I made it! Effects of perceptions of success and enhanced expectancies on motor learning and its underlying mechanisms

by

Juliana Otoni Parma

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
August 5, 2023

Keywords: target size; motivation; EEG; reward positivity; alpha power; movement preparation

Approved by

Matthew W. Miller, Chair, Professor of Kinesiology
Keith R. Lohse, Associate Professor of Physical Therapy, Neurology
William (Hank) Murrah, Associate Professor of Educational Foundations, Leadership, and Technology
Melissa Pangelinan, Associate Professor of Kinesiology
Mary Rudisill, Professor of Kinesiology

Abstract

This dissertation describes a program of research encompassing three studies that focused on examining the effects of enhanced expectancies and perceptions of success on motor learning. OPTIMAL theory (Wulf & Lewthwaite, 2016) proposes that practice manipulations that enhance a learner's expectations for future successful outcomes lead to better motor performance and learning due to increased goal-action coupling and motor memory consolidation. These effects are expected to be achieved through an increase in motivation fostered by the fulfillment of a learner's basic psychological need to feel competent. Establishing an easy criterion of success during practice is a way to decrease a learner`s perception of task difficulty and results in an enhanced expectancy of performing well, which is expected to cause reward anticipation at the neural level. Importantly, a learner's expectations tend to be fulfilled, since performance outcomes will be interpreted as successful more frequently, which can affect the likelihood and the value of the achieved rewards, as well as the quantity of cognitive resources devoted to motor programming. Given that motor performance and learning influence and are influenced by feedback-related and motor-preparatory brain activity, we developed a research program combining a series of behavioral, psychophysiological, and meta-analytical studies to uncover how task manipulations that affect learners' expectancies and perceptions of success can affect motor skill acquisition and its underlying neural processes.

The first study (Chapter 1), published in the *International Review of Sport and Exercise Psychology* (Bacelar et al., 2022), used a meta-analytical method to estimate the average and individual effect size of six types of manipulations to enhance expectancies in motor learning research. Results showed that, on average, enhancing learners' expectancies has a significant effect on skill retention ($g = 0.54$ (95% *CI* [0.38, 0.69]) that is dependent on

the type of manipulation adopted. However, evidence of reporting bias and small-study effects in this literature suggest that these effects are likely overestimated. The second study (Chapter 2), published in the *Psychology of Sport and Exercise* (Parma et al., 2023), is a behavioral experiment that sought to investigate the effect of perceived task difficulty, a manipulation to enhance expectancies, on learning. Learners with the same goal were provided with different criteria of success during the practice of a motor skill. Results showed that, contrary to the predictions of OPTIMAL theory, perceived task difficulty has a trivial, if any, behavioral effect on skill retention, even though learners with an easier criterion of success develop higher self-efficacy, perceived competence, and, for those performing more practice trials, increased intrinsic motivation. Lastly, the third study (Chapter 3) investigated how perceived and objective success affect psychophysiological measures of feedback processing (i.e., reward positivity [RewP] amplitude) and movement preparation (i.e., motor upper-alpha power). Specifically, we recorded learners' electroencephalograms while they acquired a motor skill with an easy or difficult criterion of success. Mixed-effects regression models were used to uncover how perception of success, objective success (error magnitude), and practice trial number affect RewP amplitude and motor upper-alpha power on a trial-by-trial basis. Results show that both subjective (perception of success) and objective (error magnitude) reward have a significant effect on feedback processing, with a larger effect from the former. Additionally, the relationship between feedback processing and error magnitude seems to depend on a learner's assigned criterion of success. For motor-preparatory brain activity, the effects of subjective and objective success are dependent on experience with the task, and seem to affect motor programming more than motor execution. Also, assigning learners' criteria for success moderates this relationship.

Together, this sequence of studies indicates that, although the effects of practice manipulations of expectancies of success on performance and learning are negligible, if existent, a learner's perception of success affects underlying neurophysiological and psychological mechanisms of motor skill acquisition related to feedback processing, movement preparation, and motivation.

Acknowledgments

I would like to thank my dissertation committee, Dr. Matthew Miller, Dr. Keith Lohse, Dr. Melissa Pangelinan, Dr. William (Hank) Murrah, and Dr. Mary Rudisill. Honestly, I could not have asked for a better team. Thank you all for sharing your knowledge with kindness, for being approachable, for giving me your time, and for teaching me all the "dos" of how to be a great doctor, professor, and mentor. You have also taught me unforgettable lessons that go much beyond academic knowledge. Know that I appreciate you!

Dr. Miller, I could not refrain from taking some time to talk about you in specific. You have no idea about your importance to me and how transforming our time together was in my life. Thank you for being so welcoming and generous, and for being so patient with my process as a doctoral student. Having students with backgrounds so different from yours could not have been too easy, but you never let it show. Thanks for embracing diversity and by showing that in every single action, it has never gone unnoticed and unappreciated, I hope you know! Thank you for trusting me, and at the same time never making me fear making mistakes. Thanks for picking fights that were only supposed to be mine. I was SO lucky to have you so close as mentor and a model of an outstanding human being.

To my dearest PEPlab mates, thank you for holding my hand so many times, and for taking this journey with me. Our shared experiences and knowledge were deeply valuable, and I could not have gotten to this point so smoothly without you. To all my students and research assistants, thank you for your hard work, and for teaching me more than learning from me, you are appreciated! To my volunteers, thank you for your time and commitment. I hope this dissertation makes you proud of your service!

Mom, dad, and bro, thank you for giving me everything and a little more. Your love and support were always felt even from this the distance, and that was essential. Thanks for

all the formal and informal education you always provided, for constantly pointing me in the right direction, for making me aware of all the privileges that helped to bring me here, and for letting me experience the world with the certainty I could always come back. Amo vocês.

To Auburn University and the School of Kinesiology, faculty, staff, and students, thanks for welcoming me as part of you, and for providing the best environment and support one can hope for. I am going to miss you dearly.

To my former mentor, Guilherme Lage, and my good old NNeuroM and GEDAM members, I am only here because of the bases you gave me, because of the scientific discussions we had, and because you stayed with me, even after I left. I appreciate you!

To my old friends that were always present even from a distance, thank you! To my friends that became family in this place, our bond is too special to be explained. You took this challenge with me as if it was your own. You made Auburn feel like a cozy home, and I could never thank you enough for that.

To the younger Ju, for making good use of opportunities, embracing this journey, and managing to make it beautiful, this is for you, I made it!

To every single person that somehow made it so fricking hard for me to say goodbye, THANK YOU!

Noise is part of the system. The good news is we learn how to adjust the signal-to-noise ratio!

Table of Contents

List of Tables

List of Figures

Chapter 1: Meta-analyzing enhanced expectancies on motor learning: positive effects but methodological concerns

Motor skills are a crucial part of everyone's life. Being able to effectively perform a motor action is facilitated by a thorough understanding of how motor skills are acquired and, more importantly, retained over time. Past attempts to uncover the bases of motor learning and the mechanisms underlying a variety of practice conditions (e.g., random vs. blocked practice; infrequent vs. frequent augmented feedback) relied on a cognitive perspective mainly grounded on the role played by information processing (Guadagnoli & Lee, 2004; Lee et al., 1994). More recently, however, a growing body of studies have shown that attentional and motivational factors may also need to be considered when it comes to understanding and promoting motor learning (Lewthwaite & Wulf, 2010; Pascua et al., 2015; Sanli et al., 2013), which culminated in the proposition of a new theory entitled: 'Optimizing Performance Through Intrinsic Motivation and Attention for Learning (OPTIMAL) theory of motor learning' (Wulf & Lewthwaite, 2016).

According to this theory, learning is facilitated by practice conditions promoting enhanced expectancies, autonomy, and external focus of attention (i.e., focusing on the effects of one's movement). More specifically, practice conditions wherein one's expectancies for future positive outcomes are enhanced (e.g., Lewthwaite & Wulf, 2010), the feeling of autonomy is promoted (e.g., Sanli et al., 2013), and an external focus of attention is encouraged (e.g., Lohse et al., 2010) lead learners to focus on the task goal, which enhances motor performance and learning. Although each motivational and attentional factor outlined in the OPTIMAL theory has been shown to benefit performance and learning, here we focus on studies that investigated enhanced expectancies in a motor learning context.

Different manipulations have been used to enhance learners' expectancies for future success. One of the most studied approaches consists of providing learners with feedback after more accurate trials. This approach has been shown to be effective when contrasted with both neutral (Chiviacowsky et al., 2019) and negative feedback (Chiviacowsky & Wulf, 2007). In another frequently adopted paradigm, which might be considered a manifestation of feedback after good trials, learners are led to believe they are performing better than their peers via provision of positive (false) social-comparative feedback, typically in addition to veridical feedback (Ávila et al., 2012). Manipulations of perceived task difficulty have also been used to influence learners' expectations. For instance, studies have reduced perceptions of task difficulty (i.e., made the task look easier) by implementing optical illusions (Palmer et al., 2016) or changing task criterion of success (Chiviacowsky et al., 2012). Other ways to enhance expectancies include influencing one's conceptions of ability (i.e., making one believe successful performance is achievable with practice as opposed to being a fixed capacity; e.g., Harter et al., 2019), the use of self-modeling strategies (i.e., showing edited videos with learners' best trials, e.g. (Ste-Marie et al., 2011), and extrinsic rewards (e.g., provision of monetary compensation; (Abe et al., 2011)).

The goal of the present meta-analysis was to investigate the effect of enhancing learners' expectancies for future successful outcomes on motor learning. As a secondary goal, we aimed to estimate the effect of each of the aforementioned manipulations on motor learning. To our knowledge, this meta-analysis is the first quantitative synthesis of the growing body of studies indicating enhanced expectancies facilitate motor learning. Thus, this analysis should provide the best estimate of the effect of enhanced expectancies on motor learning to date. Additionally, we use funnel plot analysis to investigate the risk that inflated effects in small studies (small-study effects) are distorting the extant literature. Our results can inform future investigations, for example by revealing shortcomings in the present research (e.g., small sample sizes). Our findings may also guide motor skill instruction, for example by providing coaches and physical therapists with the state of evidence about recommendations that are easy to implement, such as reducing perceived task difficulty. Thus, our meta-analysis has implications for researchers and practitioners.

Methods

Prior to data collection, methods and main analyses were pre-registered and made available in the Open Science Framework (OSF) repository (https://osf.io/mbux2/?view_only=2dc9697af80342ebbaf4c86f562b8bdd). The PICO (Population, Intervention, Comparison, Outcome) model was used to define the meta-analysis objectives. The population of interest was human subjects of all ages. Studies investigating people with disabilities and/or impairments were not excluded from the meta-analysis. Interventions were those Wulf and Lewthwaite (2016) indicate have shown enhanced expectancies facilitate motor learning: feedback after good trials, comparative feedback, self-modeling, perceived task difficulty, extrinsic rewards, and conceptions of ability. The main comparison of interest was between enhanced expectancies and control/neutral groups. In the absence of a control/neutral group, a comparison between enhanced expectancies and diminished or negative expectancies groups (e.g., feedback after good trials vs. feedback after poor trials) was considered. The outcome of interest was objective behavioral performance on a delayed (≥ 24-hr) retention test, which is a common and recognized learning evaluation (Kantak & Winstein, 2012).

Study Eligibility Criteria

Studies published in English and Portuguese were considered eligible if they met the following inclusion criteria: (1) it had an experimental design; (2) it used a task requiring movement to accomplish a goal that is increasingly likely to be achieved with practice (Schmidt & Lee, 2020); (3) it included at least one delayed (≥ 24-hr) retention test; (4) it was published in a peer-reviewed journal; (5) it assessed an objective behavioral measure; and (6) it included at least a positive enhanced expectancies group and a control group or a diminished (negative) expectancies group. Studies were excluded if they failed to meet the inclusion criteria and/or had insufficient data (i.e., did not report mean, standard deviation, or number of participants per group).

Literature Search Strategy

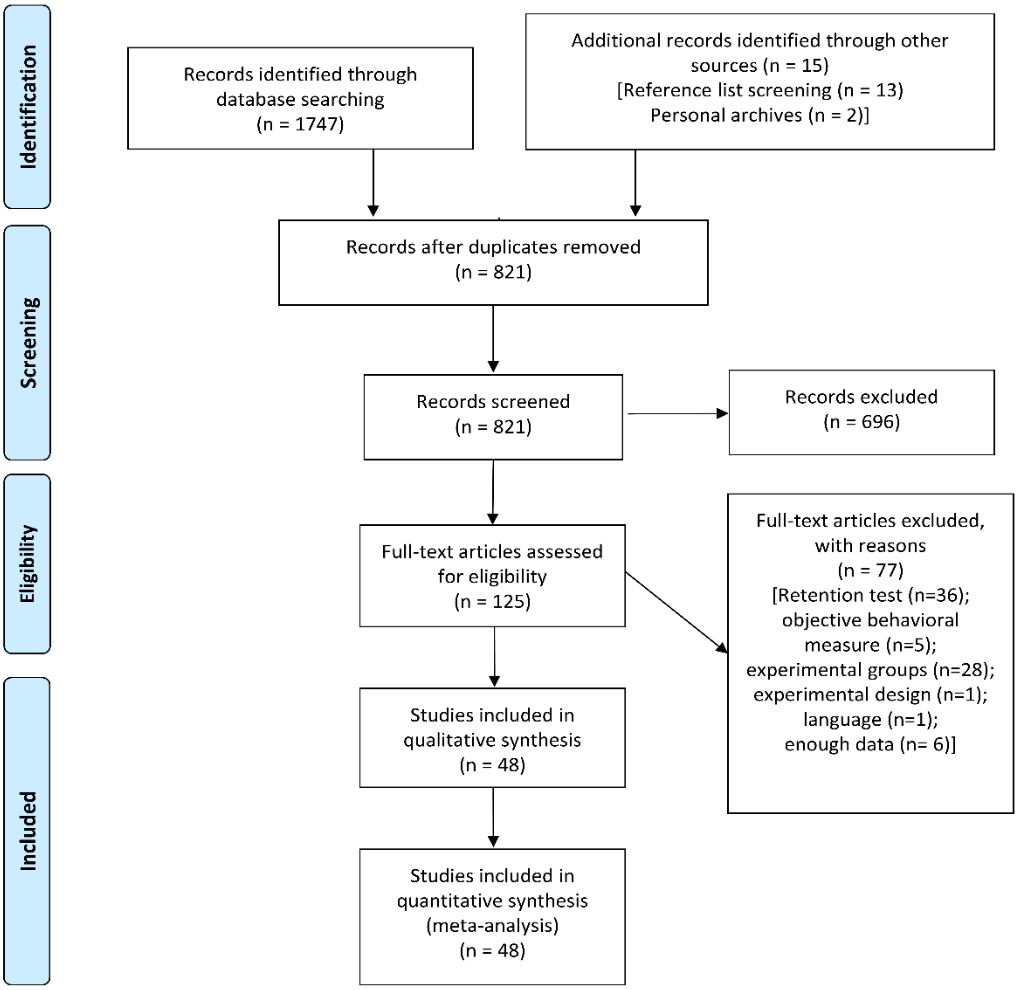
The electronic databases PsycINFO, Web of Science, and PubMed were searched from May 30, 2020, until June 19, 2020 (date of last search). Search terms included a combination of 'motor learning' or 'skill acquisition' and 'expectancies' or 'positive feedback' or 'good trial' or 'successful trial' or 'accurate trial' or 'normative feedback' or 'comparative feedback' or 'comparison feedback' or 'self-model' or 'self-as-a-model' or 'self-video' or 'video model' or 'video edit' or 'conceptions of ability' or 'ability conception' or 'inherent ability' or 'entity theory' or 'incremental theory' or 'learnable skill' or 'natural capacity' or 'acquirable skill' or 'task difficulty' or 'target size' or 'visual illusion' or 'hypnosis' or 'perceived difficulty' or 'mindset' or 'large target' or 'easy goal' or 'easy objective' or 'superstition' or 'reward' or 'incentive' or 'financial reward' or 'money'. String search was adjusted based on electronic database and intervention of interest. A detailed description of the search strategy, including limits used in each database, can be found in the OSF repository. These terms were chosen based on the terms and studies listed in the Enhanced Expectancies section of the OPTIMAL theory paper (Wulf & Lewthwaite, 2016). Further relevant papers were identified by searching through reference lists of previously selected papers and consulting personal archives. Publication period was unrestricted.

Study Selection and Data Extraction

A PRISMA flow chart with a detailed description of the study selection process can be found in Figure 1. The authors M.F.B.B and J.O.P. independently searched for studies in the databases. After removing duplicates, 821 papers were screened by title and abstract. Next, the remaining 125 papers were fully assessed for eligibility according to the inclusion criteria. When there was a disagreement regarding study eligibility, the matter was discussed with the fourth author (MWM) until agreement was reached. At the end of the study selection process, 48 studies met the inclusion criteria and were included in the meta-analysis.

Figure 1

PRISMA Flow Diagram



Note. Figure depicting the flow of information through the different steps of literature search and study selection (Moher et al., 2009).

Risk of Bias Assessment

The revised version of the Cochrane risk-of-bias tool (RoB 2) for randomized trials was used to assess the risk of bias in the studies included in the meta-analysis (Sterne et al., 2019). The tool is comprised of five bias domains, namely bias arising from the randomization process, bias due to deviations from intended interventions, bias due to missing outcome data, bias in measurement of the outcome, and bias in selection of the reported result. For the bias due to deviations from intended interventions domain, we focused on the effect of assignment to interventions (intention-to-treat effect). MFBB and JOP independently assessed the five bias domains and classified each one as low risk of bias, some concerns, or high risk of bias for each study following the proposed Cochrane algorithm. Next, an overall judgment of risk of bias was obtained for each study. Specifically, studies were classified as overall low risk of bias if they were judged to be at low risk across all individual bias domains; as "some concerns" if they raised some concerns in at least one domain but were not at high risk in any individual domain; and as overall high risk of bias if they raised some concerns in multiple bias domains or were judged to be at high risk in at least one domain. The *robvis* tool (McGuinness & Higgins, 2021) was used to plot the risk-of-bias results.

Data Extraction, Synthesis, and Analysis

The main variable of interest was performance on delayed retention test[1]. Retention test is here defined as the test performed at least 24-hr after the end of the acquisition phase, wherein all groups are tested under identical conditions and perform a task similar to the one performed during the acquisition phase (Schmidt et al., 2018). Only objective measures of performance were considered. When studies did not have a 24-hr retention test or contained

[1] Two reasons guided our decision to focus on performance on delayed retention test. First, there is no theoretical explanation as to why enhanced expectancies may affect retention and transfer test performance differently. Second, given the significant variability in types of transfers tests found in this literature, adding performance on delayed transfer test to our meta-analysis would likely introduce unnecessary heterogeneity to our data.

more than one retention test, the retention test closest to 24-hr was chosen to increase homogeneity among studies. For studies in which the 24-hr retention test was comprised of more than one block of trials, authors were contacted for data so an aggregate measure of retention test performance could be computed. In case of no response, we averaged across blocks (i.e., mean and standard deviation), which was the case for one study (Abbas & North, 2018). For studies that reported more than one behavioral measure, the measure more closely associated with accuracy (e.g., radial error as opposed to bivariate variable error [(Hancock et al., 1995)]) was chosen, since accuracy typically reflects the task objective (e.g., hitting a target). For studies in which the results of the retention test were presented as a set of individual trials as opposed to a single performance score, corresponding authors were contacted for data that would allow us to compute an aggregate measure of retention test performance. In cases where no response was obtained, we opted for the inclusion of the middle trial among a set of trials (e.g., the fourth of seven trials). The rationale behind the inclusion of the middle trial stems from the idea that this trial is less susceptible to warm-up and online learning effects, compared to the first and last trial, respectively. (Considering that averaging across trials was also an option, in the supplementary material we present the results of a sensitivity analysis using an average of retention trials.) Two authors (M.F.B.B. and J.O.P.) were responsible for extracting sample sizes, means, and standard deviations from the selected papers and entering the information into an Excel spreadsheet (Excel 2016, Microsoft). When sample sizes, means, and standard deviations were unavailable in tables or throughout the text, the R package *metaDigitise* (Pick et al., 2018) was used to extract raw data and summary statistics from figures. Corresponding authors were contacted when sufficient data and/or relevant information was not provided in the article. Only one effect size was extracted per study, except when the study had more than one experimental and control group (Ghorbani & Bund, 2020; Pascua et al., 2015; Wulf et al., 2014), was comprised of more than one experiment (Steel et al., 2016; Wulf et al., 2012), and/or assessed different populations (e.g., older vs younger adults; Drews et al., 2013; Grealy et al., 2019). In these cases, the number of effect sizes extracted exceeded the ratio one per study, but the assumption of independency among effect sizes was still met as the same experimental and/or control group was not used in multiple comparisons (Englund et al., 1999). In addition to statistical data, relevant information regarding population characteristics, study protocol, and experimental manipulation was also extracted. Table S1 provides information about experimental manipulation checks, which were conducted for 30 studies and at least somewhat successful in 22.

Hedges' *g* was chosen as the effect size metric since it considers the sample size of each study, being therefore considered an unbiased or corrected effect size (Lakens, 2013). Variables in which lower scores indicate better outcomes (e.g., radial error) were reversed in sign to ensure that effects favoring the experimental manipulation were positive (Harrer et al., 2019). Data were fitted into a random-effects model estimated using restricted maximum likelihood. Alpha level was set at .05 and effect size followed the standard guidelines (small = 0.2, medium = 0.5, large = 0.8) suggested by Cohen (1988). Heterogeneity was assessed using the Cochran's Q test. Since this test is influenced by sample size (Higgins et al., 2003), the I^2 statistic, quantified as the percentage of total heterogeneity over total variability, was also computed. The presence of small-study effects was assessed via visual inspection of the funnel plot along with Egger's regression test (Egger et al., 1997), which statistically assesses funnel plot asymmetry by predicting effect size from standard error. A trim-and-fill analysis was used to examine the sensitivity of the results to reporting bias (Duval & Tweedie, 2000). This technique iteratively trims studies from one side of the funnel plot until a criterion for symmetry is met, then fills the studies back into the plot while imputing ones that are identical except on the opposite side of the mean along the horizontal axis. The trim-and-fill analysis was carried out using the default algorithm provided by the *metafor* package (Viechtbauer, 2010) in R (cran.r-project.org) software. Since the trim-and-fill analysis assumes the decision to publish a scientific finding depends solely on the size of an effect, but reporting bias is likely more influenced by whether the effect is significant (Fanelli, 2012), we planned to *p*-curve the studies that had significant results (Simonsohn et al., 2014). However, we opted not to after determining that only 10 studies met the criteria to be included in a *p*-curve, due to the others not containing specific hypotheses, not reporting the types of post-hoc tests performed, reporting significant interactions, etc. Pre-specified moderator analyses were conducted to investigate how the type of manipulation moderated the estimated effect, and to investigate the effect of enhanced expectancies on motor learning when contrasted with different types of comparison groups (control or diminished expectancies). An exploratory moderator analysis was also conducted to investigate the effect of enhanced expectancies on learning in different populations (young adults, older adults, children/adolescents, and special populations). Visual inspection of funnel plots, studentized deleted residuals, and hat values were used to identify outliers and/or overly influential points in the dataset (Viechtbauer & Cheung, 2010). To ensure the robustness of the results, models were run with and without the studies identified as outliers and/or overly influential cases. The present meta-analysis was carried out using the *metafor* package (Viechtbauer, 2010) in R (cran.r-project.org) software. R code and dataset are available in the OSF repository.

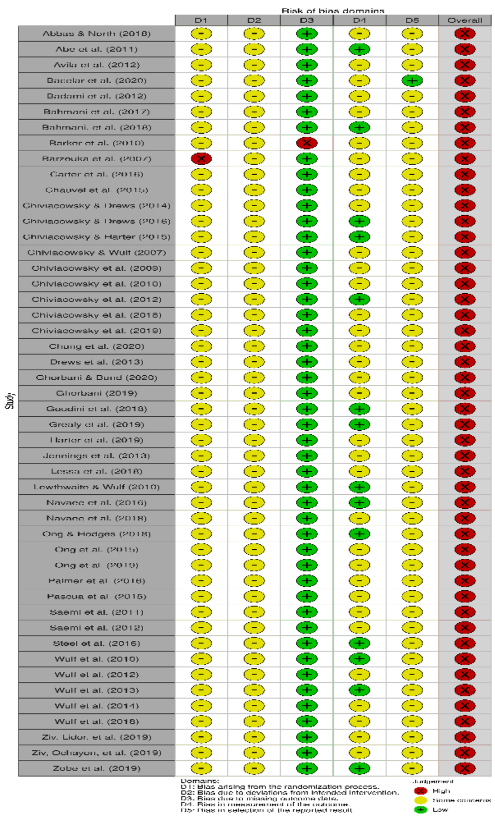
Results

Risk of Bias

Results of the risk of bias assessment are shown in Figure 2. All 48 studies included in the qualitative analysis were judged to be at high risk of bias. This was mainly due to some concerns being raised across all individual domains except for the bias due to missing outcome data domain. Specifically, some concerns were raised in the bias arising from the randomization process domain mostly due to studies not providing a detailed description of the randomization process; in the bias due to deviations from intended interventions domain due to experimenters responsible for delivering the intervention being likely aware of participants' group assignment; in the bias in measurement of the outcome domain due to the lack of information as to whether outcome assessors were aware of the intervention received by participants, which resulted in the assessment of outcome being possibility influenced by the assessors' knowledge of group assignment; and in bias in selection of the reported result domain due to the absence of pre-specified analysis plans. Except for one (Barker et al., 2010), studies were classified as being at low risk in the bias due to missing outcome data domain as there was no indication of missing data.

Figure 2

Risk of Bias Assessment Results



Note. Figure depicting all 48 studies included in the qualitative analysis and their respective risk of bias for each bias domain as well as overall risk of bias.

Descriptive Analysis

A summary of the main characteristics of the studies included in the meta-analysis can be found in Table 1. Forty-one studies contributed one data point each to the meta-analysis, whereas six contributed two data points each (Ghorbani & Bund, 2020; Grealy et al., 2019; Pascua et al., 2015; Steel et al., 2016; Wulf et al., 2012, 2014), and one study contributed three data points (Drews et al., 2013), resulting in a total of 56 effect sizes. The oldest studies included in the meta-analysis were published in 2007 (Barzouka et al., 2007; Chiviacowsky & Wulf, 2007), whereas the most recent ones were published in 2020 (Bacelar et al., 2020; Chung et al., 2020; Ghorbani & Bund, 2020), resulting in a publication period range of 14 years. The average study sample size was 14.85/group (median = 14/group), ranging from 8 to 28 participants per group.

