohse et al.*

perceptions of success, as long as these are meaningful and based on the performer's expectations.

Perceived competence is only one pathway to increasing intrinsic motivation. Autonomy and relatedness are also positively related to intrinsic motivation and may be manipulated during practice sessions. Thus, instructors should provide learners with some control over their practice environment and show support/interest in learners' progress and experiences. Autonomy can be increased in a number of different ways. For instance, learners can be given a modicum of control over the order of practice, exercises during practice, when to receive feedback, or the progression of difficulty during practice.

- Giving learners self-control over practice not only increases autonomy, but it also gives learners a more personalised practice environment. Experimental research suggests that even novice learners are good at making decisions about difficulty progression that benefit learning (e.g., Ste-Marie, Carter, Law, Vertes, & Smith, 2017).

Summary

In psychology, a prediction error is the difference between an anticipated outcome and the actual outcome. Neurologically, these prediction errors are an important stimulus for learning and updating our behavior, the details of which have been developed through decades of research on reinforcement learning. Coaches and therapists can harness reinforcement learning principles to optimise practice conditions. For instance, if outcomes become too predictable, prediction errors will move toward zero reducing the impetus to update behavior. Thus in order to maintain learning, the difficulty of practice should progressively increase in relation to the learner's skill level. While increasing the level of difficulty, perceptions of competence, relatedness, and autonomy should be kept high to maintain intrinsic motivation in challenging situations. Ultimately, we think it is important that instructors and learners understand that errors (specifically prediction errors) should not be considered a threat to practice. On the contrary, they should be embraced and explored as they can lead to better motor skill acquisition and learning.

References

- Abbas, Z. A., & North, J. S. (2017, October 21). Good-vs. poor-trial feedback in motor learning: The role of self-efficacy and intrinsic motivation across levels of task difficulty. *Learning and Instruction*. <https://doi.org/10.1016/j.learninstruc.2017.09.009>
- Abe, M., Schambra, H., Wassermann, E. M., Luckenbaugh, D. A., Schweighofer, N., & Cohen, L. G. (2011). Reward improves long-term retention of a motor memory through induction of offline memory gains. *Current Biology*, *21*, 557–562.

- Adams, J. A. (1971). A closed-loop theory of motor learning. *Journal of Motor Behavior*, 3(2), 111–150. <https://doi.org/10.1080/00222895.1971.10734898>
- Bayer, H., & Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction signal. *Neuron*, 47(1), 129–141.
- Cashaback, J. G. A., McGregor, H. R., Mohatarem, A., & Gribble, P. L. (2017). Dissociating error-based and reinforcement-based loss functions during sensorimotor learning. *PLoS Computational Biology*, 13(7), e1005623. <https://doi.org/10.1371/journal.pcbi.1005623>
- Chiviawosky, S., Wulf, G., & Lewthwaite, R. (2012). Self-controlled learning: The importance of protecting perceptions of competence. *Frontiers in Psychology*, 3, 458. <https://doi.org/10.3389/fpsyg.2012.00458>
- Cohen, M. X., Elger, C. E., & Ranganath, C. (2007). Reward expectation modulates feedback-related negativity and EEG spectra. *Neuroimage*, 35(2), 968–978.
- Dam, G., Kording, K., & Wei, K. (2013). Credit assignment during movement reinforcement learning. *PLoS ONE*, 8(2), e55352. <https://doi.org/10.1371/journal.pone.0055352>
- Daw, N. D., Kakade, S., & Dayan, P. (2002). Opponent interactions between serotonin and dopamine. *Neural Networks*, 15(4–6), 603–616. [https://doi.org/10.1016/S0893-6080\(02\)00052-7](https://doi.org/10.1016/S0893-6080(02)00052-7)
- Dayan, E., Averbeck, B. B., Richmond, B. J., & Cohen, L. G. (2014). Stochastic reinforcement benefits skill acquisition. *Learning & Memory*, 21(3), 140–142. <https://doi.org/10.1101/lm.032417.113>
- Dayan, P. (2009). *Unsupervised Learning. The MIT Encyclopedia of the Cognitive Sciences.* <https://doi.org/10.1007/BF00993379>
- Doya, K. (2000). Complementary roles of basal ganglia and cerebellum in learning and motor control. *Current Opinion in Neurobiology*, 10(6), 732–739. [https://doi.org/10.1016/S0959-4388\(00\)00153-7](https://doi.org/10.1016/S0959-4388(00)00153-7)
- Frank, M. J., Seeberger, L. C., & Reilly, R. C. O. (2004). By Carrot or by stick: Cognitive reinforcement learning in Parkinsonism by Carrot or by stick: Cognitive reinforcement learning in Parkinsonism. *Science*, 306(5703), 1940–1943. <https://doi.org/10.1126/science.1102941>
- Frank, M. J., Woroch, B. S., & Curran, T. (2005). Error-related negativity predicts reinforcement learning and conflict biases. *Neuron*, 47(4), 495–501.
- Frömer, R., Stürmer, B., & Sommer, W. (2016). The better, the bigger: The effect of graded positive performance feedback on the reward positivity. *Biological Psychology*, 114, 61–68. <https://doi.org/10.1016/j.biopsycho.2015.12.011>
- Gonzalez, D. H., & Chiviawosky, S. (2018). Relatedness support enhances motor learning. *Psychological Research*, 82(3), 439–447.
- Grand, K. F., Daou, M., Lohse, K. R., & Miller, M. W. (2017). Investigating the mechanisms underlying the effects of an incidental choice on motor learning. *Journal of Motor Learning and Development*, 5(2), 207–226
- Guadagnoli, M. A., & Lee, T. D. (2004). Challenge point: a framework for conceptualizing the effects of various practice conditions in motor learning. *Journal of Motor Behavior*, 36(2), 212–224.
- Haith, A. M., & Krakauer, J. W. (2013). Model-based and model-free mechanisms of human motor learning. *Advances in Experimental Medicine and Biology*, 782, 1–21. https://doi.org/10.1007/978-1-4614-5465-6_1