Of the 56 effect sizes included in the meta-analysis, 16 represent manipulations of feedback after good trials, 13 represent manipulations of perceived task difficulty, 15 represent manipulations of comparative feedback, 7 represent manipulations of conceptions of ability, 4 represent manipulations of extrinsic rewards/punishments, and 2 represent manipulations of self-modeling[2]. The effect sizes composing this meta-analysis were extracted from data pertaining to young adults (n = 34), older adults (n = 6), children and adolescents (n = 13), and special populations (n = 3) consisting of adults with a disability in at least one upper or lower extremity (Bahmani et al., 2018), adults with Parkinson's disease (Chung et al., 2020), and autistic children (Navaee et al., 2018). Most of the effect sizes refer to a 24-hr retention test (n = 44), whereas the remaining refer to a retention test carried out

2. If summed, the number of manipulations exceeds the total number of effect sizes included in the meta-analysis. This is because one effect size reflects two manipulations combined (i.e., feedback after good trials and conceptions of ability; Wulf et al., 2013).

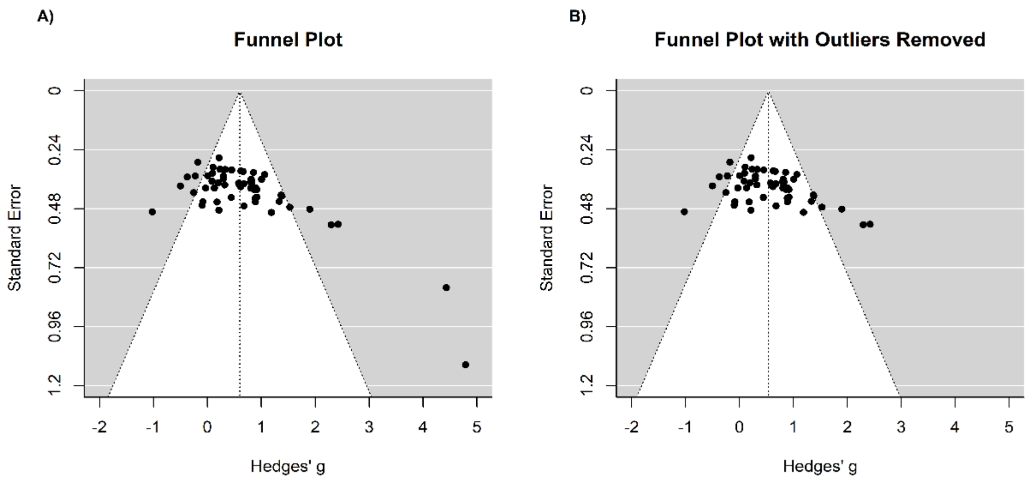
between 24-hr and one week after the acquisition phase (*n* = 7), or to a retention test carried out at least one week after the acquisition phase (*n* = 5).

Meta-analysis

Before running the random-effects model to estimate the effect of enhanced expectancies on motor learning, funnel plot visual inspection and influence diagnostics statistics were carried out to identify the presence of outliers and/or overly influential cases in the dataset. Figure 3A shows a funnel plot depicting all 56 effect sizes as a function of their standard error distribution. Visual inspection indicated the presence of two outliers (see bottom right of plot), which was confirmed by inspection of studentized deleted residuals and hat values, resulting in the removal of the studies by Goudini et al. (2018; rstudent = 4.19, hat = 0.009) and Navaee et al. (2016; rstudent = 3.44, hat = 0.005) from the subsequent analyses. (Results of the main meta-analysis with all 56 effect sizes can be found in the supplementary material.) Figure 3B shows the funnel plot after removal of outliers/influential cases.

Figure 3

Funnel Plot and Funnel Plot with Outliers Removed

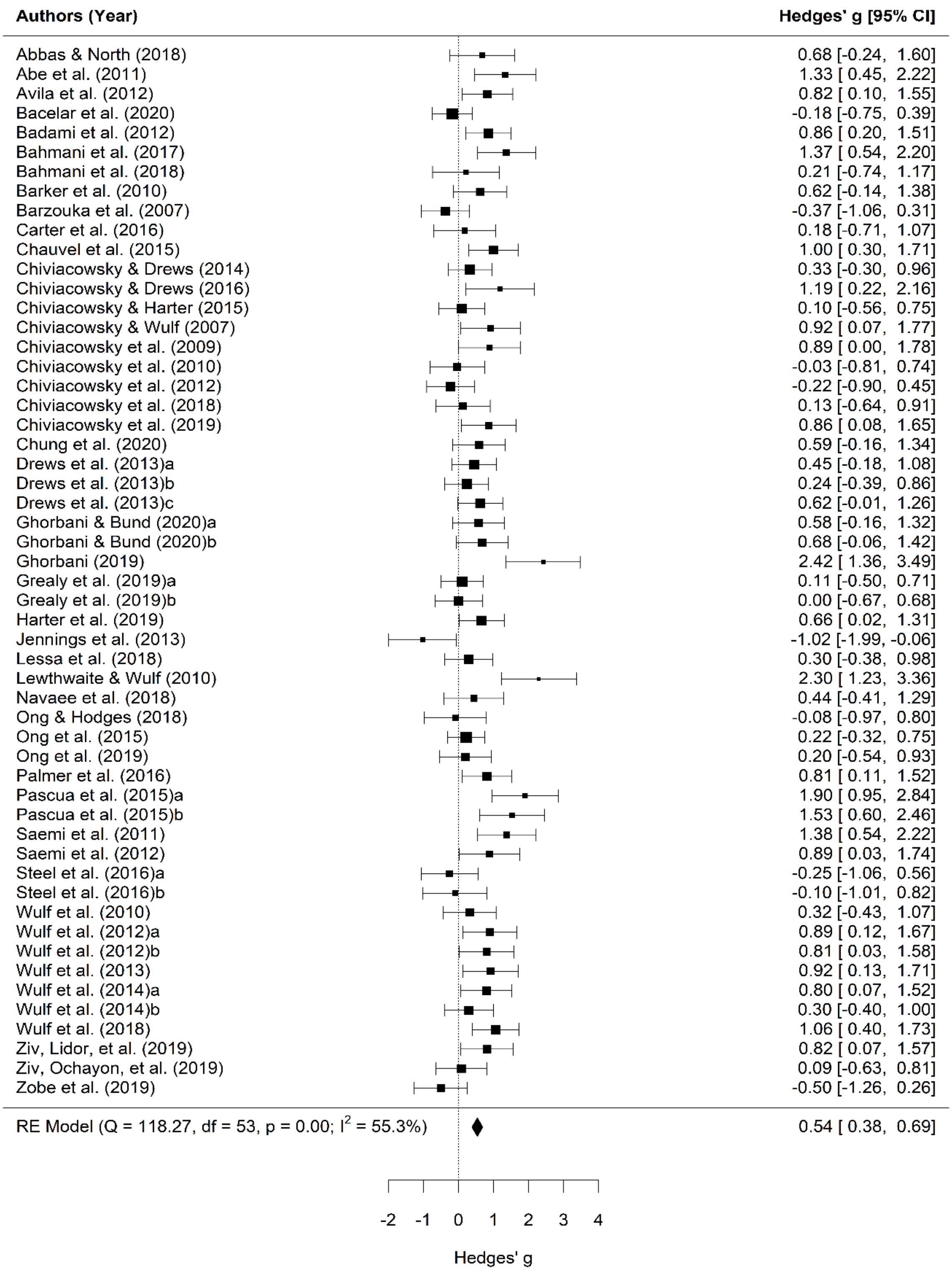


Note. A) Funnel plot depicting cases as a function of effect size and standard error of all 56 effect sizes. B) Funnel plot depicting cases as a function of effect size and standard error after outlier removal (*n* = 54).

Figure 4 depicts a forest plot with the 54 effect sizes included in the main analyses as well as a summary of the estimated effect. Results of the random-effects model revealed an overall effect size of medium magnitude (Hedges' $g = 0.54$, 95% CI [0.38, 0.69], $z = 6.85$, $p < .001$), indicating a positive effect of enhanced expectancies on motor skill learning. The Cochran's Q test was also significant ($Q(53) = 118.27$, $p < .001$), which suggests heterogeneity unlikely due to chance alone in the estimated effects across studies. This finding was corroborated by the results of the I^2 statistics, which revealed heterogeneity of $I^2 = 55.27\%$. Funnel plot visual inspection indicated asymmetry even after outlier/overly influential case removal, which was confirmed by the results of the Egger's regression test ($z = 3.49$, $p < .001$). Notably, asymmetry does not necessarily reflect small-study effects, but rather can occur by chance, sampling variation, and/or heterogeneity (Sterne et al., 2011). Since our funnel plot included 54 effect sizes, we reasoned chance and sampling variation were unlikely to have caused asymmetry. Thus, we were most concerned with exploring heterogeneity as an alternative to small-study effects as a cause of asymmetry, especially given the evidence of heterogeneity, possibly stemming from the use of studies implementing six different types of manipulations. If the asymmetry was mostly due to different types of manipulations having different effect sizes and standard errors, then funnel plots for each type of manipulation should be symmetrical. However, this does not seem to be the case, as described in the supplementary material, and depicted in Figure S2. Similarly, the asymmetry does not appear due to different populations (young adults, older adults, children/adolescents, and special populations) having different effect sizes and standard errors, as the funnel plots were not symmetrical for each population (see supplementary material and Figure S3). The trim-and-fill analysis failed to add studies to either side of the main funnel plot (Figure 3B).

Figure 4

Forest Plot



Note. Forest plot depicting all 54 effect sizes and their respective 95% confidence interval along with the overall Hedge's *g* effect size. Model summary is also presented on the bottom left side of the figure. Here, effect sizes favoring enhanced expectancies manipulations are presented on the right side of the zero Hedges' *g* line, whereas effect sizes not in favor of the manipulation in question are presented on the left side.

A moderator analysis was carried out to investigate the estimated effect size of enhanced expectancies on motor learning as a function of type of manipulation. (The same analysis with the 54 effect sizes included in the main analysis plus those considered outliers/overly influential can be found in the supplementary material.) Thus, type of manipulation (feedback after good trials, comparative feedback, self-modeling, perceived task difficulty, conceptions of ability, and extrinsic rewards) was entered into a mixed-effects model as a predictor. The moderator analysis revealed that at least one of the types of manipulations significantly differed from zero ($QM(6) = 66.15$, $p < .001$). The estimated effect of feedback after good trials was of large magnitude (Hedges' $g = 0.84$, 95% CI [0.54, 1.14], $z = 5.43$, $p < .001$, $n = 13$), indicating a beneficial effect of feedback after good trials on motor learning. In the same direction, a medium effect of comparative feedback (Hedges' $g = 0.61$, 95% CI [0.34, 0.88], $z = 4.39$, $p < .001$, $n = 15$) and a small effect of perceived task difficulty (Hedges' $g = 0.46$, 95% CI [0.18, 0.74], $z = 3.17$, $p = .002$, $n = 13$) and conceptions of ability (Hedges' $g = 0.39$, 95% CI [0.023, 0.76], $z = 2.083$, $p = .037$, $n = 7$ [3]) were found.

3. The study by Wulf et al. (2013) manipulated both conceptions of ability and feedback after good trials. Until this point, the effect size of this study reflected a combination of these two manipulations (acquirable-better group vs. inherent-worse group). However, for the purposes of this moderator analysis, we decided to categorize this study as 'conceptions of ability' by comparing the acquirable-worse group and inherent worse-group given that this manipulation had fewer cases ($n = 6$) than the feedback after good trials one ($n = 13$). (We chose to compare the acquirable- vs. inherent-*worse* groups because we reasoned the acquirable- and inherent-*better* groups may both have enhanced expectancies, with the latter believing they are naturally good at the task.) In the supplementary material, we present the results of a sensitivity analysis in which

Extrinsic rewards showed a trivial positive effect (Hedges' g = 0.15, 95% CI [-0.38, 0.68], z = 0.56, p = .577, n = 4), and self-modeling showed a moderate effect favoring the comparison group (Hedges' g = -0.64, 95% CI [-1.40, 0.12], z = -1.64, p = .101, n = 2), thus failing to provide evidence that these manipulations improve motor learning.

A second moderator analysis was conducted to identify the effects of enhanced expectancies as a function of the different types of comparison groups adopted (i.e., diminished expectancies group, n = 25; or control, n = 29). (The same analysis with the 54 effect sizes included in the main analysis plus those considered outliers/overly influential can be found in the supplementary material.) We did not find evidence that adding type of comparison to the model helped explain variability in effect sizes across studies ($QM(1)$ = 1.32, p = .251). Specifically, we observed a medium positive effect when comparing enhanced expectancies to diminished expectancies (Hedges' g = 0.63, 95% CI [0.41, 0.86], z = 5.53, p < .001) and a small effect when comparing enhanced expectancies to control (Hedges' g = 0.45, 95% CI [0.24, 0.66], z = 4.23, p < .001), but the effect of enhanced expectancies was not significantly influenced by comparison group type (β = -0.18, 95% CI [-0.49, 0.13], z = -1.15, p = .251).

Finally, to explore the effect of enhanced expectancies as a function of different populations, type of population (young adults, older adults, children/adolescents, and special populations) was entered into a mixed-effect model as the predictor. (The same analysis with the 54 effect sizes included in the main analysis plus those considered outliers/overly influential can be found in the supplementary material.) The exploratory moderator analysis revealed that at least one of the populations significantly differed from zero ($QM(4)$ = 45.29,

this study is classified as feedback after good trials (effect size reflecting the difference between the inherent-better and inherent-worse group).

$p < .001$). Specifically, a significant positive effect of medium magnitude was found for young adults (Hedges' $g = 0.61$, 95% CI [0.40, 0.81], $z = 5.70$, $p < .001$, $n = 32$), older adults (Hedges' $g = 0.48$, 95% CI [0.01, 0.96], $z = 1.99$, $p = .046$, $n = 6$), and children/adolescents (Hedges' $g = 0.44$, 95% CI [0.12, 0.75], $z = 2.71$, $p = .007$, $n = 13$), suggesting enhanced expectancies has a beneficial effect for these populations. Although a medium positive effect was observed for special populations, we did not find sufficient evidence that enhanced expectancies improve learning in this population (Hedges' $g = 0.43$, 95% CI [-0.27, 1.14], $z = 1.20$, $p = .231$, $n = 3$).

Discussion

The present meta-analysis estimated that enhancing learners' expectancies for future successful outcomes has a medium-sized benefit on motor learning ($g = 0.54$, 95% CI [0.38, 0.69]). Specifically, when analyzing different methods of enhancing expectancies, we found that manipulating feedback after good trials ($g = 0.84$, 95% CI [0.54, 1.14]) results in large benefits, while comparative feedback ($g = 0.61$, 95% CI [0.34, 0.88]) entails medium-sized benefits, and perceived task difficulty ($g = 0.46$, 95% CI [0.18, 0.74]) as well as conceptions of ability ($g = 0.39$, 95% CI [0.023, 0.76]) result in small benefits to learning. We did not find evidence that manipulating extrinsic rewards or self-modeling affect motor learning (ps $\geq$ .101), but few studies implemented these manipulations (ns $\leq$ 4), precluding reliable estimates of their effects. Thus, the effects of these manipulations should be estimated again when/if more studies in this line of investigation are conducted. (Since only 7 studies manipulated conceptions of ability and the effect of this manipulation has a wide CI that includes 0 when estimated among all 56 effect sizes (see Table S2), these results should be interpreted with caution.) Notably, enhanced expectancies benefitted motor learning similarly irrespective of whether the comparison group had diminished or neutral expectancies. This is consistent with Wulf and Lewthwaite (2016)'s suggestion that 'neutral' practice conditions are not really neutral, but rather likely elicit negative expectancies due to learners' concerns about having their performance assessed and compared with others'. Finally, we found that manipulating enhanced expectancies has a medium-sized positive effect on motor learning for young adults (g = 0.61, 95% CI [0.40, 0.81]), older adults (g = 0.48, 95% CI [0.01, 0.96]), and children/adolescents (g = 0.44, 95% CI [0.12, 0.75]). We did not find evidence to support the benefits of enhanced expectancies for special populations (p = .231), which in the present meta-analysis consist of adults with a disability in at least one upper or lower extremity (Bahmani et al., 2018), adults with Parkinson's disease (Chung et al., 2020), and children with autism (Navaee et al., 2018). However, only three studies examined these populations, preventing reliable estimates of effects in them. Future research should investigate the effect of enhanced expectancies on motor learning in these populations.

Results emphasize the role of enhanced expectancies in facilitating motor learning, so it is worth considering potential underlying mechanisms of this effect. Practice conditions that enhance expectations for successful outcomes are motivating, which increases dopamine release during motor skill practice, thereby facilitating the consolidation of motor memories (Wise, 2004). This is because successful outcomes are intrinsically rewarding, activating the dopaminergic reward system (Lutz et al., 2012), and humans are motivated to pursue rewards during motor skill practice (Moskowitz et al., 2020). Importantly, the mere expectation of dopamine release modulates the dopaminergic reward system (Schmidt et al., 2014), which is crucial for motivation (Wise, 2004).

The present meta-analysis also revealed evidence of small-study effects and underpowered studies, likely causing the effect of enhancing learners' expectancies on motor learning to be overestimated. Specifically, funnel plot visual inspection revealed asymmetry that was confirmed by a significant relationship between study effect size and standard error

(Egger's regression test). We believe the asymmetry is unlikely caused by chance or sampling variation, since 54 effect sizes were used in the funnel plot. We explored the probability that asymmetry was due to different manipulations (feedback after good trials, comparative feedback, etc...) or different populations (young adults, older adults, etc…) having different effect sizes and standard errors by constructing funnel plots for each manipulation and population. We did not observe symmetry in each manipulation and population's funnel plot (Figures S2 and S3), making it unlikely that heterogeneity between manipulations or populations explains the asymmetry in the funnel plot with all manipulations and populations (Sterne et al., 2011). Evidence that small-study effects contribute to funnel plot asymmetry can be observed in the lack of relatively imprecise studies showing negative effects (Figure 3B). In particular, asymmetry may be due to inflated effect sizes in small studies, since the median sample size was n = 14/group, and such small studies are likely to have exaggerated effect sizes (Sterne et al., 2011). Notably, the combination of small samples and small-study effects may cause effect sizes to be severely overestimated in the extant literature. This follows because small studies are likely to be underpowered such that only those drastically overestimating an effect will be statistically significant and, consequently, published (Lohse et al., 2016). However, it is important to note that the present meta-analysis did not assess the gray literature, and, therefore, does not present direct evidence of reporting bias.

The risk of bias assessment also raises the possibility that the effect of enhancing learners' expectancies on motor learning could be misrepresented. Some concerns, such as those about the randomization process, may be due to authors not reporting procedures rather than not undertaking them (The Cochrane Collaboration, 2013), and other concerns are inherent to motor learning research, such as participant awareness of group assignment. However, certain concerns can be mitigated, such as those regarding bias in selection of the reported result. Thus, to estimate the effect of enhancing learners' expectancies on motor learning more accurately, we recommend researchers conduct pre-registered studies and registered reports with a priori sample size calculations (Caldwell et al., 2020; Lohse et al., 2016). Specifically, pre-registered studies and registered reports may reduce reporting bias by committing researchers to reporting specific analyses and outcomes and journal editors to publishing studies irrespective of their results. Researchers conducting a priori sample size calculations should consider this meta-analysis' effect sizes to be overestimated and are encouraged to power their studies to detect effects close to the lower bound of the 95% CI. According to G*Power 3.1.9.4 (Faul et al., 2007), a two-tailed independent sample *t*-test with $\alpha = .05$, $\beta = .20$, equal *n*/group, and a Cohen's $d = 0.38$ (lower bound of 95% CI) requires $n = 110$/group. This number is reduced to $n = 55$/group if a Cohen's *d* is set to 0.54, consistent with the effect size (likely overestimated) in the present study. Since these sample sizes will be large increases for most researchers, they are encouraged to consider ways to make their data collections more efficient, for example by using sequential analyses (Lakens, 2014).

The present results suggesting enhanced expectancies may facilitate motor learning, with the effect possibly overestimated due to small-study effects and small sample sizes, are somewhat like other recent meta-analyses of effects predicted by OPTIMAL theory. Jimenez-Diaz et al. (2020) investigated the effect of learner control of augmented feedback during acquisition, which may promote autonomy, on motor performance and learning. The authors reported learner-controlled feedback groups exhibited superior acquisition performance relative to experimenter-regulated feedback groups, and learner-controlled feedback groups demonstrated performance stability from acquisition to retention, whereas experimenter-regulated feedback groups showed performance decrement from acquisition to retention. However, learner-controlled feedback groups did not significantly differ in performance or learning in comparison to yoked feedback groups, which consisted of participants who received augmented feedback schedules matched to a counterpart in a learner-controlled group. Thus, results provide little support for the OPTIMAL theory prediction that promoting autonomy, via giving learners control of their augmented feedback, enhances motor performance or learning. Notably, the authors reported a small median sample size of approximately *n* = 12/group as well as funnel plot asymmetry and significant Egger's regression tests for both acquisition and retention data, indicating the possibility of small-study effects. Kim et al. (2017) examined the effect of external focus (on the effects of one's movement) vs. internal focus (on one's body movements) instructions on balance performance and learning. Consistent with OPTIMAL theory (Wulf & Lewthwaite, 2016), the authors reported external focus of attention groups exhibited superior balance during acquisition, retention, and transfer relative to internal focus of attention groups. The authors reported a small median sample size of approximately *n* = 14/group as well as funnel plot asymmetry and a significant Egger's regression test in the acquisition data but not in the retention data, indicating the possibility of bias in the former. (Funnel plot asymmetry was not assessed for transfer data.) Makaruk et al. (2020) investigated the effect of external vs. internal vs. control (no) attentional focus instructions on jumping performance but not learning. Consistent with OPTIMAL theory, the authors reported external focus of attention was superior to internal focus of attention and control conditions. The authors reported a median sample size of approximately *n* = 24 (14 of 15 studies were within-subjects), which is larger than the other meta-analyses. Unlike the other meta-analyses, the authors did not assess bias. Taken together, these meta-analyses and the present one suggest that the many individual studies reporting effects consistent with OPTIMAL theory (Wulf & Lewthwaite, 2016) exaggerate the supporting evidence, due to small-study effects and underpowered studies, which is common in motor learning (Lohse et al., 2016) and other fields (e.g., (Button et al., 2013). Aggregating individual studies to estimate effects more accurately with meta-analyses is an important endeavor, but the presence of bias and an environment conducive to questionable research practices (e.g., conducting many statistical tests) in motor learning (Lohse et al., 2016) makes it difficult for OPTIMAL theory-based or other motor learning meta-analyses to establish whether even medium-sized effects, such as the one observed in the present study, are truly different from zero (Carter et al., 2019).

An important question for future research is to what degree practitioners typically implement strategies that enhance expectancies in comparison to those that are neutral or diminish expectancies. If coaches/clinicians rarely create neutral practice conditions or those that diminish expectancies, then their adoption of strategies to enhance expectancies will have little added value. Notably, researchers have investigated whether coaches use external focus of attention instructions, as recommended by OPTIMAL theory, and revealed that they usually do not (Diekfuss & Raisbeck, 2016; Porter et al., 2010; Yamada et al., 2020). Thus, it is conceivable practitioners also fail to create practice conditions that enhance expectancies.

The present meta-analysis suggests that enhancing learners' expectancies for future successful outcomes may facilitate motor learning across young adults, older adults, and children/adolescents. The meta-analysis lacked studies manipulating extrinsic rewards and self-modeling (and, to a lesser degree, conceptions of ability), and studies investigating the effect in question in special populations, so these effects should be estimated again when/if more studies in this line of investigation are conducted. As the meta-analysis indicated small-study effects and small sample sizes, pre-registered analyses and/or registered reports with greater statistical power are recommended. This final recommendation is critical to develop a body of studies conducive to accurately estimating the effect of enhanced expectancies on motor learning as well as other effects predicted by OPTIMAL theory and other motor learning theories.

Table 1

Summary of the main characteristics of the studies included in the meta-analysis.

Study	*Manipulation*	*Population*	*Sample Size*	*Comparison Group*	*Type of Task*	*Training Session*	*Retention*	*Chosen Performance Outcome*
Abbas & North (2018)	Feedback after good trials	Adults (Age: *M* = 29.67 years, *SD* = 9.36; 14 females)	KR-good: *n* = 10 KR-poor: *n* = 10 KR-neutral: *n* = 10 (Total: *N* = 30)	KR-good vs KR-neutral (Control)	Golf-putting	5 blocks of 6 trials at 2 meters 5 blocks of 6 trials at 5 meters	24-hr (1 block of 10 trials at 2 m and 1 block of 10 trials at 5 m) 1-week (1 block of 10 trials at 2 m and 1 block of 10 trials at 5 m)	Radial error
Abe et al. (2011)	Extrinsic rewards	Adults (Age: *M* = 24.3 years, *SD* = 5.2; 18 females)	Rewarded training: *n* = 13 Punished training: *n* = 12 Control training: *n* = 13 (Total: *N* = 38)	Rewarded training vs Control training (Control)	Tracking pinch force	4 blocks of 10 trials	24h and 30 days (1 block 20 trials)	Error (distance)
Ávila et al. (2012)	Comparative feedback	Children (Age: *M* = 10.4 years, *SD* = 0.36; 12 females)	Positive feedback: *n* = 16 Control: *n* = 16 (Total: *N* = 32)	Positive feedback vs Control (Control)	Non-dominant arm beanbag throwing	6 blocks of 10 trials	24-hr (1 block of 10 trials)	Accuracy score
Bacelar et al. (2020) - Main exp.	Extrinsic rewards	Adults (Age: *M* = 20.7 years, *SD* = 2.63; 55 females)	Reward: *n* = 25 Punishment: *n* = 22 Neutral: *n* = 22 (Total: *N* = 69)	Reward vs Neutral (Control)	Golf-putting	6 blocks of 8 trials	24h and 1 week (1 block of 8 trials)	Radial error

Study	*Manipulation*	*Population*	*Sample Size*	*Comparison Group*	*Type of Task*	*Training Session*	*Retention*	*Chosen Performance Outcome*
Badami et al. (2012)	Feedback after good trials	Adults (Age: *M* = 19.5 years, *SD* = 1.9; all females)	More Accurate: *n* = 20 Less Accurate: *n* = 20 (Total: *N* = 40)	More accurate vs Less accurate (diminished expectancies)	Golf-putting	10 blocks of 6 trials	24 h (1 block of 10 trials)	Putting accuracy scores
Bahmani et al. (2017)	Perceived task difficulty	Children (Age: *M* = 10.66 years, *SD* = 0.41; all males)	Perceived large hole: *n* = 15 Perceived small hole: *n* = 15 (Total: *N* = 30)	Perceived large hole vs Perceived small hole (Diminished expectancies)	Golf-putting	5 blocks of 10 trials	48-hr (1 block of 10 trials)	Deviation
Bahmani et al. (2018)	Perceived task difficulty	Adults with disability in ≥1 upper or lower extremity (Age: *M* = 37.7 years, *SD* = 9.8; 11 females)	Large illusion: *n* = 9 Small illusion: *n* = 8 (Total: *N* = 17)	Large illusion vs Small illusion (Diminished expectancies)	Aiming task (shooting - 10-m air pistol and air rifle)	5 blocks of 10 trials	24-hr (1 block of 10 trials)	Shooting accuracy
Barker et al. (2010)	Perceived task difficulty	Adults (Age: *M* = 21.50 years, *SD* = 3.25; 4 females)	Hypnosis: *n* = 14 Video attention control: *n* = 14 (Total: *N* = 28)	Hypnosis vs Video attention control (Control)	Soccer Wall-Volley	3 sessions each comprising soccer practice (3 trials), manipulation (45 min), and soccer practice (3 trials)	4 weeks (1 block of 3 trials)	Performance score
Barzouka et al. (2007)	Self-modeling	Adolescents (Age: *M* = 13.1 years, *SD* = 0.9; all females)	Other-modeling: *n* = 18 Self-modeling: *n* = 16 Control: *n* = 19 (Total: *N* = 53)	Group 2 vs Group 1 (Control)	Volleyball reception	12 practice sessions at a frequency of 2x/week; four kinds of drills with 10 repetitions each	1-week (1 block of 10 trials)	Performance outcome (score)

Study	*Manipulation*	*Population*	*Sample Size*	*Comparison Group*	*Type of Task*	*Training Session*	*Retention*	*Chosen Performance Outcome*
Carter et al. (2016)	Feedback after good trials	Adults (Age: *M* = 22.72 years, *SD* = 1.65; 22 females)	KR-good-aware: *n* = 10 KR-good-unaware: *n* = 10 KR-poor-aware: *n* = 10 KR-poor-unaware: *n* = 10 (Total: *N* = 40)	KR-good-unaware vs KR-poor-unaware (Diminished expectancies)	Mini Koosh-ball tossing	10 blocks of 6 trials	24h (2 blocks of 6 trials)	Radial error
Chauvel et al. (2015)	Perceived task difficulty	Adults (Age: *M* = 21.7 years, *SD* = 1.24; 20 females)	Perceived large hole: *n* = 18 Perceived small hole: *n* = 18 (Total: *N* = 36)	Perceived large hole vs Perceived small hole (Diminished expectancies)	Golf-putting	5 blocks of 10 trials	24-hr (1 block of 10 trials)	Deviation
Chiviacowsky & Drews (2014) – Exp. 2	Conceptions of ability	Children (Age: *M* = 10.5 years, *SD* = 0.51; 20 females)	Generic feedback: *n* = 20 Non-generic feedback: *n* = 20 (Total: *N* = 40)	Generic feedback vs Non-generic feedback (Diminished expectancies)	Non-dominant arm beanbag throwing	4 blocks of 10 trials	Retention 1: 24-hr (1 block of 10 trials) [4] Retention 2: 24-hr (1 block of 10 trials)	Accuracy score
Chiviacowsky & Drews (2016)	Comparative feedback	Adults (Age: *M* = 21.6 years, *SD* = 1.98; 4 females)	Positive self-comparison feedback: *n* = 10 Negative self-comparison feedback: *n* = 10 (Total: *N* = 20)	Positive self-comparison feedback vs Negative self-comparison feedback (Diminished expectancies)	Anticipatory coincident timing	4 blocks of 10 trials	24h (1 block of 10 trials)	Absolute error

[4] For the purposes of the present meta-analysis only Retention 1 was used.