- Hasson, C. J., Manczurovsky, J., & Yen, S.-C. (2015). A reinforcement learning approach to gait training improves retention. *Frontiers in Human Neuroscience*, 9, 459. <https://doi.org/10.3389/fnhum.2015.00459>
- Hebb, D. O. (1961). Distinctive features of learning in the higher animal. *Brain Mechanisms and Learning: A Symposium*, 37–46.
- Holroyd, C., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, 109(4), 679–709. <https://doi.org/10.1037//0033-295X.109.4.679>
- Holroyd, C., & Krigolson, O. E. (2007). Reward prediction error signals associated with a modified time estimation task. *Psychophysiology*, 44(6), 913–917. <https://doi.org/10.1111/j.1469-8986.2007.00561.x>
- Holroyd, C., Nieuwenhuis, S., Yeung, N., & Cohen, J. (2003). Errors in reward prediction are reflected in the event-related brain potential. *Cognitive Neuroscience and Neuropsychology*, 14(18), 4. <https://doi.org/10.1097/01.wnr.0000099601.41403.a5>
- Holroyd, C., Pakzad-Vaezi, K. L., & Krigolson, O. E. (2008). The feedback correct-related positivity: Sensitivity of the event-related brain potential to unexpected positive feedback. *Psychophysiology*, 45(5), 688–697. <https://doi.org/10.1111/j.1469-8986.2008.00668.x>
- Holroyd, C., & Yeung, N. (2012). Motivation of extended behaviors by anterior cingulate cortex. *Trends in Cognitive Sciences*, 16(2), 122–128. <https://doi.org/10.1016/j.tics.2011.12.008>
- Izawa, J., & Shadmehr, R. (2011). Learning from sensory and reward prediction errors during motor adaptation. *PLoS Computational Biology*, 7(3), e1002012. <https://doi.org/10.1371/journal.pcbi.1002012>
- Kantak, S. S., & Winstein, C. J. (2012). Learning–performance distinction and memory processes for motor skills: A focused review and perspective. *Behavioural Brain Research*, 228(1), 219–231. <https://doi.org/10.1016/j.bbr.2011.11.028>
- Leiker, A. M., Bruzi, A. T., Miller, M. W., Nelson, M., Wegman, R., & Lohse, K. R. (2016). The effects of autonomous difficulty selection on engagement, motivation, and learning in a motion-controlled video game task. *Human Movement Science*, 49, 326–335.
- Luck, S. J. (2014). *An Introduction to the Event-Related Potential Technique*. Cambridge, MA: MIT Press.
- Luft, C. D. B., Takase, E., & Bhattacharya, J. (2014). Processing graded feedback: electrophysiological correlates of learning from small and large errors. *Journal of Cognitive Neuroscience*, 26(5), 1180–1193.
- Ma, Q., Meng, L., Wang, L., & Shen, Q. (2013). I endeavor to make it: Effort increases valuation of subsequent monetary reward. *Behavioural Brain Research*, 261, 1–7. <https://doi.org/10.1016/j.bbr.2013.11.045>
- Meadows, C. C. C., Gable, P. A. P. A., Lohse, K. R. K. R., & Miller, M. W. M. W. (2016). The effects of reward magnitude on reward processing: An averaged and single trial event-related potential study. *Biological Psychology*, 118, 154–160. <https://doi.org/10.1016/j.biopsycho.2016.06.002>
- Meng, L., & Ma, Q. (2015). Live as we choose: The role of autonomy support in facilitating intrinsic motivation. *International Journal of Psychophysiology*, 98(3), 441–447. <https://doi.org/10.1016/j.ijpsycho.2015.08.009>
- Miltner, W. H., Braun, C. H., & Coles, M. G. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: evidence for a “generic” neural system for error detection. *Journal of Cognitive Neuroscience*, 9(6), 788–798.