Study	*Manipulation*	*Population*	*Sample Size*	*Comparison Group*	*Type of Task*	*Training Session*	*Retention*	*Chosen Performance Outcome*
Chiviacowsky & Harter (2015)	Perceived task difficulty	Adults (Age: *M* = 24.4 years, *SD* = 6.73; 24 females)	High experience of success: *n* =18 Low experience of success: *n* = 18 Control: *n* = 18 (Total: *N* = 54)	High experience of success vs Control (Control)	Anticipatory coincident timing	6 blocks of 5 trials	24-hr (1 block of 10 trials)	Absolute error
Chiviacowsky & Wulf (2007)	Feedback after good trials	Adults (Age: *M* = 21.1 years, *SD* = NA; 18 females)	KR good: *n* = 12 KR poor: *n* = 12 (Total: *N* = 24)	KR good vs KR poor (Diminished expectancies)	Non-dominant arm beanbag tossing	10 blocks of 6 trials	24 h (1 block of 10 trials)	Accuracy score
Chiviacowsky et al. (2009)	Feedback after good trials	Older adults (Age: *M* = 65.9 years, *SD* = NA; all females)	KR-good: *n* = 11 KR-poor: *n* = 11 (Total: *N* = 22)	KR-good vs KR-poor (Diminished expectancies)	Non-dominant arm beanbag tossing	10 blocks of 6 trials	72-hr (1 block of 10 trials)	Accuracy score
Chiviacowsky et al. (2010)	Feedback after good trials	Children (Age: *M* = 10 years, *SD* = NA; *ratio males/females not reported*	CRB (KR after good trials): *n* = 13 CRM (KR after poor trials): *n* = 13 (Total: *N* = 26)	CRB vs CRM (Diminished expectancies)	Pedalo	8 blocks of 4 trials (7 meters)	24h (1 block of 4 trials)	Time
Chiviacowsky et al. (2012)	Perceived task difficulty	Adults (Age: *M* = 21.8 years, *SD* = 3.36; 24 females)	Self-30: *n* = 17 Self-4: *n* = 17 Self: *n* = 17	Self-30 vs Self (Control)	Anticipatory timing	3 blocks of 10 trials	24h (1 block of 10 trials)	Absolute error

Study	*Manipulation*	*Population*	*Sample Size*	*Comparison Group*	*Type of Task*	*Training Session*	*Retention*	*Chosen Performance Outcome*
			(Total: *N* = 51)					
Chiviacowsky et al. (2018)	Perceived task difficulty	Older adults (Age: *M* = 66.1 years, *SD* = 4.78; all females)	Negative stereotype: *n* = 13 Positive stereotype: *n* =13 Control: *n* = 13 (Total: *N* = 39)	Positive stereotype vs Control (Control)	Stabilometer	1 block of 10 trials	24h (1 block of 5 trials)	Time in balance
Chiviacowsky et al. (2019)	Comparative feedback	Adults (Age: *M* = 23.2 years, *SD* = 6.71; 14 females)	Positive temporal-comparative feedback: *n* = 14 Control: *n* = 14 (Total: *N* = 28)	Positive temporal-comparative feedback vs Control (Control)	Golf-putting	5 blocks of 10 trials	24h (1 block of 10 trials)	Deviation
Chung et al. (2020)	Conceptions of ability	Individuals with Parkinson's Disease (Age: *M* = 62.36 years, *SD* = 9.80; 18 females)	Incremental theory: *n* = 15 Incremental theory plus success criteria: *n* = 15 Control: *n* = 14 (Total: *N* = 44)	Incremental theory vs Control (Control)	Stabilometer	1 block of 14 trials (30-s trial)	24-hr (1 block of 7 30-s trials)	Time in balance
Drews et al. (2013)	Conceptions of ability	Children (Age 6: *M* = 6.2 years, *SD* = 0.24; Age 10: *M* = 10.1 years, *SD* = 0.30; Age 14: *M* = 14.4 years, *SD* = 0.34; 54 females)	Acquirable-skill-6: *n* = 20 Inherent-ability-6: *n* = 20 Acquirable-skill-10: *n* = 20 Inherent-ability-10: *n* = 20	Acquirable-skill-6 vs Inherent-ability-6 (Diminished expectancies – Drew et al. (2013[a])) Acquirable-skill-10 vs Inherent-ability-10 (Diminished	Overhand non-dominant arm beanbag throwing	4 blocks of 10 trials	24-hr (1 block of 10 trials)	Accuracy score

Study	*Manipulation*	*Population*	*Sample Size*	*Comparison Group*	*Type of Task*	*Training Session*	*Retention*	*Chosen Performance Outcome*
			Acquirable-skill-14: *n* = 20 Inherent-ability-14: *n* = 20 (Total: *N* = 120)	expectancies – Drew et al., 2013[b]) Acquirable-skill-14 vs Inherent-ability-14 (Diminished expectancies – Drew et al., 2013[c])				
Ghorbani & Bund (2020)	Feedback after good trials	Adults (Age: *M* = 21.35 years, *SD* = 1.86; all males)	Good KR and High Self-Efficacy (SE): *n* = 15 Poor KR and High SE: *n* = 15 Good KR and Low SE: *n* = 15 Poor KR and Low SE: *n* = 15 (Total: *N* = 60)	Good KR and High SE vs Poor KR and High SE (Diminished expectancies- Ghorbani & Bund., 2020[a]) Good KR and Low SE vs Poor KR and Low SE (Diminished expectancies- Ghorbani & Bund., 2020[b])	Non-dominant arm beanbag throwing	10 blocks of 6 trials	24h (1 block of 10 trials)	Accuracy scores
Ghorbani (2019) – Exp. 1	Feedback after good trials	Adults (Age range: 18-24 years; all males)	KR-good: *n* = 12 KR-bad: *n* = 12 Control: *n* = 12 (Total: *N* = 36)	KR-good vs KR-bad (Diminished expectancies)	Underarm dart-throwing	10 blocks of 6 trials	24-hr (1 block of 10 trials)	Accuracy score
Goudini et al. (2018)	Feedback after good trials	Adults (Age: *M* = 24.66 years, *SD* = 1.35; 4 females)	KR after good trials: *n* = 9 KR after poor trials: *n* = 9	KR after good trials vs KR after poor	Line tracking	11 blocks of 6 trials (15-s trial)	48h (1 block of 10 trials)	Duration of errors

Study	*Manipulation*	*Population*	*Sample Size*	*Comparison Group*	*Type of Task*	*Training Session*	*Retention*	*Chosen Performance Outcome*
			(Total: N = 18)	trials (Diminished expectancies)				
Grealy et al. (2019)	Comparative feedback	Adults (Age: M = 22.38 years, SD: 2.32; 28 females – Grealy et al., 2019[a]) Older adults (Age: M = 71.65 years, SD: 4.28; 23 females – Grealy et al., 2019[b])	Young false positive: n = 21 Young veridical: n = 21 (Total: N = 42; Grealy et al., 2019[a]) Older false positive: n = 17 Older veridical: n = 17 (Total: N = 34; Grealy et al., 2019[b])	Young false positive vs Young veridical (Control – Grealy et al., 2019[a]) Older false positive vs Older veridical (Control – Grealy et al., 2019[b])	Inhibitory-action task (Simon task)	18 blocks of 50 trials completed over 6 training sessions (3 blocks/session)	Two-week (3 blocks of 50 trials)	Inhibition time
Harter et al. (2019)	Conceptions of ability	Children (Age: M = 9.6 years, SD = 0.11; all females)	Acquirable-skill: n = 20 Inherent-ability: n = 20 (Total: N = 40)	Acquirable-skill vs Inherent-ability (Diminished expectancies)	Pirouette en dehors	3 blocks of 5 trials	24-hr (1 block of 5 trials)	Punctuation scores
Jennings et al. (2013)	Self-modeling	Adolescents (Age: M = 13.6 years, SD: 1.6; 7 females)	Traditional approach: n = 10 Self-modeling intervention: n = 9 (Total: N = 19)	Traditional approach vs Self-modeling intervention (Control)	Cycling standing start	4 one-hour training sessions over a 2-week period	48-hr (1 trial)	Standing start time
Lessa et al. (2018)	Comparative feedback	Older adults (Age: M = 66.14 years, SD = 4.63; 30 females)	Positive temporal-comparative feedback: n = 17 Control: n = 17	Positive temporal-comparative vs Control (Control)	4-meter walking speed	4 blocks of 10 trials	24 h (1 block of 10 trials)	Absolute error

Study	*Manipulation*	*Population*	*Sample Size*	*Comparison Group*	*Type of Task*	*Training Session*	*Retention*	*Chosen Performance Outcome*
			(Total: *N* = 34)					
Lewthwaite & Wulf (2010)	Comparative feedback	Adults (Age: *M* = 23.0 years, *SD* = 2.26; 24 females)	Better: *n* = 12 Worse: *n* = 12 Control: *n* = 12 (Total: *N* = 36)	Better vs Control (Control)	Stabilometer	2 days with 7 trials (90-s trials)	24h (1 block of 7 trials)	Root Mean Square Error
Navaee et al. (2016)	Comparative feedback	Adults (Age: *M* = 22.60 years, *SD* = 1.89; *information about gender not reported*)	Normative positive feedback: *n* = 10 Normative negative feedback: *n* = 10 Control: *n* = 10 (Total: *N* = 30)	Normative positive feedback vs Control (Control)	Balance	16 blocks of 10 trials for 4 consecutive days (40 trials/day)	24-hr (*number of trials not reported)*	Overall stability
Navaee et al. (2018)	Comparative feedback	Autistic children (Age [5] range: 6-10, *M* = NA, *SD*: NA; *information about gender not reported*)	Normative feedback: *n* = 10 Control: *n* = 10 (Total: *N* = 20)	Normative feedback vs Control (Control)	Non-dominant arm overhead beanbag throwing	6 blocks of 10 trials	24-hr (1 block of 10 trials)	Mean score
Ong & Hodges (2018) - Exp 2a.	Comparative feedback	Adults (Age: *M* = 21.1 years, *SD* = 3.4; all females)	Positive: *n* = 10 Positive-control: *n* = 10 (Total: *N* = 20)	Positive vs Positive-control (Control)	Stabilometer	1 block of 7 trials (60-s trial)	24 h (1 block of 7 trials)	Root Mean square Error

[5] This paper reported mean and SD by group as follows: Normative feedback: 8.40±0.96, Control: 8.50±0.84.

Study	*Manipulation*	*Population*	*Sample Size*	*Comparison Group*	*Type of Task*	*Training Session*	*Retention*	*Chosen Performance Outcome*
Ong et al. (2015)	Perceived task difficulty	Adults (Age: *M* = NA, *SD* = NA; all females)	Large target: *n* = 28 Small target: *n* = 27 (Total: *N* = 55)	Large target vs Small target (Diminished expectancies)	Dart-throwing	10 blocks of 9 trials	1-week (block of 9 trials)	Radial error
Ong et al. (2019)	Perceived task difficulty	Adults (Age: *M* = 21.4 years, Age range: 18-31 years; all females)	Large-target: *n* = 14 Small-target: *n* = 15 (Total: *N* = 29)	Large-target vs Small-target (diminished expectancies)	Dart-throwing	10 blocks of 9 trials	24-hr (1 block of 6 trials – no-vision retention test) [6] 24-hr (1 block of 9 trials – with vision)	Absolute error
Palmer et al. (2016)	Perceived task difficulty	Adults (Age: *M* = 24.6 years, *SD* = 5.20; 22 females)	Large-target: *n* = 17 Small-target: *n* = 17 (Total: *N* = 34)	Large-target vs Small-target (Diminished expectancies)	Golf-putting	5 blocks of 10 trials	24-hr (1 block of 12 trials)	Deviation
Pascua et al. (2015)	Comparative feedback	Adults (Age: *M* = 21.5 years, *SD* = 1.22; 31 females)	External focus/enhanced expectancy: *n* = 13 External focus: *n* = 13 Enhanced expectancy: *n* = 13 Control: *n* = 13 (Total: *N* = 52)	Enhanced expectancy vs Control (Control – Pascua et al. 2015[a]) & External focus/enhanced expectancy vs External focus (Control - Pascua et al. 2015[b])	Non-dominant arm overarm throwing (tennis ball)	6 blocks of 10 trials	24-hr (1 block of 10 trials)	Throwing accuracy scores

[6] For the purposes of the present meta-analysis only the 24-hr retention test with no vision was used.

Study	*Manipulation*	*Population*	*Sample Size*	*Comparison Group*	*Type of Task*	*Training Session*	*Retention*	*Chosen Performance Outcome*
Saemi et al. (2011)	Feedback after good trials	Children (Age: *M* = 10.61 years, *SD*: 0.88; *information about gender not reported*)	KR-good: *n* = 14 KR-poor: *n* = 14 (Total: *N* = 28)	KR-good vs KR-poor (Diminished expectancies)	Overhand non-dominant arm beanbag throwing	10 blocks of 6 trials	24-hr (1 block of 10 trials)	Accuracy score
Saemi et al. (2012)	Feedback after good trials	Adults (Age: *M* = 19.51 years, *SD* = 1.09; all males)	KR after good trials: *n* = 12 KR after poor trials: *n* = 12 (Total: *N* = 24)	KR after good trials vs KR after poor trials (Diminished expectancies)	Non-dominant arm tennis ball tossing	10 blocks of 6 trials	24h (1 block of 10 trials)	Accuracy scores
Steel et al. (2016)	Extrinsic rewards	Adults (Age: *M* = 25 years, *SD* = 4.25; 47 females) [7]	Serial Reaction Time Task (SRTT) (Steel et al., 2016[a]): Reward: *n* = 12 Punishment: *n* = 12 Control: *n* = 12 (Total: *N* = 36) Force-Tracking Task (FTT) (Steel et al., 2016[b]): Reward: *n* = 9 Punishment: *n* = 11	Reward vs Control (Control)	SRTT FTT	SRTT (Steel et al., 2016[a]): Training: 6 blocks of 96 trials FTT (Steel et al., 2016[b]): Training: 6 blocks of 8 trials (12-s trial)	SRTT (Steel et al., 2016[a]): 24-hr and 30-day (3 blocks of 96 trials; sequence: random-fixed-random) FTT (Steel et al., 2016[b]): 24-hr and 30-day (3 blocks of 8 trials; sequence: random-fixed-random)	SRTT (Steel et al., 2016[a]): Reaction time FTT (Steel et al., 2016[b]): Squared error

[7] Authors did not provide information about age (mean and standard deviation) and gender separately for each task. Thus, the information presented is based on the total sample size of 72 participants.

Study	*Manipulation*	*Population*	*Sample Size*	*Comparison Group*	*Type of Task*	*Training Session*	*Retention*	*Chosen Performance Outcome*
			Control: *n* = 10 (Total: *N* = 30)					
Wulf et al. (2010)	Comparative feedback	Adults (Age: *M* = 20.8 years, *SD* = 3.53; 12 females)	Better: *n* = 14 Worse: *n* = 14 (Total: *N* = 28)	Better vs Worse (Diminished expectancies)	Computerized sequential timing	8 blocks of 10 trials	24-hr (1 block of 10 trials)	Overall timing error
Wulf et al. (2012) – Exp 1.	Comparative feedback	Older adults (Age: *M* = 71.1 years, *SD* = 5.25; all females)	Normative feedback: *n* = 15 Control: *n* = 14 (Total: *N* = 29)	Normative feedback vs Control (Control – Wulf et al., 2012[a])	Stabilometer	1 block of 10 trials (30-s trial)	24-hr (1 block of 5 trials)	Time in balance
Wulf et al. (2012) – Exp 2.	Perceived task difficulty	Older adults (Age: *M* = 63.6 years, *SD* = 3.40; all females)	Enhanced expectancies: *n* = 14 Control: *n* = 14 (Total: *N* = 28)	Enhanced expectancies vs Control (Control - Wulf et al., 2012[b])	Stabilometer	1 block of 10 trials (30-s trial)	24-hr (1 block of 5 trials)	Time in balance
Wulf et al. (2013)	Conceptions of ability and feedback after good trials	Adults (Age: *M* = 22.3 years, *SD* = 2.25; 36 females)	Inherent-ability better: *n* = 14 Inherent-ability worse: *n* = 14 Acquirable-skill better: *n* = 14 Acquirable-skill worse: *n* = 14 (Total: N = 56)	Acquirable-skill better vs Inherent-ability worse (Diminished expectancies)	Stabilometer	2 days with 7 trials (90-s trials)	24h (1 block of 7 trials)	Root Mean Square Error

Study	*Manipulation*	*Population*	*Sample Size*	*Comparison Group*	*Type of Task*	*Training Session*	*Retention*	*Chosen Performance Outcome*
Wulf et al. (2014)	Comparative feedback	Adolescents (Age: *M* = 16.7 years, *SD* = 1.14; 28 females)	Autonomy support/enhanced expectancies: *n* = 16 Autonomy support: *n* = 16 Enhanced expectancies: *n* = 16 Control: *n* = 16 (Total: *N* = 64)	Enhanced expectancies vs Control (Control – Wulf et al. 2014[a]) & Autonomy support/enhanced expectancies vs Autonomy support (Control - Wulf et al. 2014[b])	Non-dominant arm overhand throwing (beach tennis ball)	6 blocks of 10 trials	24-hr (1 block of 10 trials)	Accuracy score
Wulf et al. (2018)	Comparative feedback	Adults (Age: *M* = 22.8 years, *SD* = 3.87; 20 females)	Enhanced expectancy and autonomy support: *n* = 15 Enhanced expectancy and external focus: *n* = 15 Autonomy support and external focus: *n* = 15 Enhanced expectancy, autonomy support, and external focus: *n* = 15 (Total: *N* = 60)	Enhanced expectancy, autonomy support, and external focus vs Autonomy support and external focus (Control)	Beach tennis-ball throwing	6 blocks of 10 trials	24h (1 block of 10 trials)	Accuracy scores
Ziv, Lidor, et al. (2019)	Perceived task difficulty	Adults (Age: *M* = 23.90 years, *SD* = 2.7; 32 females)	Large circle: *n* = 15 Small-circle: *n* = 15 Control: *n* = 15 (Total: *N* = 45)	Large circle vs Control (Control)	Golf-putting	5 blocks of 10 trials	48-hr (1 block of 12 trials)	Radial error

Study	*Manipulation*	*Population*	*Sample Size*	*Comparison Group*	*Type of Task*	*Training Session*	*Retention*	*Chosen Performance Outcome*
Ziv, Ochayon, et al. (2019)	Perceived task difficulty	Adults (Age: *M* = NA, *SD* = NA; all males)	Large-circle: *n* = 15 Small-circle: *n* = 15 Control: *n* = 15 (Total: *N* = 45)	Large-circle vs Control (Control)	Golf-putting	5 blocks of 10 trials	48-hr (1 block of 12 trials)	Absolute error
Zobe et al. (2019)	Comparative feedback	Adults (Age: *M* = 22.5 years, *SD* = 2.8; 20 females)	Normative-Positive-Group: *n* = 14 Normative-Negative-Group: *n* = 14 Passive-Control-Group: *n* = 14 (Total: *N* = 42)	Normative-positive-group vs Normative-negative-group [8] (Diminished expectancies)	Elbow-extension-flexion sequence with three movement reversals at 70°, 20°, and 70°	5 sessions (15 blocks total): session 1 was comprised of 3 blocks of 38 trials and sessions 2-5 were comprised of 3 blocks of 48 trials per session	48-72-hr (1 block of 6 trials)	Absolute error

Note. NA indicates information was not available. *KR* indicates knowledge of results. The retention test closest to 24-hr was used in the meta-analysis.

[8] The Passive-Control-Group did not go through the training session, hence our decision to compare the Normative-Positive- Group to the Normative-Negative-Group.

Chapter 2: That Looks Easy! Evidence against the benefits of an easier criterion of success for enhancing motor learning

Designing motor skill practice to optimize learning is critical to many domains including clinical practice, military training, and sports, thus, researchers aim to uncover conditions that maximize learning. The Optimizing Performance Through Intrinsic Motivation and Attention for Learning (OPTIMAL) theory of motor learning predicts that practice conditions that foster motivation facilitate motor learning (Wulf & Lewthwaite, 2016). Motivation, which refers to factors implicated in the direction and intensity of one's effort, can enhance motor learning by driving learners to engage in additional practice (Wulf et al., 2014) and, during a fixed amount of practice, by increasing dopamine release, thereby facilitating the consolidation of motor memories (Wise, 2004).

Practice conditions that enhance learners' expectations for successful outcomes are expected to be motivating. This is because successful outcomes are inherently rewarding, activating the dopaminergic reward system (Lutz et al, 2012) and humans are motivated to pursue rewards during motor skill practice (Moskowitz et al., 2020). Importantly, the mere expectation of dopaminergic medication release modulates the dopamine reward system (Schmidt et al., 2014), which is crucial for motivation (Wise, 2004).

Expectations for successful outcomes can be enhanced by lowering the criterion for success. This can be done by instructing learners that practice trials landing within a large area surrounding a target are considered good. A larger zone of success is hypothesized to influence several factors associated with enhanced expectancies and motivation. It should increase learners' self-efficacy, specifically their confidence they can achieve successful outcomes, while decreasing psychological pressure related to their perceived ability to perform well. Importantly, it indeed adds successful outcomes, further enhancing self-efficacy and reducing psychological pressure. Increases in task-related self-efficacy can also lead to greater task effort (Frömer et al., 2021) and a reduced internal focus on body movements (Wulf & Lewthwaite, 2016) and reductions in psychological pressure are associated with less conscious processing of movements (more automaticity) (Baumeister, 1984; DeCaro et al., 2011; Masters & Maxwell, 2008). Greater effort leads to better performance due to an increased allocation of neural resources toward the task (Frömer et al., 2021), and an external focus of attention and automaticity are too linked with superior performance (Wulf & Lewthwaite, 2016; Baumeister, 1984), likely due to more efficient muscle activation and correlated effector movement (Lohse et al., 2010; Lohse et al., 2014; Lohse & Sherwood, 2012). This higher performance results in rewarding outcomes, activating the dopaminergic reward system, and improving learning (Wulf & Lewthwaite, 2016). Further, successful performance may preclude learners from testing hypotheses about how to correct performance errors, consequently limiting the accrual of declarative knowledge about the learned motor skill, making it less susceptible to deterioration under conditions such as pressure and secondary task demands (Masters & Maxwell, 2008). Finally, successful performance increases learners' perceived competence, thereby promoting intrinsic motivation (Deci & Ryan, 2000).

Several studies have tested the prediction that a larger zone of success enhances motor learning, with mixed results (Chiviacowsky & Harter, 2015; Chiviacowsky et al., 2012; Iwatsuki & Regis, 2021; Ong et al., 2015, 2019, Palmer et al., 2016; Trempe et al., 2012; Ziv et al., 2019, 2021; Ziv & Lidor, 2021). Although zone of success has been manipulated in different ways, for example with modulation of the temporal bandwidth for a coincident-timing task (Chiviacowsky & Harter, 2015; Chiviacowsky et al., 2012) or the size of an area surrounding a target in a golf putting task (Palmer et al., 2016; Ziv et al., 2019, 2021), results do not seem to depend on these features. Rather, the learning benefit of practicing with a larger zone of success may be just a result of chance factors; associated with small number of participants or uneven differences in allocation to groups, or perhaps another covarying factor, such as the absolute number of successful outcomes achieved by the group practicing with the smaller zone of success. For example, Palmer et al. (2016)'s small zone group averaged only ~4 successful outcomes over 50 practice trials and had significantly worse learning than the large zone group, which averaged 11 successful outcomes. Conversely, Ong et al. (2019)'s small zone group averaged ~10 successful outcomes over 90 practice trials and exhibited similar learning in comparison to the large zone group, which averaged ~40 successful outcomes. Importantly, the large zone groups in both studies had increased success compared to the small zone groups, with Ong et al. reporting a greater increase (290.36%) relative to Palmer et al. (178.48%). These results, wherein a learning benefit for the large zone group versus the small zone group is conditioned on the absolute number of successful outcomes achieved by the latter, reflects a pattern in the literature as detailed in Table 1.

If the number of successful trials for the small zone groups is a factor moderating target size effects, this result would be incompatible with the OPTIMAL theory prediction that enhanced expectancies facilitate motor learning (Wulf & Lewthwaite, 2016). Because the large zone groups consistently achieve successful outcomes more often than the small zone groups, as shown in the Table, this should enhance the former's expectancies. A possible explanation for why small and large zone groups do not differ in terms of learning outcomes is that a minimum number of trials is sufficient for learners to associate the outcome with the precipitating action. Specifically, each time an action leads to a successful outcome, a positive reward-prediction error occurs, bringing about a dopaminergic reward signal that increases the value of the action (Lohse et al., 2019). Once this process occurs a minimum number of times, the value of the action may be consolidated through dopaminergic activity (Wise, 2004). Importantly, a large zone group should have diminishing returns from additional successful outcomes because they become more predictable, reducing the magnitude of reward-prediction errors (Lohse et al., 2020). This reduction in reward prediction errors reduces the value added to the precipitating action and the dopamine released for consolidation.

Table 1

Absolute Number of Successes per Group in Studies Manipulating Zone of Success

Study	Absolute Success	
	Small	Large
Trempe et al. (2012)	0.72	15.8
Chiviacowsky & Harter (2015)	1.9	17.3
Ziv et al. (2019)	2.5	7.4
Ziv & Lidor (2021) [1]	3.5	23.10
Palmer et al. (2016)	3.95	11
Iwatsuki & Regis (2021)	6.67	43.47
Ong et al. (2019)	10.17	39.7
Ziv et al. (2021)	14.4	25.6
Ong et al. (2015)	29	74

Note. Studies marked in bold showed a significant difference ($p<.05$) between the small and large success zones on at least one learning test.

To test the hypothesis that benefits associated with practicing with a larger zone of success depends on the absolute number of successful outcomes achieved by the small zone group, we manipulated both target size (large or small zone) and number of trials

[1] Ziv & Lidor (2021) did not use a truly neutral condition during retention. Participants observed the change in target either from a large to a medium target or from a small to a medium target and were informed that the task would be harder or easier than the previous day. This procedure might have threatened the perception of competence of the large target group.

(50 or 100 trials). This manipulation meant that half the participants practicing with the small zone would achieve few successful outcomes (small zone/50-trial group) and the others would achieve more successful outcomes (small zone/100-trial group). A main effect of zone size, where the large zone groups exhibit superior learning compared to the small zone groups, would indicate enhanced expectancies may explain the zone-size effect, consistent with OPTIMAL theory (Wulf & Lewthwaite, 2016). An interaction, such that the zone-size effect is moderated by the number of trials (i.e., only seen for 50-trial groups), would indicate that the zone-size effect is limited to conditions where small-zone learners fail to achieve a minimum number of successful outcomes.

Methods

Sample Size Calculation

We used G*Power 3.1.9.4 (Faul et al., 2009) to calculate the sample size required to detect main effects and interactions in an ANCOVA. We set $\alpha = .05$ [2], power = .80, numerator $df = 1$; number of groups = 4 (small zone/50-trial, small zone/100-trial, large zone/50-trial, and large zone/100-trial); and covariates = 1 (pretest). To set the effect size, we computed the difference between the average zone-size effect in studies wherein the small zone group had few absolute successes (<10; Iwatsuki & Regis, 2021; Palmer et al., 2016; Ziv et al., 2019) and relatively many absolute successes (≥10; Ong et al., 2015, 2019; Ziv et al., 2021), reflecting the hypothesis that the zone-size effect is moderated by the small zone group's absolute number of successes. This effect size computation yielded $f = .226$, and details of the computation can be found in the study's pre-registration form (https://osf.io/9djrx/?view_only=0429bae1daaf4d53b77ca66a89f71a47, Pre-Registration).

[2] The alpha level should have been set to .0294 for the sample size calculation to be consistent with the alpha level that would have been used at the final analysis.

The power calculation yielded a sample size of 156, n = 39/group, which was rounded up to 160, n = 40/group, to account for missing data. A sequential analysis with an interim analysis at N = 80 (n = 20/group) was conducted using the Pocock boundary (interim and final $\alpha \leq .0294$). The interim analysis was exclusively conducted on the main outcome measure and only involved the assessment of learning (see details below). Data collection would terminate at the interim analysis under any of the following conditions: (1) Zone Size x Number of Trials interaction is significant; (2) zone size main effect is significant AND Zone Size x Number of Trials interaction is $f < .10$; (3) Zone Size x Number of Trials interaction is $f < .10$ AND zone size main effect is $f < .10$. The interim analysis revealed no statistically significant main effect of zone size nor a Zone Size x Number of Trials interaction (see Posttest section). Taking a conservative approach, we compared the Zone Size x Trial Number x Posttest interaction effect size against our stopping criteria, since this effect size was larger (f = .095) than the main effect of zone size and the Zone Size x Trial Number interaction effect size. Since this effect size still met our pre-established criteria to stop collecting data (i.e., $f < .10$), we stopped data collection.

Participants

Due to meeting our early stopping criterion, the final sample was composed of 80 participants (M age = 21.52 years, SD = 2.70, 33 males). Participants were healthy undergraduate and graduate students with the preference to throw with their right hand, between the ages of 19 and 40 years, novices to the task, and persons who reported not being allergic to conductive gel, colorblind, at high-risk for serious complications from Covid-19 infection, or having physical impairments precluding comfortable left-arm movements from a seated position. The study was approved by the University Institutional Review Board (Protocol #19-046 EP 1902) and was conducted in agreement with the 1964 Declaration of

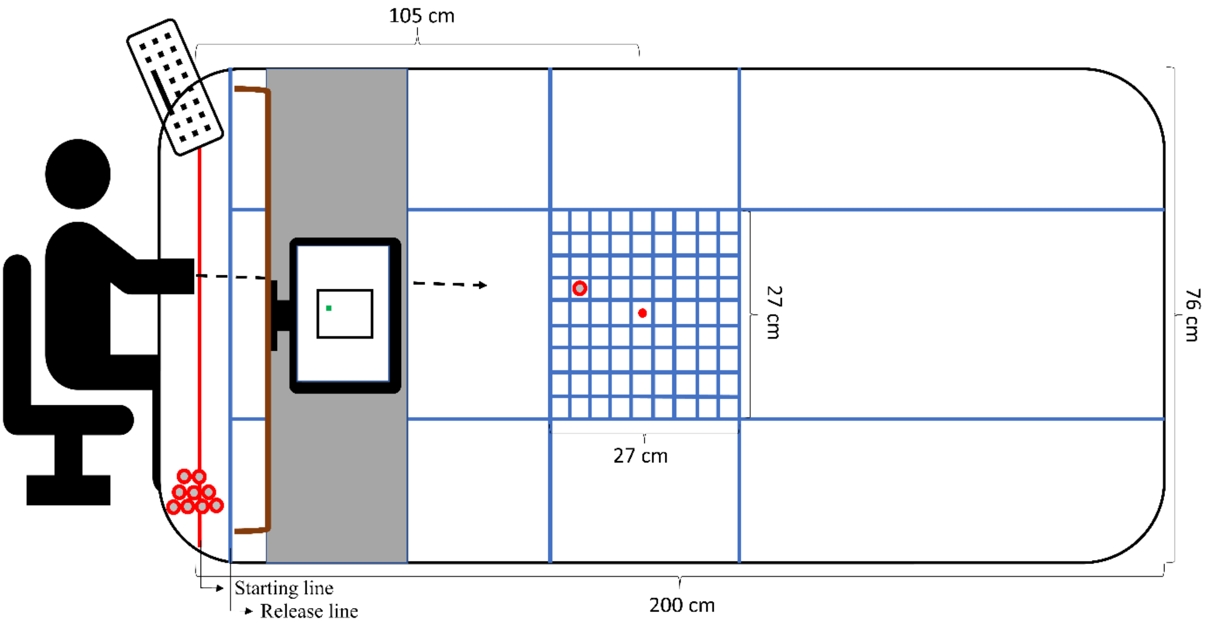
Helsinki. Written informed consent was provided by all participants prior to the beginning of the experiment.

Task

Participants performed a mini-shuffleboard learning task using 10 mini-shuffleboard pucks, 1 at a time (as shown in Figure 1). The pucks consisted of a 1.5 cm diameter red plastic ring encompassing a metal sphere. Participants slid the pucks lengthwise on a rectangular table (213 cm long x 76 cm wide) covered in low-friction adhesive paper with auto-adhesive foam tape (1 cm x 1 cm x 1 cm) lining the side and back edges to prevent the pucks from falling off the table. The start line was drawn with a red marker 13 cm from the edge of the table and followed 10 cm later by the release line, drawn in blue. A 27 cm x 27 cm grid divided in 81 squares of 3 cm x 3 cm was drawn in blue with its center 95 cm past the release line. The grid was connected to the edges of the table by two parallel horizontal lines and two parallel vertical lines that were extensions of the outer lines of the grid. Participants remained seated throughout the whole experiment in a chair positioned so that their left wrist could comfortably reach the release line. A laptop table supporting a computer screen and an occlusion board was positioned after the release line, restricting the participant's vision to about 25 cm of the puck's trajectory after release. The number pad of a keyboard rested on the left side of the shuffleboard table and was used by the participants to initiate the trial and receive feedback (see Procedures below). To perform the task, participants were asked to grip the sides of the puck with their left index finger and thumb and slide the puck under the occlusion board by extending their left arm in a straight line (complete instructions given to participants can be found at https://osf.io/9djrx/?view_only=0429bae1daaf4d53b77ca66a89f71a47, Instructions). Before each shot, participants prepared the puck by positioning it on the start line and were instructed to release it once it reached the release line.

Figure 1

Illustration of Experimental Set-Up



Procedures

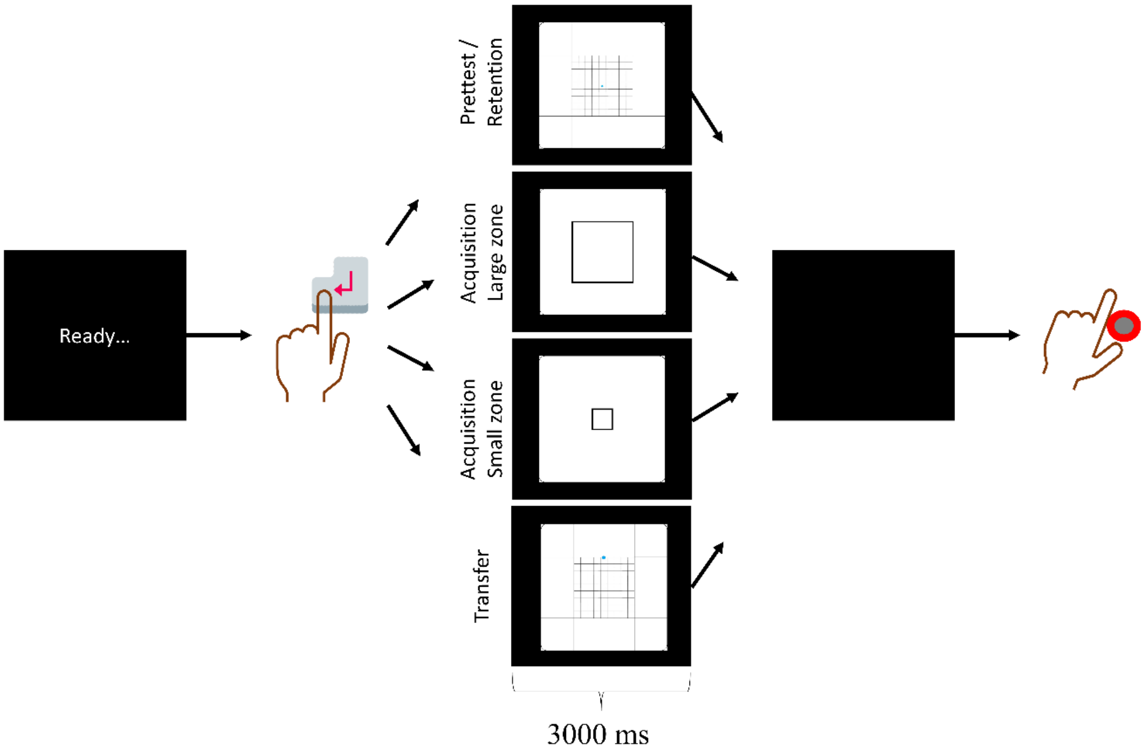
Pretest

Before pretest, participants read and signed the informed consent and completed the Edinburgh Handedness Inventory (Oldfield, 1971). In the pretest, participants were instructed to position the puck when prompted by the word "Ready" from the computer monitor (Figure 2). They then pressed the "enter" key on the keyboard with their left hand, which triggered the appearance of a representation of the grid and target on the screen. The image disappeared from the screen after 3000 ms, at which point participants were allowed to shoot the puck. In the pretest, participants were informed that the goal was to make the puck stop as close to the target as possible. The target was a red dot in the center of the grid. Before the first trial, the occlusion board was removed for approximately 5 s, allowing participants to observe the grid and the target. Next, two non-recorded trials without post-shot augmented feedback were performed, so the participants could get familiar with the procedures and movement.

Participants then completed the pretest self-efficacy questionnaire, marking their confidence, on a scale from 0 to 100, in making the puck stop on the red target (Bandura, 2006). Finally, participants performed 10 trials with no augmented feedback. (All questionnaires can be found at https://osf.io/9djrx/?view_only=0429bae1daaf4d53b77ca66a89f71a47, Questionnaires.)

Figure 2

Trial Initiation



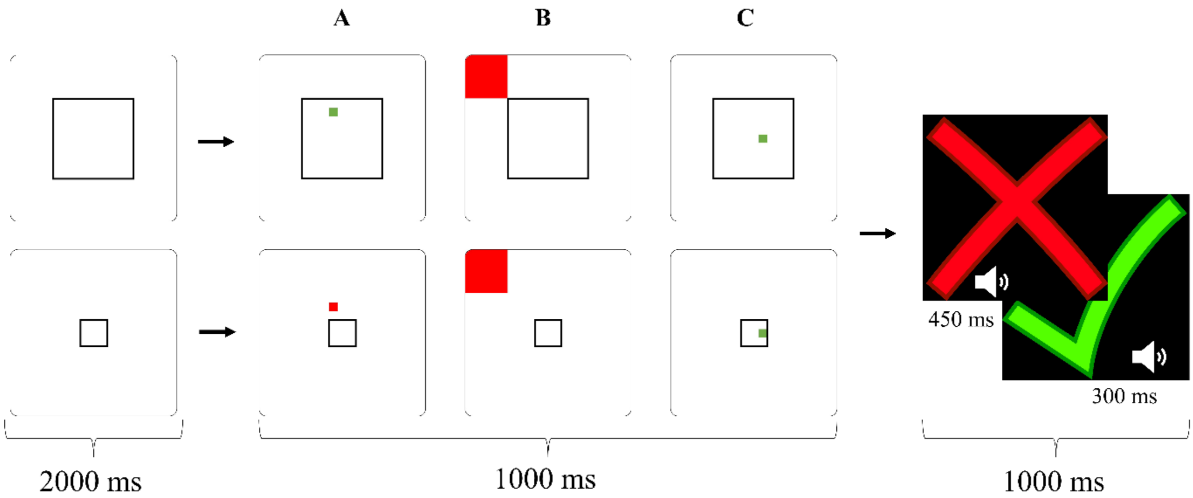
Note. This figure depicts the moments before participant was allowed to start a new puck shot. Once the "Ready…" image was shown on the screen by the experimenter, the participant was able to press "enter" on the keyboard positioned to their left and returned their hand to the puck. A representation of the grid with the target (represented as a blue dot) in the middle (during pretest and retention test) or beyond (during transfer test), or the participant's assigned target zone (large or small during acquisition) then appeared on the screen for 3000 ms. The participant was allowed to start the movement any time after the image disappeared from the screen.

Acquisition

Each participant was assigned to one of the four groups (large zone/50-trial, large zone/100-trial, small zone/50-trial, small zone/100-trial, 20 participants per group) and one of two posttest orders (retention-transfer/transfer-retention) based on a pre-determined list. Participants were randomly assigned to the eight group x order combinations, stratified by gender. Participants in the 50-trial groups practiced five blocks of ten trials and this was 10 blocks of 10 trials for the 100-trial groups. Participants were informed there was a target zone they were trying to achieve, such that trials where the puck stopped inside this zone would be considered "good". For participants in the small zone groups, the zone was a 9 cm x 9 cm square centered on the target in the middle of the grid, while for participants in the large zone groups, the zone was a 27 cm x 27 cm square centered on the target. Zone sizes and number of trials were defined after two pilot studies described in the supplementary material at the OSF repository (https://osf.io/9djrx/?view_only=0429bae1daaf4d53b77ca66a89f71a47, Supplementary Material). Before the first trial of acquisition, participants were shown the zone surrounding the target overlayed on the grid. Specifically, a green cardboard square the size of the participant's assigned zone was positioned on the grid, and the occlusion board was removed for approximately 5 s. Participants then completed the acquisition self-efficacy questionnaire, marking their confidence, on a scale from 0 to 100, in making the puck stop in the zone. The experimenter pointed out the difference between the pretest self-efficacy questionnaire, which referred to the target, and the acquisition self-efficacy questionnaire, which referred to the zone. During acquisition, participants performed the task just like they did on the pretest, with a couple of exceptions. First, the image they saw on the computer screen before shooting was a square outlined in black representing the zone, instead of the image of the grid and target (Figure 2). Second, augmented feedback was provided after every trial. To receive feedback, participants pressed “Enter” on the keyboard when prompted by the word “Ready” on the computer screen. Then, the representation of the zone appeared on the computer screen for 2000 ms (Figure 3). Next, the square in the grid where the puck stopped was highlighted in green or red for 1000 ms, depending on whether the puck landed within the zone (a good shot) or not (a bad shot), respectively. That is, participants in both groups were shown the square in the grid where the puck landed on every trial, unless it landed outside the grid. In this case, a rectangle was highlighted in red to indicate the shot was far-left, far-center, far-right, left, right, short-left, short-center, or short-right (see an example in Figure 3 column B). Subsequently, participants saw a green checkmark for 1000 ms and heard a “correct” sound (sound length ≈300 ms), or a red “X” for 1000 ms with an “invalid” sound (sound length ≈450 ms), depending on whether the trial was good or not (all stimuli, including sound files can found at https://osf.io/9djrx/?view_only=0429bae1daaf4d53b77ca66a89f71a47, Stimuli). Participants had a 1-min break between blocks.

Figure 3

Feedback Presentation



Note. This figure depicts the moments after participants pressed enter on the keyboard to receive feedback. First, an image depicting their assigned target zone was presented for 2000 ms, and then the square where the puck stopped was highlighted in green or red for 1000 ms. Column A represents the same outcome being considered positive for the large zone groups but negative for the small zone groups, while column B represents a negative outcome for all groups and column C depicts an outcome considered positive for all groups. Lastly, participants were presented with a red cross (1000 ms) and an invalid sound (≈450 ms) if the trial was considered negative, or a green checkmark (1000 ms) and correct sound (≈ 300 ms) if the trial was considered successful.

Post-Acquisition Questionnaires

After the last block, participants completed several questionnaires. They completed a task adapted version of the Conscious Motor Processing subscale of the Movement Specific Reinvestment Scale (Vine et al., 2013). There were six questions asking participants to indicate how often they had certain thoughts while shooting (e.g., “I reflected about my technique”). The scale had five points and was anchored by “1-Never” and “5-Always” with

“3-Sometimes” in the middle. A focus of attention questionnaire asked participants to best describe where they were focusing their attention while shooting the puck: arm, hand, fingers, puck, the path they wanted the puck to take, or the target. Participants could choose up to 3 answers. Participants completed the Interest/Enjoyment (intrinsic motivation); Effort/Importance; Perceived Competence; and Pressure/Tension subscales of the Intrinsic Motivation Inventory (IMI; McAuley et al., 1989). Each subscale had five to seven statements about participants’ experience with the task and asked them to indicate how true the statements were on a seven-point scale anchored by 1 = “Not True at All” and 7 = “Very True” (with 4 = “Somewhat True” in the middle). The Interest/Enjoyment subscale included statements such as “I enjoyed doing this task very much”; the Effort/Importance subscale included statements such as “I put a lot of effort into this activity”; the Perceived Competence subscale included statements such as “I think I am pretty good at this activity”; and the Pressure/Tension subscale included statements such as “I felt pressured while doing this activity”. Finally, participants indicated their agreement with six statements about their objectives while performing the task and how they assessed their performance (e.g., “During the task, I was aiming to make the puck stop anywhere inside the target zone”; “During the task, I thought my performance was good when the puck stopped anywhere inside the target zone”). The order of the questions was randomized across participants. Participants responded with a five-point scale anchored by 1 = “Not at All” and 5 = “Completely”. This questionnaire was implemented for potential exploratory analyses.

Posttests

Approximately 24-hr after acquisition, participants returned to the lab to perform retention and transfer tests. The retention test was performed exactly like the pretest, except that the red dot representing the target was moved 13.5 cm farther away from the participant during transfer test. The order of the tests was counterbalanced across participants. The grids and the respective targets were shown to the participants for approximately 5 s before each test. Before retention, after seeing the target, participants completed the retention test self-efficacy questionnaire, which was the same as the one used before the pretest. After the posttests, participants completed a free recall questionnaire, which asked them to report, in as much detail as possible, any rules, methods, or techniques they recalled using to shoot the pucks on the second day. This questionnaire assessed declarative knowledge about the skill.

Data Processing and Statistical Analysis

Our main outcome measure of performance and learning was radial error (RE), representing accuracy (Hancock et al., 1995). As a secondary measure, we used bivariate variable error (BVE), given by the square root of the 10 shots' mean squared distance from the centroid (Hancock et al., 1995), to assess precision. To extract the x and y coordinates of the puck's stopping position on the table, an iPad was fixated to the ceiling above the table, and photographs were taken after each trial with a wireless clicker. These photographs were then analyzed with LabView® software using the virtual instrument ScorePutting (Neumann & Thomas, 2008) to determine the distance between the center of the puck and the target along the x- and y-axis and use these distances to calculate RE and BVE. RE and BVE were then averaged across trials for each participant on the tests (1 block of 10 trials for each test) and acquisition phase (5 or 10 blocks of 10 trials, depending on the participant's group). To assess learning, a 2 (Zone Size: small/large) x 2 (Trial Number: 50/100) x 2 (Posttest: retention/transfer) ANCOVA with repeated measures on the last factor was conducted, with pretest RE or BVE serving as the covariate, depending on the outcome measure. To assess acquisition performance for the 50-trial groups, acquisition phase RE and BVE were analyzed with 2 (Zone Size) x 5 (Blocks) ANCOVAs with repeated measures on the last factor, with pretest RE or BVE serving as the covariate. A similar analysis was conducted for the 100-trial groups, but with 10 (Blocks). A one-tailed paired samples *t*-test was used to evaluate whether there was change between pretest RE and retention RE, regardless of group, to assess whether participants learned the task.

For the questionnaires, we averaged across single items in the Conscious Motor Processing subscale and IMI subscales as well as calculated their reliability with Cronbach's alpha (described in detail in the supplementary material). These subscales were submitted to 2 (Zone Size) x 2 (Trial Number) ANOVAs. Results of the self-efficacy scale given before pretest, acquisition, and retention test were analyzed with a 2 (Zone Size) x 2 (Trial Number) x 2 (Time: acquisition/posttest) ANCOVA with repeated measures on the last factor and pretest serving as the covariate. Results of the focus of attention questionnaire and the questionnaire about participants' objectives while performing the task and how they assessed their performance were averaged by group and reported descriptively. Two indices of declarative knowledge use were extracted from participants' responses on the free recall test. The first index, 'all concepts,' refers to the number of statements about a concept (rule) (e.g., "I would start with the puck as centered as possible in the exact position as the puck before it."), ignoring statements irrelevant to technical performance (e.g., "Elbow sits comfortable before each shot"). The second index, 'hypothesis testing', refers to statements indicating the participant tested hypotheses related to their putting movement (e.g., "I tried to not extend my wrist flick as hard, so the puck had a better chance of gliding into the center."). That is, hypothesis testing statements are those that indicate the participant made a prediction about the relationship between their putting movement and outcome (Maxwell et al., 2001). We ignored retrospective statements (e.g., "I used fingers of my left hand to push") that may not have been used or thought about while shooting, and/or that were included in the task instructions. We planned to analyze results from the Free Recall questionnaire using a 2 (Zone Size) x 2 (Trial Number) MANOVA, with the two indices serving as dependent variables. However, given that the "all concepts" and "hypothesis testing" indices of declarative knowledge did not meet the assumptions of multivariate normality (Mardia's skeweness = 45.14, $p < .001$; Mardia's Kurtosis = 2.67, $p = .001$) and homogeneity of variance-covariance matrices ($M(9) = 17.17$, $p = .046$) required to run a MANOVA, we ran two separate 2 (Zone Size) x 2 (Trial Number) ANOVAs.

Considering that motivation is the main mechanism through which the OPTIMAL theory predicts criterion of success to affect performance and learning, we conducted an exploratory mixed-effects regression model with RE at the posttests and last block of acquisition as the dependent variables, to further investigate the influence of intrinsic motivation on learning tests and acquisition performance at the individual level. For the model, pretest RE, condition (zone size, trial number, and their interaction), time (last block of acquisition/retention/transfer), the interaction between condition and time, intrinsic motivation (IMI Interest/Enjoyment subscale score), and the interaction between intrinsic motivation and time were entered as fixed-factors, while participant served as a random-effect. For this model, all continuous variables were mean-centered, and all categorical variables were contrast-coded.

For all inferential analyses, alpha was set to .05, and Tukey HSD was used for post-hoc tests when necessary. The Greenhouse-Geisser correction was applied when sphericity was violated, and we evaluated whether all other assumptions of statistical tests were met. All statistical analysis were conducted in R (cran.r-project.org) and can be found at the OSF repository (https://osf.io/9djrx/?view_only=0429bae1daaf4d53b77ca66a89f71a47, R Project).

Results

Success Rate

As shown in Table 2, our manipulation was successful in producing more successful trials for the large zone and 100-trial groups, as compared to the small zone and 50-trial groups, respectively.

Table 2

Average Number of Successful Trials Achieved by Each Group During the Acquisition Phase.