- Nikooyan, A. A., & Ahmed, A. A. (2015). Reward feedback accelerates motor learning. *Journal of Neurophysiology*, *113*(2), 633–646. <https://doi.org/10.1152/jn.00032.2014>
- Niv, Y. (2009). Reinforcement learning in the brain. *Journal of Mathematical Psychology*, *53*(3), 139–154. <https://doi.org/10.1016/j.jmp.2008.12.005>
- Proudfit, G. H. (2015). The reward positivity: From basic research on reward to a biomarker for depression. *Psychophysiology*, *52*(4), 449–459. <https://doi.org/10.1111/psyp.12370>
- Reinhart, R. M., & Woodman, G. F. (2014). Causal control of medial–frontal cortex governs electrophysiological and behavioral indices of performance monitoring and learning. *Journal of Neuroscience*, *34*(12), 4214–4227.
- Rescorla, R., & Wagner, A. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical Conditioning II: Current Research and Theory*, *2*, 64–99.
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68–78.
- Saemi, E., Porter, J. M., Ghotbi-varzaneh, A., Zarghami, M., & Maleki, F. (2012). Knowledge of results after relatively good trials enhances self-efficacy and motor learning. *Psychology of Sport & Exercise*, *13*(4), 378–382. <https://doi.org/10.1016/j.psychsport.2011.12.008>
- Sambrook, T. D., & Goslin, J. (2014). Mediofrontal event-related potentials in response to positive, negative and unsigned prediction errors. *Neuropsychologia*, *61*(1), 1–10. <https://doi.org/10.1016/j.neuropsychologia.2014.06.004>
- Schmidt, R., & Lee, T. (2011). *Motor Control and Learning* (5th ed.). Champaign, IL: Human Kinetics.
- Schultz, W. (2015). Neuronal reward and decision signals: From theories to data. *Physiological Reviews*, *95*(3), 853–951. <https://doi.org/10.1152/physrev.00023.2014>
- Seidler, R. D., Kwak, Y., Fling, B. W., & Bernard, J. A. (2013). Neurocognitive mechanisms of error-based motor learning. *Advances in Experimental Medicine and Biology*, *782*, 39–60. https://doi.org/10.1007/978-1-4614-5465-6_3
- Singh, S. P., Barto, A. G., & Chentanez, N. (2004). Intrinsically motivated reinforcement learning. *Advances in Neural Information Processing Systems*, 1281–1288.
- Steel, A., Silson, E. H., Stagg, C. J., & Baker, C. I. (2016). The impact of reward and punishment on skill learning depends on task demands. *Scientific Reports*, *6*(1), 36056. <https://doi.org/10.1038/srep36056>
- Ste-Marie, D. M., Carter, M. J., Law, B., Vertes, K., & Smith, V. (2016). Self-controlled learning benefits: exploring contributions of self-efficacy and intrinsic motivation via path analysis. *Journal of Sports Sciences*, *34*(17), 1650–1656.
- Suri, R. E., & Schultz, W. (1998). Learning of sequential movements by neural network model with dopamine-like reinforcement signal. *Experimental Brain Research*, *121*(3), 350–354. <https://doi.org/10.1007/s002210050467>
- Sutton, R. S., & Barto, A. G. (1998). *Introduction to Reinforcement Learning*. Cambridge, MA: MIT Press. <https://doi.org/10.1.1.32.7692>
- Thorndike, E. L. (1927). The law of effect. *The American Journal of Psychology*, *39*, 212–222. <https://doi.org/10.2307/1415413>
- Tobler, P. N., Fiorillo, C. D., & Schultz, W. (2014). Adaptive coding of reward value by dopamine neurons. *Science*, *307*(5614), 1642–1645. <https://doi.org/10.1126/science.1105370>
- Tolman, E. C., & Hoznik, C. H. (1930). Degrees of hunger, reward and non-reward, and maze learning in rats. *University of California Publications in Psychology*, *4*, 241–256.

- van der Helden, J., Boksem, M. A., & Blom, J. H. (2010). The importance of failure: Feedback-related negativity predicts motor learning efficiency. *Cerebral Cortex*, *20*(7), 1596–1603.
- Wachter, T., Lungu, O. V., Liu, T., Willingham, D. T., & Ashe, J. (2009). Differential effect of reward and punishment on procedural learning. *Journal of Neuroscience*, *29*(2), 436–443. <https://doi.org/10.1523/JNEUROSCI.4132-08.2009>
- Walsh, M. M., & Anderson, J. R. (2012). Learning from experience: Event-related potential correlates of reward processing, neural adaptation, and behavioral choice. *Neuroscience & Biobehavioral Reviews*, *36*(8), 1870–1884. <https://doi.org/10.1016/j.neubiorev.2012.05.008>
- Wise, R. R. A. (2004). Dopamine, learning and motivation. *Nature Reviews. Neuroscience*, *5*(6), 483–494. <https://doi.org/10.1038/nrn1406>
- Wulf, G., & Lewthwaite, R. (2016). Optimizing performance through intrinsic motivation and attention for learning: The OPTIMAL theory of motor learning. *Psychonomic Bulletin & Review*, *23*(5), 1382–1414. <https://doi.org/10.3758/s13423-015-0999-9>
- Yarrow, K., Brown, P., & Krakauer, J. W. (2009). Inside the brain of an elite athlete: The neural processes that support high achievement in sports. *Nature Reviews Neuroscience*, *10*(8), 585. <https://doi.org/10.1038/nrn2672>