Group	Successful Trials	% Successful Trials
Large Zone/50-Trial	22.75 (*SD* = 7.03)	45.5 (*SD* = 14.06)
Large Zone/100-Trial	47.05 (*SD* = 11.28)	47.05 (*SD* = 11.28)
Small Zone/50-Trial	4.57 (*SD* = 2.61)	9.14 (*SD* = 5.22)
Small Zone/100-Trial	10.00 (*SD* = 3.09)	10 (*SD* = 3.09)

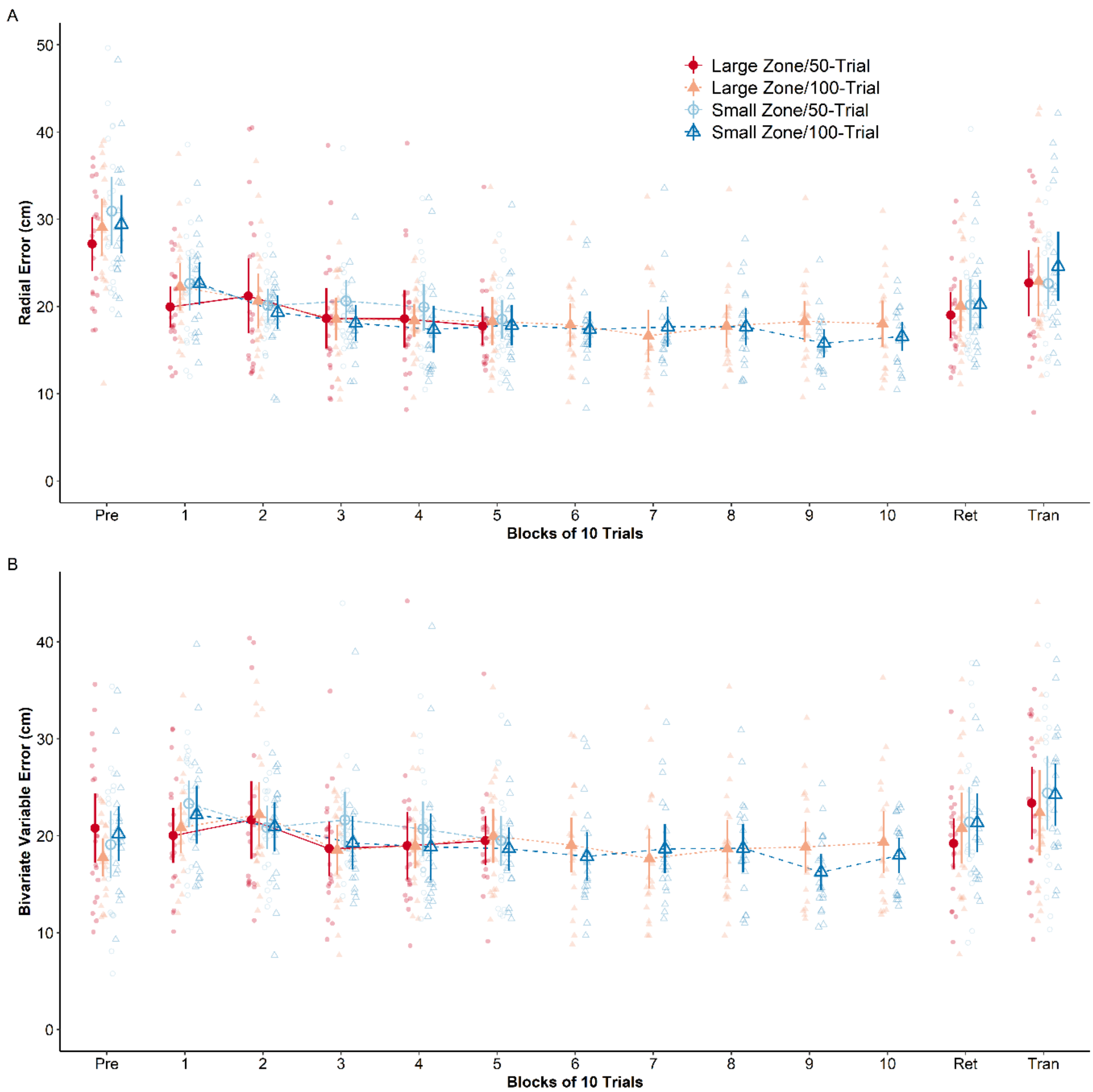
Posttests

Radial Error

To assess whether participants learned the task, we ran a one-tailed paired samples *t*-test to assess the difference in performance between pretest RE and retention RE, regardless of training condition. Participants performed with less error on the retention test ($M = 19.88$ cm, $SD = 5.95$) compared to the pretest ($M = 29.15$ cm, $SD = 7.33$, $t(79) = 9.85$, $p < .001$, Hedges' $g = 1.37$, 95% CI [0.99, 1.75]), suggesting learning (Figure 4A).

Figure 4

Shuffleboard Task Accuracy (A) and Precision (B) as a Function of Study and Group



Note. Each large data point represents the group average, while smaller points represent individual data. Error bars represent 95% CIs. Lower values on the y-axis indicate greater accuracy or precision. Pre = pretest; Ret = retention test; Tran = transfer test.

For the primary confirmatory analysis of interest, the mixed-factor ANCOVA assessing the effect of training conditions on posttest RE did not reveal main effects of zone size ($F(1, 75) = 0.08$, $p = .782$, $\eta^2_p < .01$), trial number ($F(1, 75) = 0.37$, $p = .544$, $\eta^2_p = .01$) or posttest type ($F(1, 75) = 0.29$, $p = .593$, $\eta^2_p < .01$). Also, no Zone Size x Trial Number

(F(1, 75) = 0.16, p = .694, $\eta^2_p < .01$), Zone Size x Posttest (F(1,75) < 0.01, p = .970, $\eta^2_p <$ .01), Trial Number x Posttest (F(1,75) = 0.08, p = .772, $\eta^2_p < .01$), or Zone Size x Trial Number x Posttest interactions (F(1, 75) = 0.68, p = .411, $\eta^2_p = .01$) were detected.

Equivalence Test

After completing the study, it came to our attention that manually specifying the effect size at which to stop data collection for futility, risks a type 2 error (Lakens et al., 2021). To address this risk, we conducted an equivalence test for ANOVA. Specifically, we used the TOSTER package in R (Lakens et al., 2018) to compare the largest observed between-subjects effect size of interest (zone size or Zone x Trial Number)[3] against $\eta^2_p =$ .048594 (f = .226), which is the expected effect size used in our sample size calculation that yielded a sample size approximately as large as we were willing to collect. Based on this equivalence test, we rejected the presence of effects more extreme than $\eta^2_p = .049$ (p =.047).

Bivariate Variable Error

The mixed-factor ANCOVA assessing the effect of training conditions on posttest BVE did not reveal a main effect of zone size (p = .339, $\eta^2_p = .01$), trial number (p = .859, η^2_p < .01) or posttest type (p = .136, $\eta^2_p = .03$). Also, no Zone Size x Trial Number (p = .691, η^2_p < .01), Zone Size x Posttest (p = .961, $\eta^2_p < .01$), Trial Number x Posttest (p = .455, $\eta^2_p =$ .01), or Zone Size x Trial Number x Posttest interactions (p = .464, $\eta^2_p = .01$) were found (Figure 4B).

Acquisition Phase

Radial Error

[3] Caldwell (2022, https://aaroncaldwell.us/TOSTERpkg) explains that TOSTER's equivalence test for F-tests can be extended from one-way ANOVA to factorial ANOVA, but we believe this is limited to between-subjects factors.

For the 100-trial groups, the mixed-factor ANCOVA assessing the effect of zone size on accuracy in the acquisition blocks did not reveal a main effect of zone size (p = .524, η^2_p = .01), block (p = .209, η^2_p = .04), or Zone Size x Block interaction (p = .765, η^2_p = .02). Similarly, for the 50-trials groups, no main effect of zone size (p = .803, η^2_p < .01), block (p = .784, η^2_p = .01), or Zone Size x Block interaction (p = .410, η^2_p = .03) were found (Figure 4A). Our ability to detect block effect in these analyses was likely constrained by including a covariate of pretest radial error, which explained a substantial amount of variance in radial error during practice (p < .138, $\eta^2_p \geq$ .06), reducing the amount of variance that could be explained by block.

Bivariate Variable Error

For the 100-trials groups, the mixed-factor ANCOVA assessing the effect of zone size on precision in the acquisition blocks revealed no main effect of zone size (p = .307, η^2_p =.03), block (p = .195, η^2_p = .04) or Zone Size x Block interaction (p = .743, η^2_p = .02). Similarly, for the 50-trials groups, no main effects of zone size (p = .239, η^2_p = .04), block (p = .599, η^2_p = .02), or Zone Size x Block interaction (p = .210, η^2_p = .04) were found (Figure 4B).

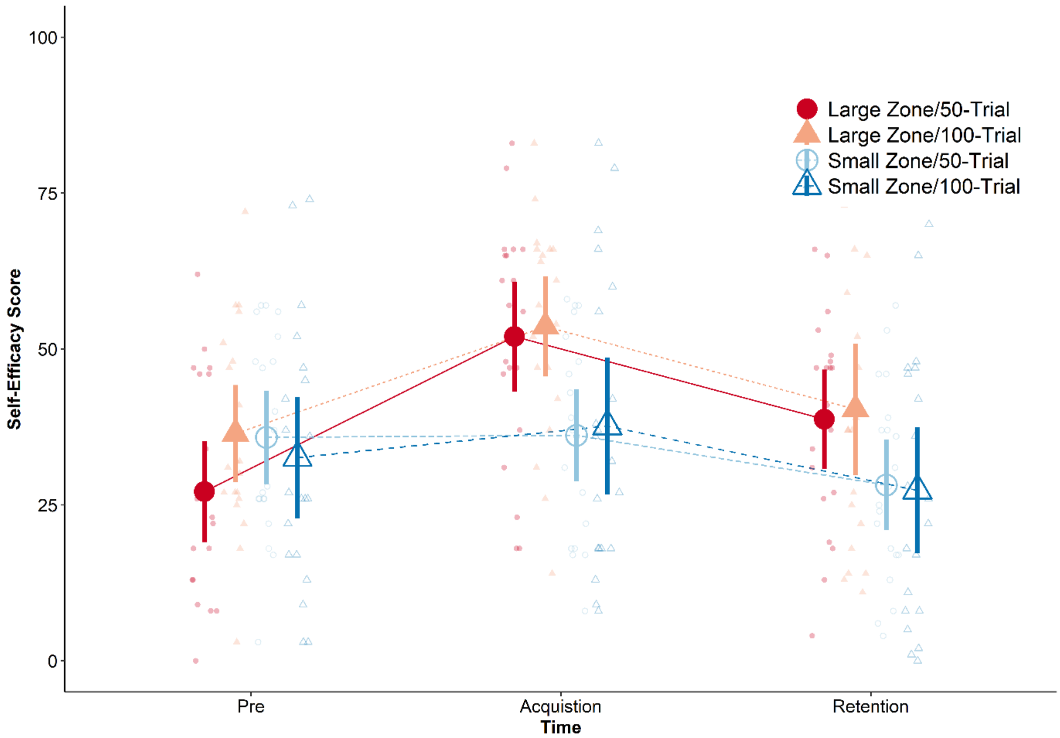
Questionnaires

Self-Efficacy

The mixed-factor ANCOVA assessing the effect of training conditions and time on participants' self-efficacy revealed a significant zone size effect, $F(1, 75) = 33.94$, $p < .001$, η^2_p = .31 as illustrated in Figure 5. As predicted, participants practicing with the large zone reported increased self-efficacy than those practicing with the small zone. Increased self-efficacy was also reported before acquisition, as compared to before retention, $F(1, 75)$ = 8.05, p = .006, η^2_p = .10. There was no significant effect of trial number (p = .690, η^2_p < .01), nor interactions (p's > .189, η^2_p's < .02).

Figure 5

Self-Efficacy as a Function of Phase and Group



Note. Each large data point represents the group average, while smaller points represent individual data. Error bars represent 95% CIs. Higher scores indicate greater self-efficacy. Pre = pretest.

Intrinsic Motivation Inventory

For the Perceived Competence subscale (Figure 6A), participants who practiced with the large zone perceived themselves as more competent in the task than participants with the small zone $F(1, 76) = 28.56$, $p < .001$, $\eta^2_p = .27$. There was, however, no effect of trial number ($p = .379$, $\eta^2_p = .01$) or Zone Size x Trial Number interaction ($p = .647$, $\eta^2_p < .01$).

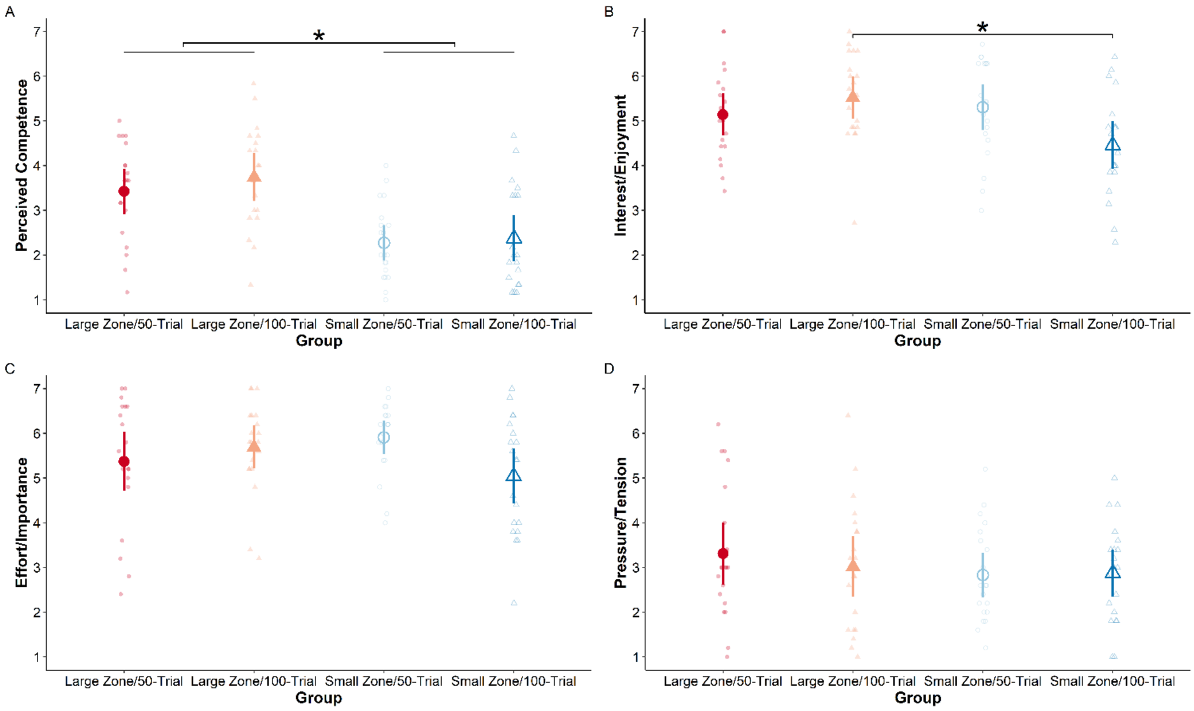
For the Interest/Enjoyment subscale (Figure 6B), there was no main effect of zone size ($p = .063$, $\eta^2_p = .04$) or trial number ($p = .326$, $\eta^2_p = .01$), but there was a significant interaction, $F(1, 76) = 6.65$, $p = .012$, $\eta^2_p = .08$. Follow-up Tukey HSD tests indicated that participants with the large zone reported higher levels of intrinsic motivation than those with the small zone, when a total of 100 acquisition trials were practiced (p = .012), but not with a total of 50 trials (p = .962). No other significant pairwise differences were found (p's > .064).

Regarding the Effort/Importance subscale (Figure 6C), there were no group differences as a function of zone size (p = .833, $\eta^2_p < .01$) or trial number (p = .294, η^2_p = .01), but there was a significant interaction, $F(1, 76) = 5.23$, p = .025, η^2_p = .06. Tukey HSD post hoc test was not sensitive (p's > .093) to differences between groups.

For the Pressure/Tension subscale (Figure 6D), there were no significant main effects of zone size (p = .281, η^2_p = .02), trial number (p = .668, $\eta^2_p < .01$), nor an interaction (p = .571, $\eta^2_p < .01$).

Figure 6

Intrinsic Motivation Inventory Subscale Scores as a Function of Group



Note. Perceived competence (A), Interest/Enjoyment (B), Effort/Importance (C), and Pressure/Tension (D). Each large data point represents the group average, while smaller points represent individual data. Error bars represent 95% CIs. * indicates significant differences.

Conscious Motor Processing Subscale of the Movement Specific Reinvestment Scale

The reliability was poor for this scale, thus raising issues about the validity of the data. There were no zone size (p = .898, η^2_p < .01) nor trial number (p = .971, η^2_p < .01) main effects, but there was an interaction, $F(1, 73) = 9.91$, $p = .002$, $\eta^2_p = .12$. However, no pairwise comparisons were statistically significant following Tukey's HSD test (p's > .100).

Free Recall

The two-way ANOVA assessing the effect of training conditions on the "hypothesis testing" index of the declarative knowledge revealed a significant effect of zone size (F(1, 76) = 4.85, p = .031, η^2_p = .06), such that participants in the large zone group engaged in more hypothesis testing than participants with the small zone. No significant effect of trial number (p = .127, η^2_p = .03) or Zone Size x Trial Number interaction (p = .274, η^2_p = .01) was found. Regarding the second index, "all concepts" (Figure 7B), no main effect of zone size (p = .800, η^2_p < .01), trial number (p = .078, η^2_p = .04), or Zone Size x Trial Number interaction (p = .554, η^2_p < .01) were found.

Focus of Attention and Participant's Objectives Questionnaires

Descriptively, participants in the small-zone groups seem to distribute their attention more than participants in the large-zone groups, but no clear preference regarding the direction of their attentional focus was detected across groups. Lastly, participants accepted their assigned zone of success. [4] Given the lack of relevant group differences found, the

[4] Based on the suggestion of an anonymous reviewer, we conducted a sensitivity analysis to determine whether the degree to which participants reported their objective was to make the puck stop in the center of the target explained learning or moderated the effect of zone size or trial number on learning. Specifically, we added participants' response to the item stating, "During the task, I was

detailed descriptive analysis of these questionnaires can be found in the supplementary material.

Motivation x Performance – Exploratory Analysis

Intrinsic motivation at the end of the acquisition phase did not predict practice or posttest performance. This conclusion was based on the absence of a significant main effect of intrinsic motivation (β = 0.19, SE = 0.50, t = 0.39, p = .699) and an Intrinsic Motivation x Time interaction (β = 0.079, SE = 0.40, t = 0.20, p = .84), after controlling for pretest RE, zone size, trial number, time, and Zone Size x Trial Number x Time interaction.

Discussion

OPTIMAL theory predicts that providing a learner with an easier criterion of success during practice should enhance learning (Wulf & Lewthwaite, 2016). Although some studies show benefits of practicing with a large zone of success (Iwatsuki & Regis, 2021; Palmer et al., 2016; Ziv et al., 2019), other studies, in which the group with the small zone achieved a

---

aiming to make the puck stop in the center of the target zone" as an independent variable in the primary confirmatory analysis of interest (the ANCOVA that had radial error as the dependent variable, pretest radial error as the covariate, and zone size, trial number, and posttest type as independent variables). We did not find evidence that the degree to which participants reported their objective was to make the puck stop in the center of the target had a significant main effect (p = .169) or significantly moderated the effects of the other independent variables (ps ≥ .239). (We considered creating an independent variable based on the average of the item we used and two other items: "If the puck was off-center, I felt like I made a mistake even if I was in the target zone" and "The farther I was from the center, the more I tried to improve on the next trial", but reliability among the items was not good (Cronbach's α = .61). Thus, we proceeded with the item we believed to be most consistent with the reviewer's suggestion to assess the effect of task strategy.)

relatively large number of successes, have failed to find an effect (Ong et al., 2015, 2019; Ziv et al., 2021; Ziv & Lidor, 2021). This pattern in the literature could indicate that the number of successes achieved by the small zone group moderates the zone size effect on learning, which is incompatible with the predictions of the OPTIMAL theory. Thus, to investigate this possibility, we manipulated quantity of practice to affect the absolute number of successes achieved by learners practicing with different criteria of success.

As expected, the large zone groups achieved more successful trials than the small zone groups during practice, despite showing the same overall accuracy and precision. Contrary to the predictions of OPTIMAL theory, we did not find evidence that practicing with a large zone of success facilitated learning. An equivalence test revealed that, even if the zone size effect existed, it would be considered small ($f < .226$, i.e., less than 2.62 cm [5]). Importantly, we were unable to find benefits of the large zone even when the number of successes achieved by learners with the small zone was substantially restricted (average of 4.57 good trials) and similar to experiments wherein the criterion of success effect was found (Chiviacowsky & Harter, 2015; Chiviacowsy et al., 2012, Iwatsuki & Regis, 2021; Palmer et al., 2016; Trempe et al., 2012; Ziv et al., 2019). These results add to the previous literature questioning the benefits of practicing with an easier criterion of success (Ong et al., 2015, 2019; Ziv & Lidor, 2021; Ziv et al., 2021), and are supported by a recent meta-analysis on the effects of enhanced expectancies on learning. This meta-analysis showed that manipulations of perceived task difficulty, which include manipulations of criteria of success, have at best a small effect on learning (Hedges' $g = 0.46$, Bacelar, Parma, Murrah, et al., 2022).

[5] To get to this estimation, we converted $f = 0.226$ to $d = 0.452$, and then multiplied 0.452 by the standard deviation of the average post-test performance across participants. The average post-test performance across participants was calculated by averaging retention and transfer performance for each participant and then taking the average of this value across participants.

Importantly, this effect was deemed overestimated by the authors due to evidence of reporting bias in the literature, consistent with our finding that the zone size effect is small, if existent.

The zone size effect was not present, despite the success of our manipulation in increasing self-efficacy and perceptions of competence, which were also enhanced in previous studies on criterion of success (Chiviacowsky et al., 2012; Chiviacowsky & Harter, 2015; Iwatsuki & Regis, 2021; Ong et al., 2015, 2019; Trempe et al., 2012; Ziv & Lidor, 2021). According to OPTIMAL theory, increased self-efficacy and perceptions of competence should result in increased intrinsic motivation (Wulf & Lewthwaite, 2016). Indeed, the learners who performed 100 practice trials with the easier criterion of success were found to have higher levels of intrinsic motivation than those practicing with the difficult criterion. This result adds to mixed evidence regarding the effects of an easier criterion of success on intrinsic motivation (Chiviacowsky et al., 2012; Ong et al., 2019).

Despite the proposed role of motivation in mediating the effect of a lower criterion of success on performance and learning, increasing motivation did not result in better performance or learning. Similarly, we also did not find evidence that, at the individual level, motivation predicted learning or acquisition performance. To the best of our knowledge, ours is the only study to assess the effects of motivation on performance and learning at the individual level in criterion of success paradigm. There are other studies, with different manipulations related to OPTIMAL theory, where motivation has been assessed and here the results are also quite mixed. Specifically, intrinsic motivation was positively associated with performance at the end of practice in Bacelar et al. (2020) and Grand et al. (2017), but most studies have failed to show a relationship between motivation and learning as assessed in delayed posttest (Bacelar et al., 2020; Grand et al., 2017; Leiker et al., 2016; and Leiker et al., 2019), with the exception of Bacelar, Parma, Cabral et al. (2022), which is, notably, the study with the highest sample size ($N = 200$). These results suggest that the relationship between intrinsic motivation and learning may be much smaller than that between intrinsic motivation and performance, thus requiring larger sample sizes to detect.

Besides motivation, self-efficacy, and perceived competence, other psychological factors have been hypothesized to underlie learning effects in zone size studies, but we did not find evidence supporting these predictions. For instance, a large zone was not linked to increased effort or decreased tension during practice. Regarding focus of attention, although our data suggest the groups with the easier criterion distributed their attention to fewer factors, we did not find any clear pattern distinguishing the direction (external vs. internal) of the attentional focus among the different groups. The accrual of declarative knowledge or use of conscious processing were not significantly increased for the small zone group. In fact, participants with the large zone were shown to engage in activities of hypothesis testing to a greater extent than those with the small zone, indicating that they accrued more declarative knowledge. This result could be due to the large zone groups having approximately equal numbers of unsuccessful as successful trials, which afforded them the opportunity to compare movements strategies that precipitated both types of trial outcomes and hypothesize about the relationship between strategies and outcomes. Although mechanistic assumptions regarding effort, pressure, focus of attention, conscious processing, and declarative knowledge are often drawn in zone size studies and by proponents of OPTIMAL theory, this is one of the first studies to measure these variables (see Ong et al. 2015, 2019, who also used free recall questionnaires to assess explicit knowledge, showing no differences between groups).

This study has strengths and a limitation worth noting. Although this study was pre-registered and had the largest sample size among those investigating the effect of criterion of success on motor learning, a post-hoc sensitivity analysis revealed that we only had 80%

power to detect effect sizes of f > .343 (Faul et al., 2009)[6]. However, this limitation is mitigated by the result of the equivalence test, that showed that, even if the zone size effect is real, it is likely smaller than $f = .226$ ($\eta^2_p = .049$).

Conclusion

Based on the current methods and data, we question the learning benefits of easing success criteria. We present evidence showing that the zone size effect is small, *if existent*. Although easing criteria of success resulted in increased self-efficacy, perceptions of competence, and, for participants with more practice trials, intrinsic motivation, the manipulation of these key psychological variables did not entail increased motor learning or performance. Moreover, at the individual level, intrinsic motivation did not explain motor learning or performance. Therefore, our results challenge key tenets of OPTIMAL theory and prevent us from broadly recommending easing criteria of success during practice, given that other theories hypothesize that 'optimal' learning may be associated with making errors (Lohse et al., 2019) and only modest success during practice (Guadagnoli & Lee, 2004; Hodges & Lohse, 2022).

These data show that any direct effect of relaxing the criteria for success on long term learning is trivially small. As such, these manipulations have little utility for practitioners looking to improve learning. The manipulation did, however, increase motivation which might be valuable tool for practitioners in and of itself. That is, if one learner is struggling with motivation, relaxing the criteria for success will not improve learning, but may increase their motivation, allowing them to persist in practice longer. These trivial effect sizes also have theoretical implications for researchers. Perceived competence, motivation, and learning

[6] For the sensitivity analysis, we used G*Power 3.1.9.4. We set the statistical test to ANCOVA, selected the option "sensitivity analysis", and inputted $\alpha = .0294$, power = .80, N = 80, numerator df = 1, number of groups = 4, and number of covariates = 1.

do not appear to be as tightly coupled as originally hypothesized in OPTIMAL theory. This is not to say that motivation or competence are unimportant for learning, but the moderating effect of motivation appears to be more complicated than we first thought and not easily manipulated.

Chapter 3: It's Subjective! Effects of Perceptions of Success on Neural Correlates of Feedback Processing and Movement Preparation

Motor performance is crucial for daily-life activities, labor skills, as well as sports, and many factors influence how one executes and adapts their movement to achieve task goals. One of the most important factors is how feedback is processed in the brain because it has been shown to have a bidirectional relationship with performance. For example, Lohse et al. (2020) showed that feedback about a performance outcome affects brain activity, which, in turn, can predict the learner's likelihood of maintaining or changing their subsequent behavior. Similarly, prior motor performance influences the brain activity in preparation for the subsequent motor performance (motor-preparatory brain activity) (Cooke et al., 2015), and motor performance accuracy is affected by the motor-preparatory brain activity that precedes it (Dyke et al., 2014). Given the interrelationship between performance and feedback-related as well as motor-preparatory brain activity, it is important to understand how motor skill practice conditions affect these neural processes.

Manipulating practice conditions, instructions, or feedback delivery to provide learners with an increased perception of success on the task is a strategy with the potential to affect feedback-related and motor-preparatory brain activity and, consequently, motor learning and performance. One way to give learners an increased perception of success is by providing them with easier criteria for success. Consider the following practical example: a golfer is practicing tee shots. Their instructor wants them to drive the ball to the center of the fairway and informs them that only such shots will be considered a success. If the golfer hits a shot on the fairway but off center, then they will likely process the outcome of the shot as a failure and reprogram their next swing in attempt to hit the ball in the center of the fairway. Conversely, if the instructor informs them that any shot on the fairway will be considered a success, then a shot on the fairway but off center will likely be processed as a success and the golfer will not reprogram their next swing to a large degree. In that example, even though the goal of the task, its level of difficulty, and the error magnitude, represented by deviation of the ball from the center of the fairway, were unchanged in the two scenarios, the golfer's perception of success would likely alter feedback processing and subsequent movement preparation. Importantly, these alterations would result from a simple manipulation of the task instruction given by the instructor. The effects of criteria of success on performance and learning have been studied at the behavioral and psychological level (e.g., Parma et al., 2023). However, despite the potential effects of the manipulation of perceptions of success on feedback-related and motor-preparatory brain activity, to the best of our knowledge no study has investigated these effects nor how they interact with error magnitude and task experience.

At the neurophysiological level, feedback processing can be investigated using the reward positivity (RewP), an event-related potential (ERP) component of the electroencephalographic (EEG) signal. RewP is characterized by a positive deflection in the EEG signal that occurs 230 ms to 350 ms after feedback onset and is maximal at fronto-central electrodes (Sambrook & Goslin, 2015). RewP is suggested to reflect the activation of the midbrain reward circuit (Proudfit, 2015) and is a proxy for positive reward-prediction error (RPE; Holroyd & Coles, 2002). Given that RPE is the difference between the predicted and the actual outcome (Lohse et al., 2019), RewP amplitude is expected to be positively correlated with feedback valence (positive vs. negative) and magnitude (large vs. small), and inversely correlated with outcome likelihood. As such, feedback about more positive outcomes and/or more surprising positive outcomes is predicted to elicit larger RewPs than feedback about poorer and/or less surprising outcomes (Margraf et al., 2022).

These predictions have been confirmed in the literature. Regarding valence, Meadows et al. (2016) revealed that participants exhibited larger RewPs after receiving positive feedback than negative feedback during a response time task. Concerning magnitude, Frömer et al. (2016) had participants perform a virtual throwing task and demonstrated that, among successful (on-target) trials, throws closer to the center of the target resulted in larger RewP amplitudes than those farther from the center of the target. RewP responsiveness to error magnitude, however, seems to depend on feedback valence, such that the effect of magnitude is conditional on positive valence. For example, Meadows et al. (2016) demonstrated that RewP was responsive to feedback magnitude only when its valence was positive, but not negative, highlighting an interaction between these factors.

The effect of outcome expectation on RewP has been confirmed both through the observation of different levels of expertise and manipulations of task difficulty. Regarding expertise, Frömer et al. (2016) demonstrated that, throughout practice, the higher the participant's on-target frequency, the smaller the RewP, presumably because the participant starts to expect positive outcomes. Similarly, Williams et al. (2018) showed smaller RewPs at the end of the acquisition phase of a cognitive task, as compared to the beginning of practice. The effect of task difficulty was addressed in Williams et al. (2017), in which the same participants engaged in a time-estimation task under different conditions. In some of the conditions, participants' responses were only considered correct if they fell within a narrow time window surrounding the target time, whereas in other conditions the responses were considered correct if they fell within a wider time window. In the easier conditions (wider time-windows), when participants had correct responses, smaller RewPs were observed in comparison to correct trials in harder conditions (narrower time windows), likely due to participants' lower expectations for success in the latter.

Although Williams et al. (2017) demonstrated that establishing different criteria of success can alter feedback processing, feedback was only provided in a qualitative binary way (correct vs. incorrect) and participants were not informed about their quantitative error (deviation from the target time), confounding the effects of error magnitude and feedback valence. Interestingly, in Wilhelm et al. (2019), the perception of cognitive task difficulty, rather than task difficulty itself, affected the RewP. In this study, participants observed fictitious data regarding the performance of other people, leading them to believe that some blocks of trials were more difficult than others. During blocks participants believed to be difficult, positive feedback elicited larger RewPs than positive feedback during blocks considered to be easy, while perceived difficulty did not affect RewP on trials with negative feedback. These results suggest that the mere *perception* of success by the performer can affect feedback processing, and indicate that, like error magnitude, expectations for success only modulate the RewP for positive outcomes.

Motor-preparatory brain activity, in turn, can be inferred from power in the upper-alpha frequency bandwidth (10 – 13 Hz) recorded at electrodes overlying frontocentral and central scalp locations in the seconds preceding movement. Upper-alpha power reflects neuronal inhibition, so increased neural activity manifests as decreased power (Babiloni et al., 2008). Therefore, when cortical resources are dedicated to motor programming, upper-alpha power decreases. For example, Daou et al. (2018) showed that upper-alpha power progressively decreased at frontocentral and central electrodes during the seconds preceding the backswing of a golf putt. Cooke et al. (2015) demonstrated that this effect is moderated by performance on the previous trial, such that this decrease in pre-movement upper-alpha power was larger after missed putts relative to made putts, indicating that more resources were allocated to movement preparation following errors.

Interestingly, Cooke et al. (2014) showed that expertise also affects alpha power. In their experiment, experts exhibited higher pre-movement upper-alpha power than novices, likely because the former executed the putts with greater automaticity, resulting in a lower demand for cortical resources during motor programming. This effect is consistent with del Percio et al. (2010), which showed higher alpha power over motor areas during the pre-movement and movement periods for elite athletes compared to non-athletes, suggesting that athletes have greater neural efficiency. Together, these studies suggest that motor upper-alpha power reflects neural resources allocated to movement preparation and is modulated by performance outcome and expertise. However, to the best of our knowledge, no study has investigated how perceptions of success may influence motor upper-alpha power.

Since motor learning and performance are influenced by feedback-related and motor-preparatory brain activity, understanding how they are affected by a simple task instruction that affects learners' perceptions of success is important. Therefore, we recorded participants' EEG while they practiced a motor skill under an easy or hard criterion of success. Specifically, we used mixed-effects models to analyze, on a trial-by-trial basis, how RewP and motor upper-alpha changed according to: (1) whether the outcome was considered good or bad based on the participant's criterion of success; (2) error magnitude; (3) experience (trial number), and (4) the interaction of these factors.

Regarding RewP amplitude, based on previous literature, we expected a main effect of success, such that trials within the participant's zone of success would result in larger RewPs than trials outside of this zone; a main effect of error, such that larger errors would result in smaller RewPs; and a main effect of trial, such that RewP would become smaller with increased trial number, since experience should make movement outcomes less surprising. However, we also predicted interactions between these factors. Based on the idea that RewP amplitude is expected to be more responsive to error magnitude on successful trials than on unsuccessful trials, we predicted a success by error interaction. Specifically, successful trials should result in a stronger relationship between RewP amplitude and error than unsuccessful trials, since, for a small error, RewP should be large only if it falls within the participant`s zone of success; conversely, if a trial is considered unsuccessful, any error, even a small one, should result is a small RewP. We also predicted an error by trial interaction and a success by trial interaction, such that later trials would result in a weaker relationship between error and RewP as well as success and RewP. This follows because successful trials and smaller errors should become more frequent and expected, resulting in smaller RewPs for a given outcome. Finally, we hypothesized a success by error by trial interaction, such that, in later trials, the relationship between error and RewP would be weaker, but this weakening should be more prominent for successful trials versus unsuccessful ones. In other words, unsuccessful trials were expected to have small RewPs for all magnitudes of error and across all trials. However, for successful trials, a large difference in RewP amplitude was expected between trials with smaller and larger errors, but mostly in the beginning of practice. With habituation brought by experience, a decrease in RewP amplitude was expected for small errors, making the elicited RewP amplitude less distinguishable from that elicited by larger errors.

Regarding motor upper-alpha power on the current trial, we expected a main effect of success on the prior trial, such that alpha would be higher when the prior trial was successful, since the learner would not feel compelled to allocate more neural resources to motor programming. Similarly, a main effect of error on the prior trial was expected, such that larger errors would result in more resources allocated to motor programming and, consequently, less alpha power. Finally, we also predicted a main effect of trial, such that greater alpha power was expected with increased trial number, due to acquired automaticity. Although we had no precedent results in the literature to predict interactions among prior trial success, prior trial error, or trial number, interactions between these variables were tested in exploratory analyses.

Methods

Participants

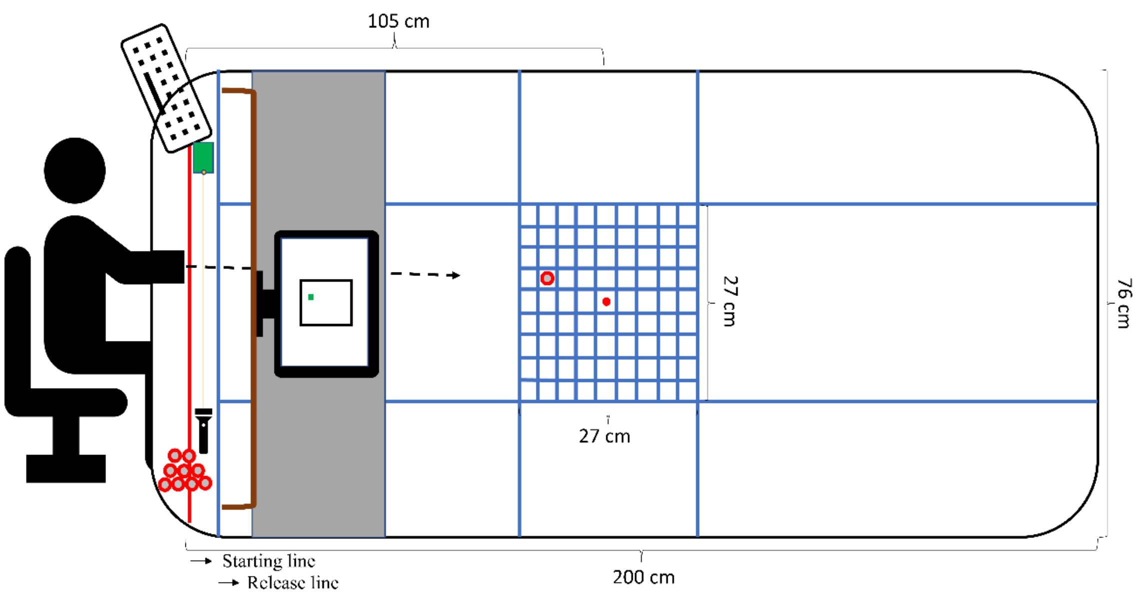
Fifty-two individuals (*M* age = 21.52 years, *SD* = 2.70, 33 males) who were part of a larger project were included in this study, but two were excluded due to poor EEG recording. Participants were healthy undergraduate and graduate students with the preference to throw with their right hand, between the ages of 19 and 40 years, novices to the task, and persons who reported not being allergic to conductive gel, colorblind, at high-risk for serious complications from Covid-19 infection, or having physical impairments precluding comfortable left-arm movements from a seated position. The study was approved by the Auburn University Institutional Review Board (Protocol #19-046 EP 1902) and was conducted in agreement with the 1964 Declaration of Helsinki. Written informed consent was provided by all participants prior to the beginning of the experiment.

Task

Participants performed a mini-shuffleboard learning task using 10 mini-shuffleboard pucks, 1 at a time (as shown in Figure 1). The pucks consisted of a 1.5 cm diameter red plastic ring encompassing a metal sphere. Participants slid the pucks lengthwise on a rectangular table (213 cm long x 76 cm wide) covered in low-friction adhesive paper with auto-adhesive foam tape (1 cm x 1 cm x 1 cm) lining the side and back edges to prevent the pucks from falling off the table. The start line was drawn with a red marker 13 cm from the edge of the table and followed 10 cm farther by the release line, drawn in blue. A 27 cm x 27 cm grid divided in 81 squares of 3 cm x 3 cm was drawn in blue with its center 95 cm past the release line. The grid was connected to the edges of the table by two parallel horizontal lines and two parallel vertical lines that were extensions of the outer lines of the grid. Participants remained seated throughout the whole experiment in a chair positioned so that their left wrist could comfortably reach the release line. A laptop table supporting a computer screen and an occlusion board was positioned after the release line, restricting the participant's vision to about 25 cm of the puck's trajectory after release. The number pad of a keyboard rested on the left side of the shuffleboard table and was used by the participants to initiate the trial and receive feedback (see Procedures below). A photosensor was positioned on the left side of the table opposite a flashlight, both at approximately 6 cm from the starting line. To perform the task, participants were asked to grip the sides of the puck with their left index finger and thumb and slide the puck under the occlusion board by extending their left arm in a straight line (complete instructions given to participants can be found at https://osf.io/9djrx/, Instructions). Before each shot, participants prepared the puck by positioning it on the start line and were instructed to release it once it reached the release line.

Figure 1

Illustration of Experimental Set-Up



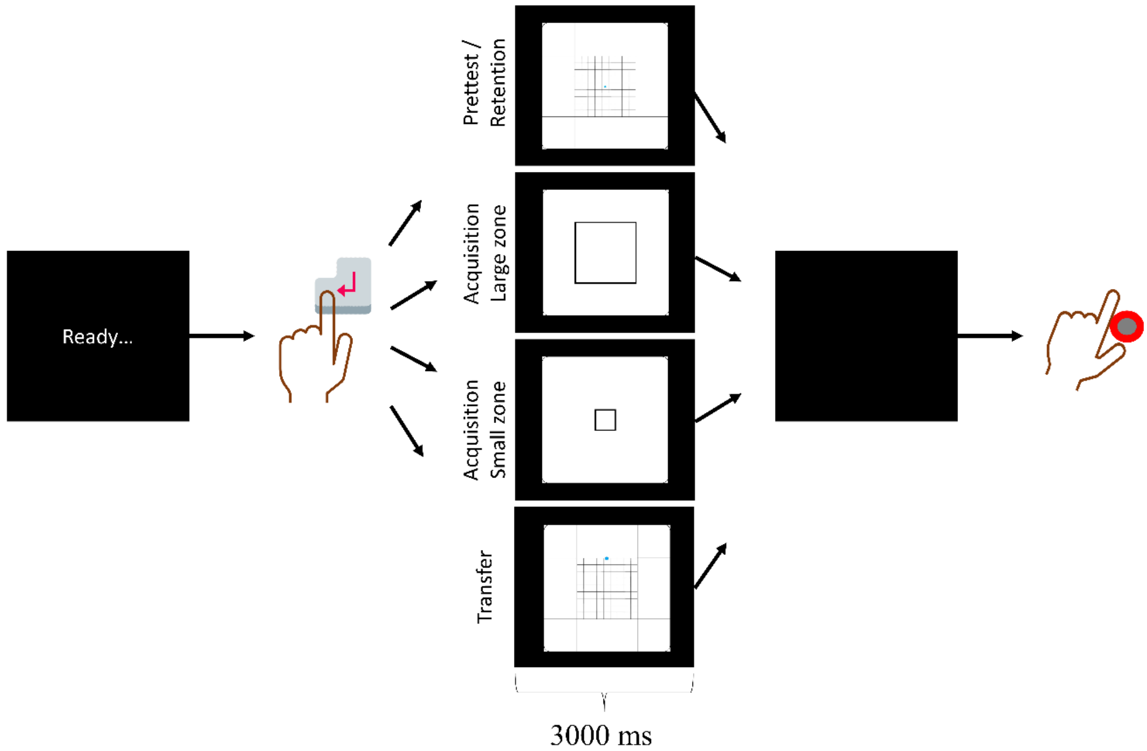
Procedures

Before Acquisition

Before practice, participants read and signed the informed consent and completed the Edinburgh Handedness Inventory (Oldfield, 1971). Participants were then prepared for EEG recording and had their resting brain activity recorded with their eyes open and closed for three minutes each while seated. Participants practiced the task 10 times without augmented feedback after observing the grid and the target for 5 s. To do so, participants were instructed to position the puck when prompted by the word “Ready” from the computer monitor (Figure 2). They then pressed the “enter” key on the keyboard with their left hand, which triggered the appearance of a representation of the grid and target on the screen. The image disappeared from the screen after 3000 ms, at which point participants were allowed to shoot the puck. At this point in the experiment, participants were informed that the goal was to make the puck stop as close to the target as possible. The target was a red dot in the center of the grid.

Figure 2

Trial Initiation



Note. This figure depicts the moments before the participant was allowed to start a new puck shot. Once the “Ready…” image was shown on the screen by the experimenter, the participant was able to press “enter” on the keyboard positioned to their left and returned their hand to the puck. A representation of the grid with the target (represented as a blue dot) in the middle (during pretest and retention test) or beyond (during transfer test), or the participant's assigned target zone (large or small during acquisition) then appeared on the screen for 3000 ms. The participant was allowed to start the movement any time after the image disappeared from the screen.

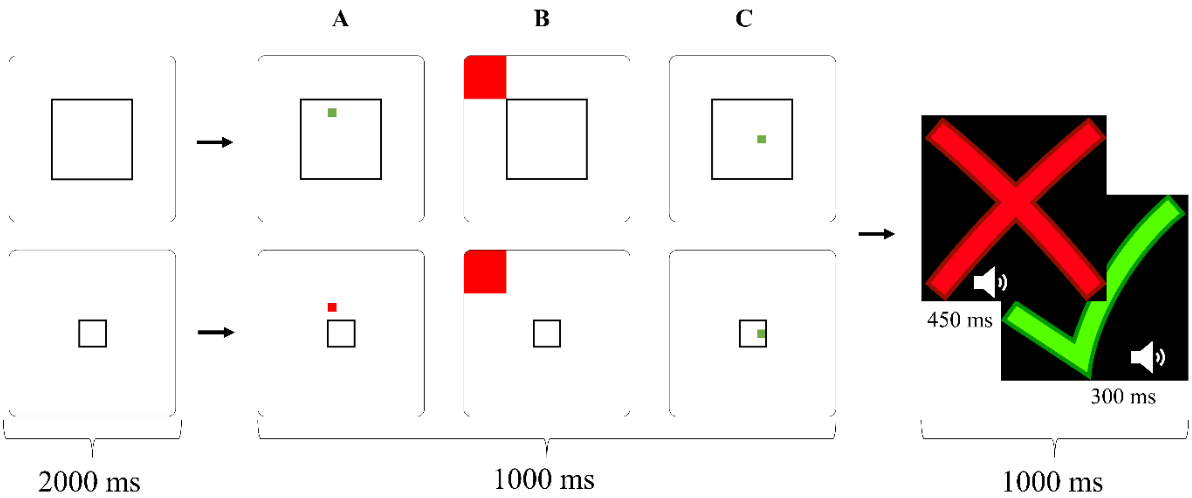
Acquisition

Each participant was randomly assigned to one of two groups (large zone, *n* = 28; small zone, *n* = 22) based on a pre-determined list stratified by gender. Participants performed 10 blocks of 10 trials, before which they were informed there was a target zone they were trying to achieve, such that trials where the puck stopped inside this zone would be considered "good". For participants in the small zone group, the zone was a 9 cm x 9 cm square centered on the target in the middle of the grid, while for participants in the large zone group, the zone was a 27 cm x 27 cm square centered on the target. Zone sizes and number of trials were defined after two pilot studies conducted before the start of this project, described in the supplementary material at the OSF repository, (https://osf.io/9djrx/, Supplementary Material).

Before the first trial of acquisition, participants were shown the zone surrounding the target overlayed on the grid. Specifically, a green cardboard square the size of the participant's assigned zone was positioned on the grid, and the occlusion board was removed for approximately 5 s. During acquisition, participants performed the task as in the 10 trials with no augmented feedback, with a couple of exceptions. First, the image they saw on the computer screen before shooting was a square outlined in black representing the zone, instead of the image of the grid and target (Figure 2). Second, augmented feedback was provided after every trial. To receive feedback, participants pressed "Enter" on the keyboard when prompted by the word "Ready" on the computer screen. Then, the representation of the zone appeared on the computer screen for 2000 ms (Figure 3). Next, the square in the grid where the puck stopped was highlighted in green or red for 1000 ms, depending on whether the puck landed within the zone (a good shot) or not (a bad shot), respectively. That is, participants in both groups were shown the square in the grid where the puck landed on every trial, unless it landed outside the grid. In this case, a rectangle was highlighted in red to indicate the shot was far-left, far-center, far-right, left, right, short-left, short-center, or short-right (see an example in Figure 3 column B). Subsequently, participants saw a green checkmark for 1000 ms and heard a "correct" sound (sound length ≈300 ms), or a red "X" for 1000 ms with an "invalid" sound (sound length ≈450 ms), depending on whether the trial was good or not (all stimuli, including sound files can found at https://osf.io/9djrx/, Stimuli). Participants had a 1-min break between blocks. By the end of the acquisition phase, participants with the large zone had an average of 47.75 (± 11.32) good trials, while participants with the small zone had 10.05 (± 3.00) good trials.

Figure 3

Feedback Presentation



Note. This figure depicts the moments after participants pressed enter on the keyboard to receive feedback. First, an image depicting their assigned target zone was presented for 2000 ms, and then the square where the puck stopped was highlighted in green or red for 1000 ms. Column A represents the same outcome being considered positive for the large zone groups but negative for the small zone groups, while column B represents a negative outcome for all groups and column C depicts an outcome considered positive for all groups. Lastly, participants were presented with a red cross (1000 ms) and an invalid sound (≈450 ms) if the trial was considered negative, or a green checkmark (1000 ms) and correct sound (≈ 300 ms) if the trial was considered successful.

EEG Recording

EEG was recorded during the acquisition phase with 19 channels of an EEG cap housing a 64 channel BrainVision actiCAP system (Brain Products GmbH, Munich, Germany) labeled in accordance with an extended international 10–20 system (Oostenveld & Praamstra, 2001). EEG data were sampled at 250 Hz. EEG data were online referenced to the left earlobe, and the FPz electrode site was employed as the ground electrode. Electrode impedances were set below 25 kΩ before recording started and a high-pass filter was set at 0.016 Hz. The EEG signal was amplified and digitized with a BrainAmp DC amplifier (Brain Products GmbH) linked to BrainVision Recorder software (Brain Products GmbH).

Several events were marked every trial in the EEG signal, including the moment the image of the zone of success appeared or disappeared from the computer screen prior to trial initiation (see Figure 2). Since we aimed to investigate motor-preparatory brain activity, we also marked movement onset via the photosensor. The photosensor was connected to a BrainVision StimTrak device (BrainProducts GmbH) and was triggered every time a participant broke the light beam shining on it. Since participants broke the light beam multiple times each trial (e.g., when they pushed their hand forward and when they pulled their hand backward), we defined movement onset as the first photosensor marker after the disappearance of the zone of success image from the screen (see Figure 2). Finally, to investigate feedback processing, we marked the onset of feedback presentation (column A/B/C in Figure 3).

Data Processing

EEG Processing

Resting Data and Individualized Alpha Frequency. Resting data with eyes closed were first visually inspected to determine whether any electrode needed to be interpolated, and data were re-referenced to an averaged ears montage. Then, a 0.1 – 40 Hz band-pass filter with 4th order roll-offs and a 60 Hz notch filter was applied to the data. Next, data were segmented in 1-s epochs with 0.5 s overlaps, and epochs containing any of the following in the midline electrodes (Fz, FCz, Cz, CPz, or Pz) were removed: change of more than 50 µV from one data point to the next, a change of 100 µV within a moving 200-ms window, or a change of less than 0.5 µV within a moving 200-ms window. Then, data were fast Fourier transformed with a 25% Hanning window and 0.977 Hz bin resolution. Next, data were averaged across segments for each electrode. Then, individual alpha frequency (IAF) was identified to account for individual differences in alpha frequency (Klimesch, 1996). Specifically, in the resting brain data with eyes closed, the spectral peak within the alpha bandwidth (8-13 Hz) at electrode Pz was used to determine the IAF, since alpha tends to peak at posterior electrodes.

Acquisition Data Cleaning. All EEG data processing was conducted with BrainVision Analyzer 2.2 software (BrainProducts GmbH). EEG data were first visually inspected to determine whether any electrode needed to be interpolated, and data were re-referenced to an averaged ears montage. Then, data were prepared for independent component analysis (ICA) cleaning. First, a 1 – 40 Hz band-pass filter with 4th order roll-offs and a 60 Hz notch filter was applied. Next, data from blocks 4 to 6 were visually inspected and non-stereotypical artifacts marked. Then, an ICA was conducted on the same blocks to identify stereotypical artifacts, such as blinks and saccades. The stereotypical artifacts identified by the ICA were then removed from all blocks of the unfiltered data.

RewP Processing. To assess feedback-related brain activity, we extracted single-trial RewP amplitude. Cleaned data were band-passed filtered between 0.1 and 30 Hz with 4th order roll-offs, and a 60 Hz notch filter was applied. Next, to define the individualized RewP time window for each participant, data were segmented from 200 ms prior to 800 ms after feedback onset. Then, these epochs were baseline corrected based on the pre-stimulus interval (-200 – 0 ms). Next, epochs containing any of the following in the midline electrodes (Fz, FCz, Cz, CPz, or Pz) were removed: change of more than 50 μV from one data point to the next, a change of 100 μV within a moving 200-ms window, or a change of less than 0.5 μV within a moving 200-ms window. Then, epochs time-locked to positive feedback (trials stopping within the zone of success) and negative feedback (trials stopping outside of the zone of success) were separately averaged. Then, the average of the negative feedback epochs was subtracted from the average of the positive feedback epochs to create a difference wave for each participant. We centered each participant's RewP time window (length = 40 ms) on their peak latency between 230 and 350 ms at the frontocentral electrode (Fz, FCz, or Cz) at which it peaked (Parma et al., in press). We also confirmed that this individualized window included a negative deflection in the negative feedback waveform. If it did not, we centered the window on the maximal negativity between 230 and 350 ms in the negative feedback waveform (Parma et al., in press). Then, we computed mean amplitude in each participant's time window at Fz, FCz, and Cz for each epoch and then averaged across these electrodes, yielding one RewP for each trial. If Fz, FCz, or Cz malfunctioned during recording, it was not included in the average.

Motor Upper-Alpha Processing. To assess motor-preparatory brain activity, we extracted single-trial motor upper-alpha power. Specifically, cleaned data were band-passed filtered between 0.1 and 40 Hz with 4^{th} order roll-offs, and a 60 Hz notch filter was applied. Next, for each trial, four epochs were created within a motor preparatory time window: 1) 3-s to 2-s prior to movement onset, 2) 2-s to 1-s prior to movement onset, 3) 1-s prior to movement onset to movement onset, and 4) movement onset to 1-s after movement onset (Daou et al., 2018). Then, epochs containing any of the following in the midline electrodes (Fz, FCz, Cz, CPz, or Pz) were removed: change of more than 50 µV from one data point to the next, a change of 100 µV within a moving 200-ms window, or a change of less than 0.5 µV within a moving 200-ms window. Next, a fast Fourier transformation was employed using 0.977 Hz bins and a Hanning window (50% taper). Spectral power was averaged from IAF to IAF + 2 Hz (Wang et al., 2020), and then averaged across the right frontocentral electrodes, yielding one motor upper-alpha power value for each of the four epochs within each trial. If one of the frontocentral electrodes malfunctioned during recording, it was not included in the average.

Motor Performance

Our main outcome measure of performance was radial error (radial error = $(x^2 + y^2)^{1/2}$, representing accuracy (Hancock et al., 1995). To extract the x and y coordinates of the puck's stopping position on the table, an iPad was affixed to the ceiling above the table, and photographs were taken after each trial with a wireless clicker. These photographs were then analyzed with LabView® software using the virtual instrument ScorePutting (Neumann & Thomas, 2008) to determine the distance between the center of the puck and the target along the x and y axis and use these distances to calculate radial error.

Statistical Analysis

For all inferential analyses, alpha was set to .05. All statistical analyses were conducted in R (cran.r-project.org), and all models used can be found in Appendix A and B.

Prior to statistical analyses, we visually inspected density plots for single-trial RewP, motor upper-alpha, and radial error. We also inspected scatter and spaghetti plots for the effects of radial error, success, trial, epoch, and their interactions on RewP amplitude and/or motor upper-alpha power. Before running mixed-effects regressions, we created the prior success and prior radial error variables by pairing trials with the values of success and radial error of the previous trial. Then, we contrast coded all categorical variables (success, prior success, and zone assignment ([small or large]), participant-mean centered radial error and prior radial error, and linear contrast coded trial. After centering these variables, we created the quadratic and cubic trial variables (i.e., $trial^2$ and $trial^3$, respectively), the quadratic and cubic radial error variables (i.e., radial $error^2$ and radial $error^3$, respectively), as well as the quadratic and cubic prior radial error variables (i.e., prior radial $error^2$ and prior radial $error^3$, respectively).

Feedback Processing Models

Before running models, we excluded trials with missing data on the single-trial RewP (due to artifact rejection) and/or radial error (due to missed photographs of the puck's landing position), which led to the loss of 78 of 5000 trials (1.56% of the data).

Then, before building the main models related to feedback processing, we compared models fitted with maximum likelihood to explore whether the relationship between RewP amplitude and radial error as well as RewP amplitude and trial number should be modelled in a linear, quadratic, or cubic fashion. For both radial error and trial number, the models with quadratic terms to both the fixed- and random-effects were considered better fits to the data than linear or cubic models, or models with quadratic terms only on the fixed effects. Specifically, the models reduced AIC by ≥ 2 points in comparison to simpler models, whereas the models with cubic terms did not reduce AIC by ≥ 2 points in comparison to the models with quadratic terms. Lohse et al. (2020) used this approach to employ AIC for model selection, approximating Burnham and Anderson (2002)'s method that used effect sizes for model selection.

Regarding main models, to analyze how feedback processing changed during practice according to perception of success, error magnitude, and experience, we used three separate mixed-effects models with single-trial RewP amplitude as their dependent variable. The first model was fitted with maximum likelihood and included fixed effects of success (in/out of the zone of success), trial number (1 – 100), its quadratic term, and their interactions, and included a random intercept of participant, and random slopes of success, trial number, and its quadratic term.

The second model was fitted with maximum likelihood and included fixed effects of radial error, its quadratic term, trial number, its quadratic term, and their interactions, and included a random intercept of participant, and random slopes of radial error, its quadratic term, trial number, and its quadratic term.

We then used Wald likelihood ratio tests to assess the change in deviance between these models and a reference model that included only trial number and its quadratic term as fixed-effects, and random intercept of participant with random slopes of trial and its quadratic term to determine whether the additional RewP variance explained by success was different from the additional RewP variance explained by radial error. Complementary, to investigate which one (success or radial error) explained more of RewP's variance, we also estimated the effect size of the Reference Model and Models 1 and 2 by using the conditional pseudo-R^2 computed using the MuMin package (Barton, 2018). These strategies were used to compare Models 1 and 2 given that they have different levels of complexity. Model 2 includes more terms, due to the quadratic term of radial error, making it more penalized by AIC, if that metric was used.

Finally, we ran a third model to assess whether the learner's perception of success could modulate the effect of objective success (as indexed by radial error) on the RewP. Initially, we tried to include success and radial error in the same model to investigate their possible interaction. However, because of the presence of high multicollinearity between these variables, we ended up removing the variable success, and replacing it with the between-subjects zone variable (large group vs. small group). Therefore, for the final model, fitted with restricted maximum likelihood, we included fixed effects of zone, radial error, its quadratic term, trial number, its quadratic term, and their interactions, and included a random intercept of participant, and random slopes of radial error, its quadratic term, trial number, and its quadratic term.

Movement Preparation Models

Before running models, we excluded 8 participants for whom photosensor data were not recorded. For the remaining participants, we excluded the first trial because prior radial error and prior success (predictors in the main models) would be undefined for this trial. Next, we excluded epochs with missing data on prior radial error (due to missed photographs of the puck's landing position) and/or motor upper-alpha power (due to artifact rejection or photosensor malfunction). These exclusions led to the loss of 1,791 of 16,800 total epochs (10.23% of the data).

Then, before building the main models related to movement preparation, we compared models fitted with maximum likelihood to explore whether the relationship between motor upper-alpha power and prior radial error as well as motor upper-alpha power and trial number should be modelled in a linear, quadratic, or cubic fashion. For trial number, the model with quadratic terms to both the fixed- and random-effects was considered a better fit to the data (based on the 2 points criterion) than linear or cubic models, or than the model with the quadratic term only as a fixed- effect. For prior radial error, the model with cubic terms to both the fixed- and random-effects was considered a better fit to the data than the models with linear or quadratic terms, or the model with the cubic term only as a fixed-effect.

Regarding main models, to analyze how movement preparation changed during practice according to perception of success on the previous trial, error magnitude on the previous trial, and experience, we used three separate mixed-effects models with single-trial motor upper-alpha power as their dependent variable. The first model was fitted with maximum likelihood and included fixed effects of prior success (previous trial in/out of the zone of success), epoch (1: -3 s to -2 s, 2: -2 s to -1 s, 3: -1 s to 0 s, 4: 0 s to +1 s), trial number (1 – 100), its quadratic term, and their interactions, and included a random intercept of participant, with a random slope of trial number, and a random intercept of participant by epoch[1].

The second model was fitted with maximum likelihood and included fixed effects of prior radial error, its quadratic and cubic terms, trial number, its quadratic term, epoch, and their interactions, and included a random intercept of participant, with a random slope of trial number, its quadratic term, prior radial error, and its quadratic term, and a random intercept of participant by epoch[2].

We then used Wald likelihood ratio tests to assess the change in deviance between these models and a reference model that included only epoch, trial number and its quadratic term as fixed-effects, random intercept of participant with random slopes of trial and its quadratic term, and random intercept of participant by trial, to determine whether the

[1] Initially our model also let the slope of trial^2 to vary across participant and included a random intercept of participant by prior success. However, boundary warnings revealed that trial^2 was highly correlated to trial in the random-effects, and that there was minimal variance for the effect of prior success across participant. Thus, we dropped these terms from our random-effects.

[2] Initially our model also let the slope of prior radial error^3 to vary across participant. However, due to the lack of convergence, that revealed that prior radial error^3 was perfectly correlated to prior radial error^2 in the random-effects, we dropped the cubic term from our random-effects.

additional alpha variance explained by prior success was different from the additional alpha variance explained by prior radial error. Complementary, to investigate which one (prior success or prior radial error) explained more of RewP's variance, we also estimated the effect size of the Reference Model and Models 1 and 2 by using the conditional pseudo-R^2 computed using the MuMin package (Barton, 2018).

Finally, we ran a third model to assess whether the learner's perception of success in the previous trial could modulate the effect of objective success in the previous trial (as indexed by prior radial error) on alpha power. Initially, we tried to include prior success and prior radial error in the same model to investigate their possible interaction. However, because of the presence of high multicollinearity between these variables, we ended up removing the variable prior success, and replacing it with the between-subjects zone variable (large group vs. small group). Therefore, for the final model, fitted with restricted maximum likelihood, we included fixed effects of zone, prior radial error, its quadratic and cubic terms, trial number, its quadratic term, epoch, and their interactions, and included random intercepts of participant, with random slopes of prior radial error, trial number, and its quadratic term, and random intercepts of participant by epoch [3].

Results

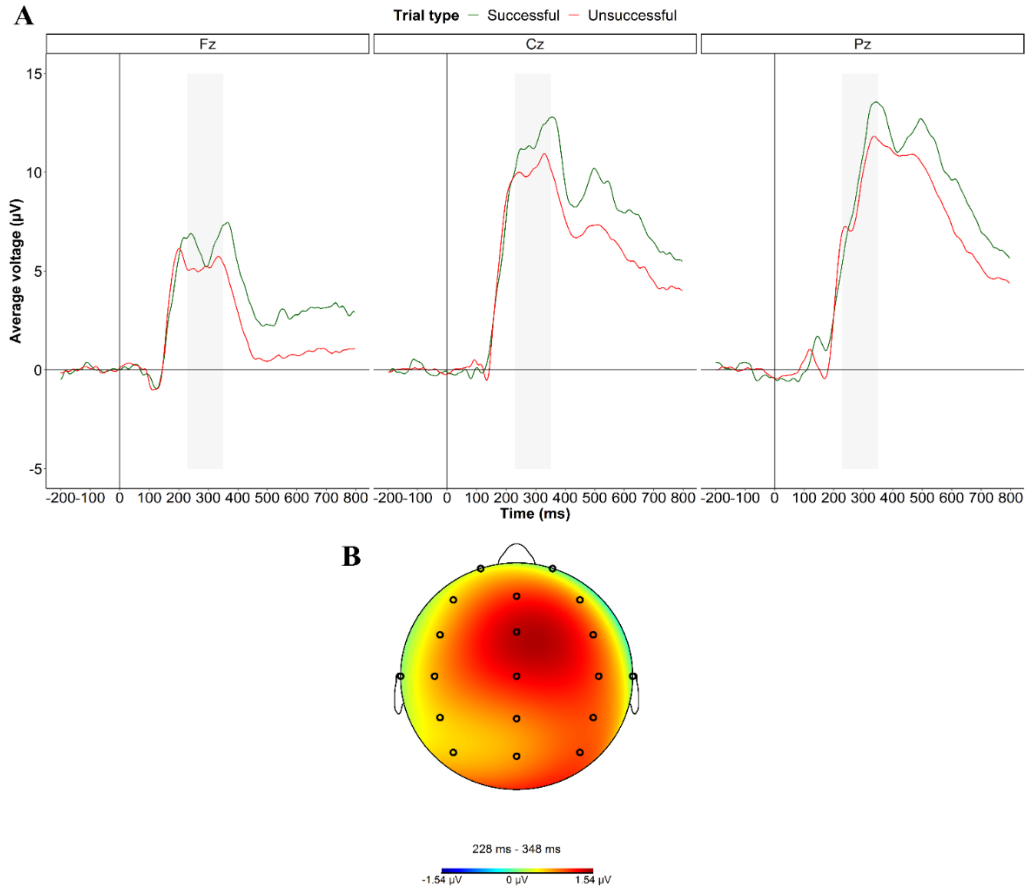
Feedback processing

Figure 4A depicts the grand average ERPs at electrodes Fz, Cz, and Pz for successful and unsuccessful trials. Figure 4B depicts the scalp topography of the grand average successful minus unsuccessful difference waveform during the RewP time window.

[3] Initially our model also let the slope of prior radial error^2 and prior radial error^3 to vary across participant. However, due to the lack of convergence, we dropped the quadratic and cubic terms from our random-effects.

Figure 4

Grand Average RewP Waveforms for Successful and Unsuccessful Trials



Note. A: Grand average waveforms for the RewP time-locked to the onset of augmented feedback (time 0) at electrodes Fz, Cz, and Pz after successful (green line) and unsuccessful (red line) trials, as determined, respectively, by trials that stopped within or outside of the participant's zone of success. Shaded area represents the RewP time window (230ms-350ms). B: Scalp topography of the grand average successful minus unsuccessful difference waveform during the RewP time window.

Model 1 - Effects of Perception of Success on the RewP

Results of the analysis of the effect of perceptions of success on RewP amplitude are presented in Table 1.

Table 1

Random and Fixed Effects for the Analysis of the Effect of Success on the RewP

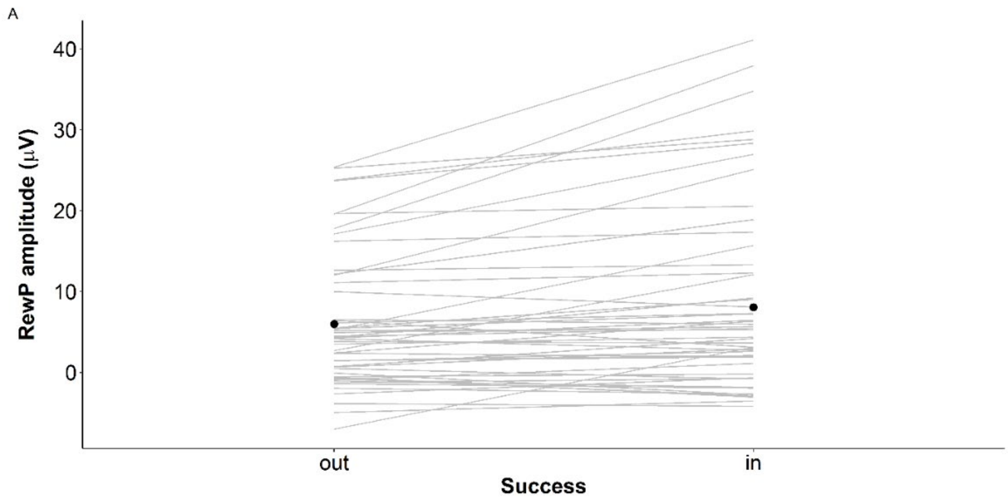
Random Effects					
Group	*Effect*	*SD*	*Corr*		
Participant	Intercept	9.72			
	Success	3.96	.75		
	Trial	0.38	-.28	-.34	
	Trial²	0.10	-.41	-.35	.58
Residual		9.87			
Fixed Effects					
Effects	*β*	*SE*	*df*	*t-value*	*p-value*
Intercept	7.50	1.40	4867	5.36	<.001*
Success	2.86	0.76	4867	3.76	<.001*
Trial	-0.11	0.08	4867	-1.50	.135
Trial²	-0.02	0.03	4867	-0.86	.390
Success:Trial	0.12	0.11	4867	1.00	.315
Success:Trial²	-0.03	0.04	4867	-0.59	.556

Note. Number of observations: 4922, groups: 50. * indicates significant differences. SD = standard deviation. Corr = correlation. SE = standard error. df= degrees of freedom.

The analysis revealed a significant main effect of success ($p < .001$), confirming the impression from Figures 4 and 5 that successful (in) trials resulted in more positive RewPs than unsuccessful (out) trials, even after controlling for trial. No main effect of trial number (ps ≥ .135) or interactions (ps ≥ .315) were found. Figure 5 depicts the relationship between success and RewP. Figure 6 depicts the relationship between trial number and RewP.

Figure 5

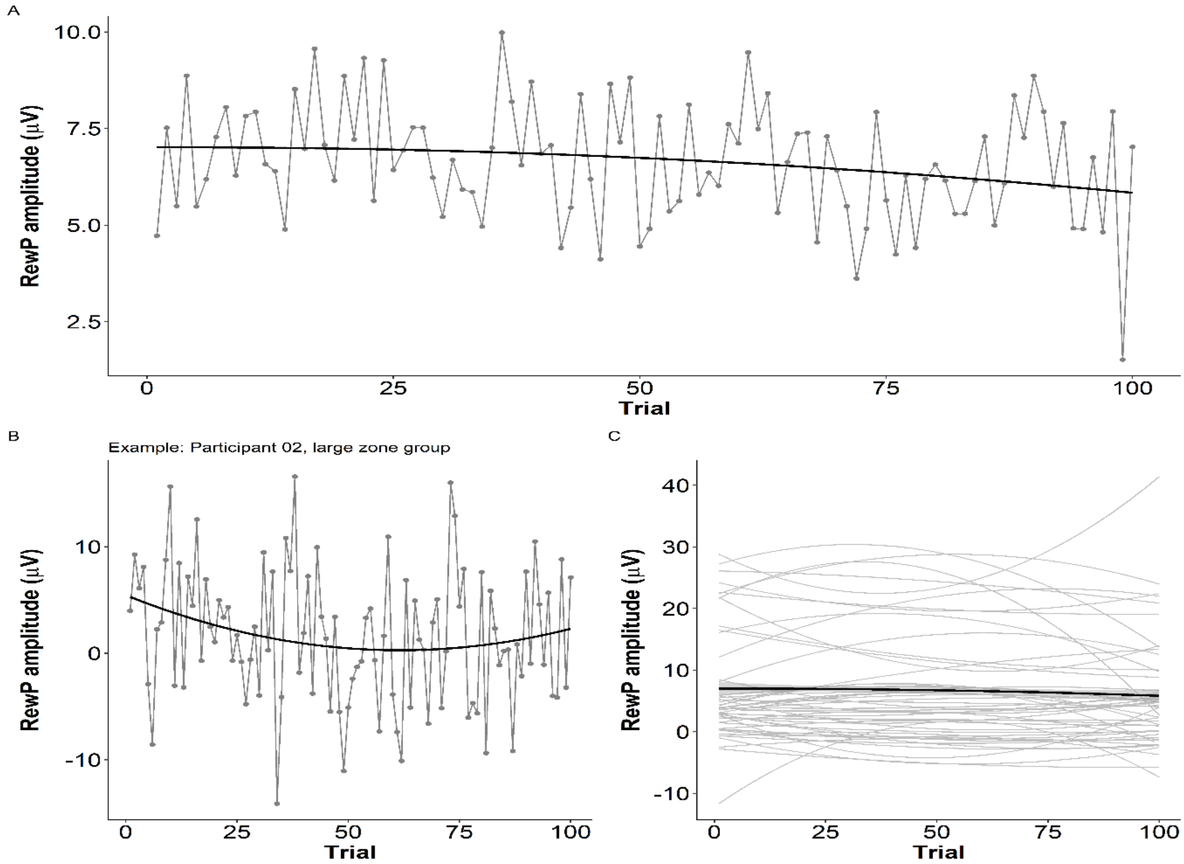
Relationship Between Success and RewP Amplitude



Note. Figure shows the positive relationship between RewP amplitude and trial. The black dots represent the average RewP for the sample for successful (in) and unsuccessful (out) trials, whereas the gray lines represent the intercept and slope for each participant.

Figure 6

Relationship Between Trial Number and RewP Amplitude



Note. A: Figure shows the average RewP across participant for every trial in gray and the quadratic line of best fit in black. B: Figure shows, as an example, the RewP amplitude for every trial of a single, randomly chosen participant (participant 2 of the large zone group). The black line represents a well-fitted quadratic relationship between RewP and trial for this specific participant. C: Figure shows the relationship between RewP amplitude and trial. The black line represents the quadratic relationship between RewP and trial for the sample, whereas the gray lines represent the slope for each participant.

Model 2 - Effects of Error Magnitude on the RewP

Results of the analysis of the effect of error magnitude on RewP amplitude are presented in Table 2.

Table 2

Random and Fixed Effects for the Analysis of the Effect of Radial Error on the RewP

Random Effects						
Group	*Effect*	*SD*	*Corr*			
Participant	Intercept	8.57				
	Radial error	0.15	-.65			
	Radial error^2	0.00	.54	-.94		
	Trial	0.41	-.25	.54	-.49	
	Trial^2	0.11	-.33	.53	-.51	.62
Residual		9.88				

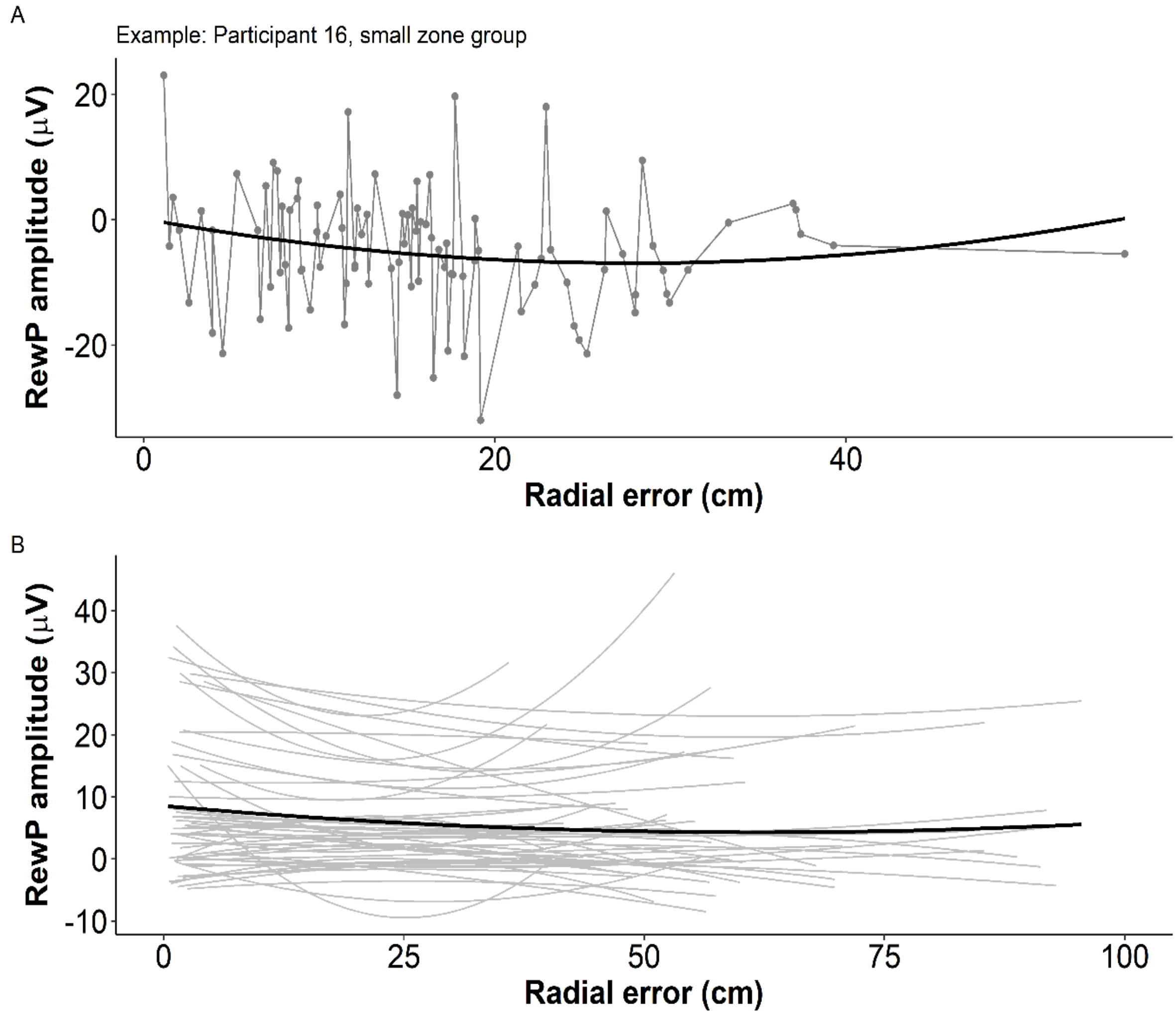
Fixed Effects					
Effects	*β*	*SE*	*df*	*t-value*	*p-value*
Intercept	6.54	1.24	4864	5.29	<.001*
Radial error	-0.10	0.03	4864	-3.27	.001*
Radial error^2	0.00	0.00	4864	1.87	.061
Trial	-0.12	0.08	4846	-1.46	.145
Trial^2	-0.01	0.03	4864	-0.34	.735
Radial error:Trial	-0.01	0.01	4846	-0.93	.355
Radial error:Trial^2	-0.00	0.00	4846	0.36	.718
Radial error^2:Trial	-0.00	0.00	4846	-0.68	.503
Radial error^2:Trial^2	-0.00	0.00	4846	-0.62	.534

Note. Number of observations: 4922, groups: 50. * indicates significant differences. SD = standard deviation. Corr = correlation. SE = standard error. df= degrees of freedom.

The analysis revealed a significant negative main effect of radial error ($p = .001$), such that trials with higher errors (lower accuracy) resulted in smaller RewPs than trials with smaller errors. No main effect of trial number (ps $\geq .145$) or interactions (ps $\geq .355$) were found. Figure 7 depicts the relationship between radial error and RewP.

Figure 7

Relationship Between Radial Error and RewP Amplitude



Note. A: Figure shows, as an example, the RewP amplitude as a function of the radial error of each trial of a single, randomly chosen participant (participant 16 of the small zone group). The black line represents a well-fitted quadratic relationship between RewP and radial error for this specific participant. B: Figure shows the relationship between RewP amplitude and radial error for the sample. The black line represents the quadratic relationship between RewP and radial error for the sample, whereas the gray lines represent the slope for each participant.

Comparing Model 1 and Model 2

The Wald likelihood ratio tests indicated that both models (1 and 2) are an improvement above the reference (trial-only) model, and that Model 1's likelihood ratio is significantly different from Model 2's likelihood ratio (Table 3).

Table 3

Wald Likelihood Ratio Test for the Reference Model, Model 1, and Model 2 on RewP Amplitude

Model	*df*	*AIC*	*logLik*	*Test*	*L.Ratio*	*p-value*
Reference	10	36949.36	-18464.68			
Model 1	17	36849.36	-18407.68	Reference vs. Model 1	114.00	<.001*
Model 2	25	36881.95	-18415.97	Model 1 vs. Model 2	16.59	.035*

Note. * indicates significant differences. df= degrees of freedom. AIC= Akaike Information Criterion. logLik= log-likelihood. L.Ratio= Absolute likelihood-ratio.

The conditional pseudo-R^2 revealed that, overall, Model 1 has a higher effect size than Model 2, suggesting that perception of success explains more of the RewP variance than error magnitude (Table 4).

Table 4

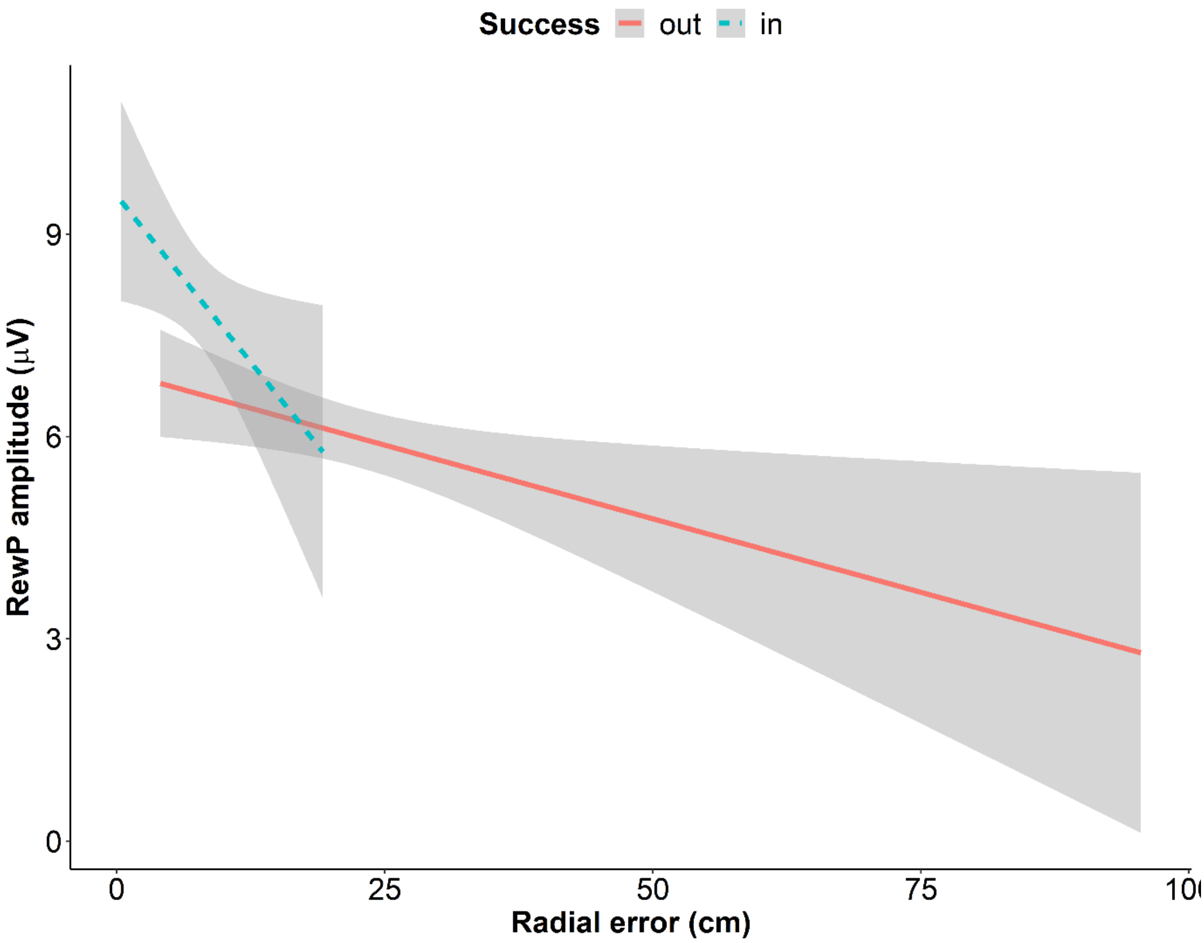
Conditional Pseudo-R^2 Test for the Reference Model, Model 1, and Model 2 on RewP Amplitude

Model	*Conditional R^2*
Reference	.43
Model 1	.47
Model 2	.45

Figure 8 depicts the relationship between success, radial error, and RewP.

Figure 8

Relationship of Success and Radial Error on RewP Amplitude



Note. Figure shows the relationship between RewP amplitude and radial error for successful (in) and unsuccessful (out) trials in traced blue and solid pink lines, respectively. The shaded areas represent 95% confidence intervals.

Model 3 - Effects of Zone of Success Assignment and Error Magnitude on the RewP

Results of the analysis of the effect of assigned zone of success and error magnitude on RewP amplitude are presented in Table 5.

Table 5

Random and Fixed Effects for the Analysis of the Effect of Assigned Zone of Success and Radial Error on the RewP

Random Effects

Group	*Effect*	*SD*	*Corr*			
Participant	Intercept	8.73				
	Radial error	0.15	-.63			
	Radial error^2	0.00	.52	-.93		
	Trial	0.43	-.25	.55	-.50	
	Trial^2	0.12	-.33	.52	-.50	.61
Residual		9.88				

Fixed Effects

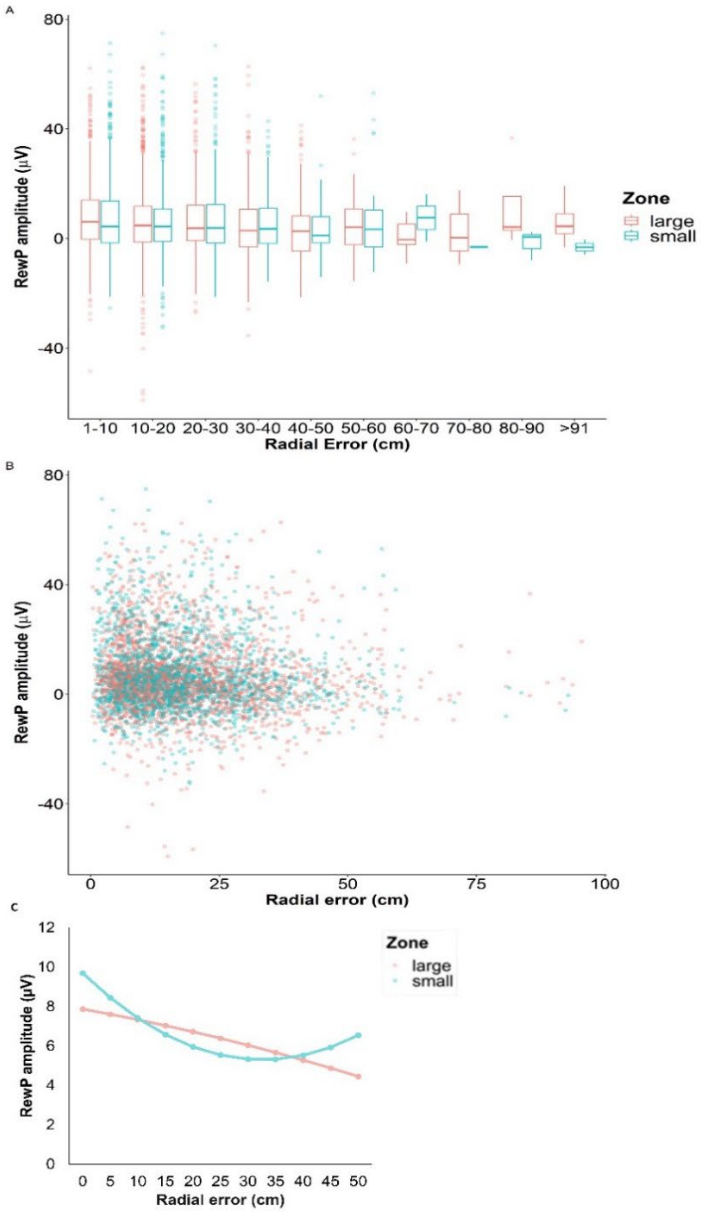
Effects	*β*	*SE*	*DF*	*t-value*	*p-value*
Intercept	6.48	1.27	4856	5.11	<.001*
Radial error	-0.11	0.03	4856	-3.35	.001*
Radial error^2	0.00	0.00	4856	2.24	.025*
Zone	-0.59	2.54	48	-0.23	.818
Trial	-0.11	0.08	4856	-1.33	.183
Trial^2	-0.00	0.03	4856	-0.16	.877
Radial error:Zone	-0.06	0.06	4856	-0.91	.365
Radial error:Trial	-0.00	0.01	4856	-0.79	.433
Radial error:Trial^2	0.00	0.00	4856	0.48	.634
Radial error^2:Zone	0.01	0.00	4856	1.98	.048*
Radial error^2:Trial	-0.00	0.01	4856	-0.83	.409
Radial error^2:Trial^2	-0.00	0.01	4856	-0.91	.361
Zone:Trial	0.03	0.17	4856	0.16	.874
Zone:Trial^2	0.06	0.06	4856	1.07	.284
Zone:Radial error:Trial	0.004	0.011	4856	0.375	.708
Zone:Radial error:Trial^2	0.002	0.004	4856	0.465	.642
Zone:Radial error^2:Trial	-0.0001	0.0004	4856	-0.319	.750
Zone:Radial error^2:Trial^2	-0.0002	0.0002	4856	-1.364	.173

Note. Number of observations: 4922, groups: 50. * indicates significant differences. SD = standard deviation. Corr = correlation. SE = standard error. DF= degrees of freedom.

The analysis revealed a significant negative main effect of radial error ($p = .001$), such that trials with higher errors resulted in smaller RewPs than trials with smaller errors. A positive significant main effect of radial error^2 ($p = .025$) was also found, but it was superseded by a significant interaction of Zone by Radial Error^2 ($p = .048$). The model's estimate suggests that, in comparison to participants with a large zone of success, participants with a small zone of success exhibit larger RewPs for small errors, but smaller RewPs for average to large errors. Participants with a small zone of success exhibit larger RewPs for the largest errors ($\approx > 35$ cm), but it is important to note that relatively few datapoints were used to estimate the RewP for these errors, as shown in Figure 9B, meaning that their effect on RewP should be interpreted with caution. No main effect of zone (p =.818), trial number (ps $\geq .183$), or any other interactions (ps $\geq .173$) were found. Figure 9A depicts the effect of radial error on the RewP for each of the assigned zones of success in our data, while Figure 9B represents the model's predictions about the interaction between zone of success and radial error.

Figure 9

RewP Amplitude as a Function of Radial Error and Zone of Success



Note. A: Figure shows boxplots of RewP amplitude as a function of radial error (in increments of 10 cm) and zone of success. B: Figure shows RewP amplitude as a function of radial error and zone of success. C: Figure plots Model 3's estimates of the interaction between quadratic radial error[4] and zone on RewP amplitude.

---

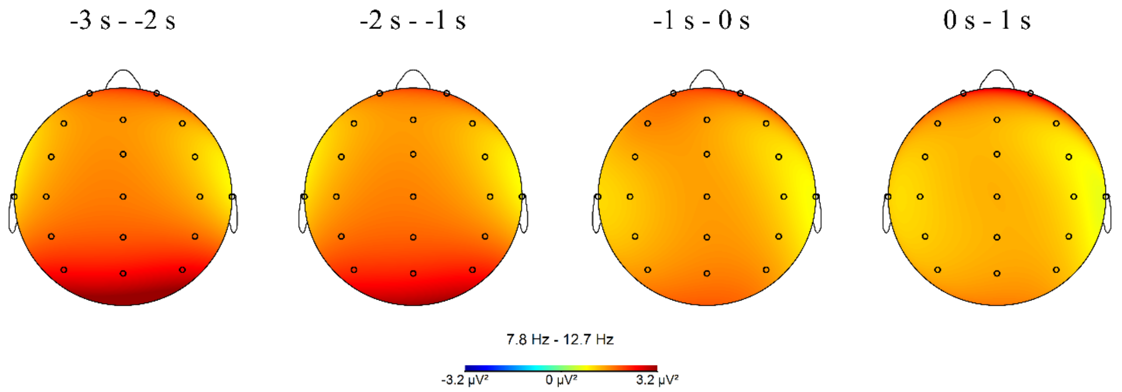
[4] For the sake of better interpretation of the results, we also ran Model 3 with uncentered radial error and radial error^2, and we used the results of this model to plot Figure 9B. Different from the original model, we had to drop Trial^2 from the random effects for the

Movement Preparation

Figure 10 depicts the average scalp topography across participants in the alpha frequency band for all four epochs.

Figure 10

Average Scalp Topography for the Alpha Frequency Band per Epoch



Note. Figure represents the average scalp topography across participants in the alpha frequency band (7.8 Hz – 12.7 Hz) for each of the four epochs (3-s – 2-s prior to movement onset, 2-s – 1-s prior to movement onset, 1-s – 0-s prior to movement onset, and 0-s – 1-s after movement onset). Warmer colors represent higher power.

Model 1 - Effects of Perception of Success on Motor Upper-Alpha

Results of the analysis of the effect of perceptions of success on motor upper-alpha power are presented in Table 6.

Table 6

Model 1 Type III Analysis of Variance Table with Satterthwaite's method

Effects	*Sum Sq*	Mean Sq	*NumDF*	*DenDF*	*F value*	*p-value*
Prior Success	3.43	3.43	1	14921.00	1.88	.171
Trial	11.37	11.37	1	44.6	6.23	.016*
$Trial^2$	3.00	3.00	1	14916.30	1.65	.200

model to converge. Importantly, the reported results in the text refer to the model with centered radial error and centered radial error² and the random effect of Trial².

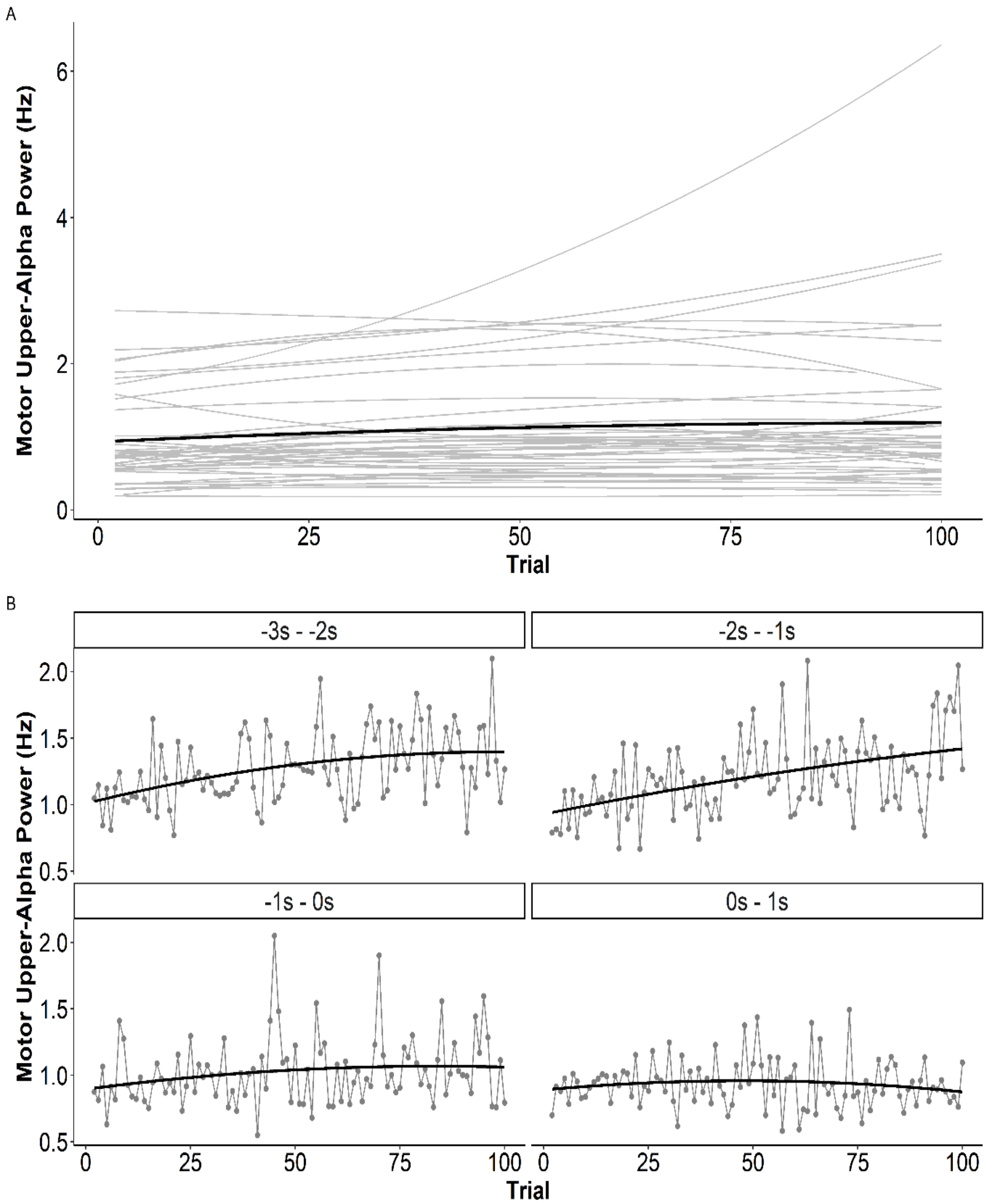
Epoch	31.367	10.46	3	193.90	5.73	<.001*
Prior Success:Trial	5.43	5.43	1	10501.00	2.98	.085
Prior Success:Trial^2	3.24	3.24	1	14920.20	1.77	.183
Prior Success:Epoch	1.74	0.58	3	14823.90	0.32	.813
Trial:Epoch	57.40	19.13	3	14900.40	10.49	<.001*
Trial^2:Epoch	1.71	0.57	3	14930.20	0.31	.817
Prior Success:Trial:Epoch	16.84	5.61	3	14908.90	3.08	.026*
Prior Success:Trial^2:Epoch	4.30	1.43	3	14937.7	0.79	.502

Note. Number of observations: 15081, participants: 42. * indicates significant differences. Sum Sq = sum of squares. Mean Sq = mean square. NumDF = numerator degrees of freedom. DenDF = denominator degrees of freedom.

The analysis revealed a significant main effect of trial (omnibus test $p = .016$), such that more experience in the task resulted in higher motor upper-alpha power. A main effect of epoch was also found, confirming the impressions from Figures 10 and 12 that later epochs resulted in decreased motor upper-alpha power (omnibus test $p < .001$). Figure 12 depicts the effect of epoch on motor upper-alpha power. These effects were superseded by significant interactions of Trial by Epoch (omnibus test $p < .001$), and Prior Success by Trial by Epoch (omnibus test $p = .026$). Figure 13 depicts the two- and three-way interactions and indicates that, over the course of practice trials, motor upper-alpha power increased during motor preparation (epochs 1 – 3), but not motor execution (epoch 4), resulting in the Trial by Epoch interaction. Crucially, this interaction was significantly stronger following successful trials, producing the Prior Success by Trial by Epoch interaction. No other main effects (ps $\geq .171$) or interactions (ps $\geq .085$) were found.

Figure 11

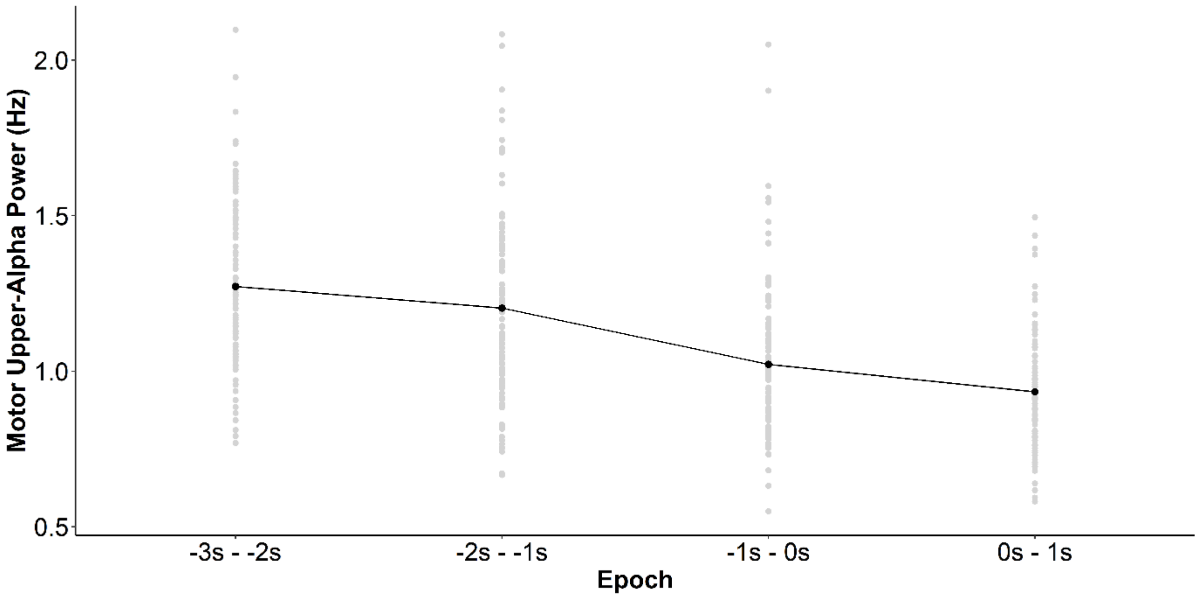
Relationship Between Trial Number and Motor-Upper Alpha Power



Note. A: Figure shows the relationship between motor upper-alpha power and trial. The black line represents the quadratic relationship between motor upper-alpha and trial for the sample, whereas the gray lines represent the slope for each participant. B: Figure shows, for every epoch, the average motor upper-alpha power across participant for every trial in gray and the quadratic line of best fit in black.

Figure 12

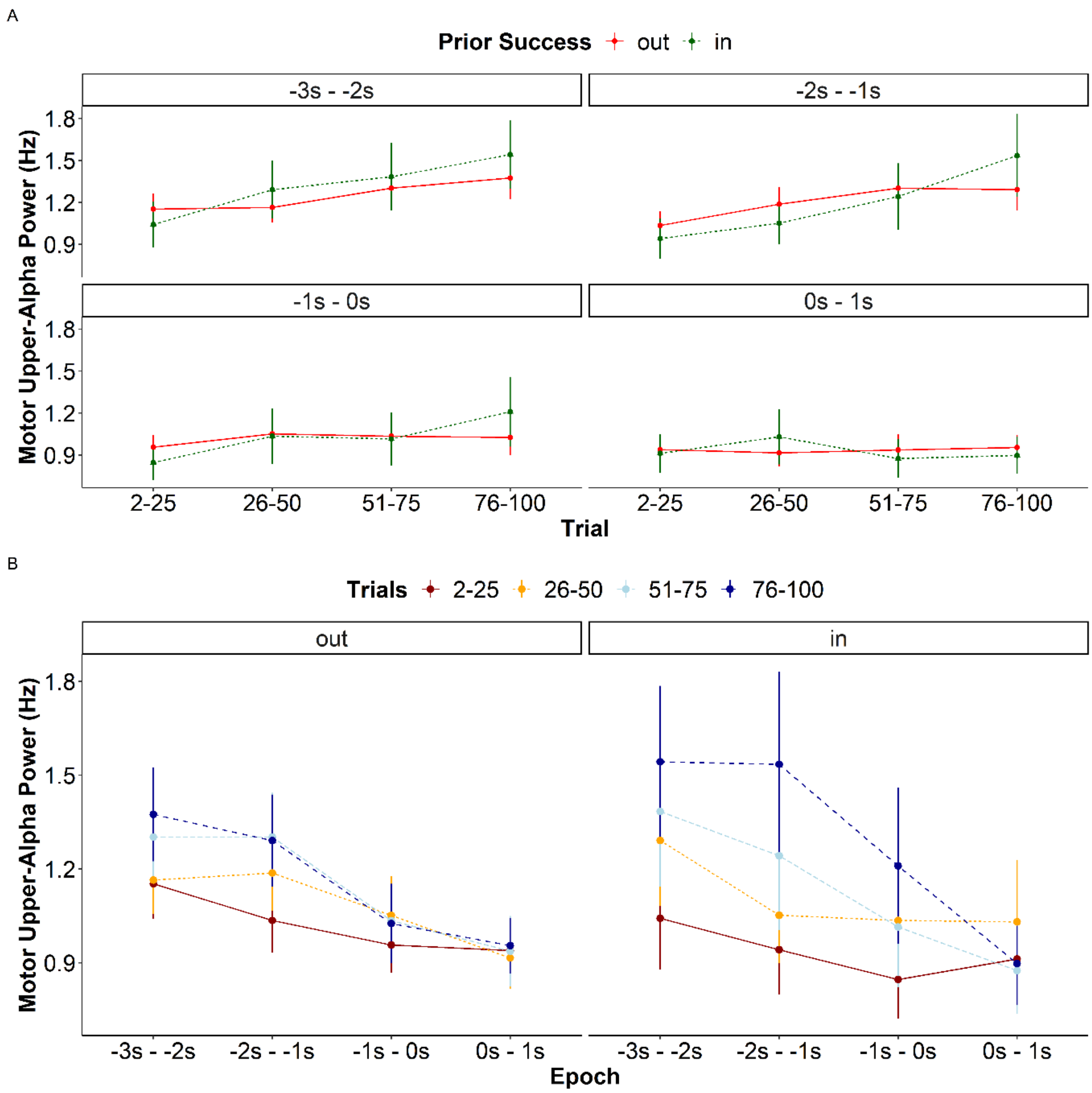
Relationship Between Epoch and Motor Upper-Alpha Power



Note. Figure shows, for each epoch, the average motor upper-alpha power across participant for every trial in gray, and the average across trial in black.

Figure 13

Motor Upper-Alpha Power as a Function of Prior Success, Trial Number, and Epoch



Note. A: Figure shows motor upper-alpha power as a function of trial number after successful (green line) and unsuccessful (red line) trials, separated by epoch. B: Figure shows motor upper-alpha power as a function of epoch in trials 2 to 25 (burgundy line), 26 to 50 (orange line), 51-75 (light blue line), and 76-100 (dark blue line), separated by prior success. Vertical bars represent 95% confidence intervals.

Model 2 - Effects of Error Magnitude on Motor Upper-Alpha

Results of the analysis of the effect of error magnitude on motor upper-alpha power are presented in Table 7.

Table 7

Model 2 Type III Analysis of Variance Table with Satterthwaite's method

Effects	*Sum Sq*	*Mean Sq*	*NumDF*	*DenDF*	*F value*	*p-value*
Prior Radial Error	1.48	1.48	1	318.70	0.82	.365
Prior Radial Error^2	2.17	2.17	1	1841.20	1.21	.272
Prior Radial Error^3	0.02	0.02	1	441.50	0.010	.920
Trial	13.96	13.96	1	48.10	7.78	.008*
Trial^2	1.44	1.44	1	168.70	0.805	.371
Epoch	39.86	13.29	3	262.30	7.41	<.001*
Prior Radial Error:Trial	4.46	4.46	1	9658.50	2.49	.115
Prior Radial Error:Trial^2	0.13	0.13	1	8829.50	0.07	.791
Prior Radial Error:Epoch	2.87	0.96	3	14876.90	0.53	.659
Prior Radial Error^2:Trial	8.78	8.78	1	8917.40	4.90	.027*
Prior Radial Error^2:Trial^2	17.67	17.69	1	8604.80	9.85	.002*
Prior Radial Error^2:Epoch	6.63	2.21	3	14946.80	1.23	.297
Prior Radial Error^3:Trial	17.55	17.55	1	730.40	9.78	.002*
Prior Radial Error^3:Trial^2	25.37	25.37	1	657.80	14.15	<.001*
Prior Radial Error^3:Epoch	0.94	0.31	3	14921.00	0.17	.914
Trial:Epoch	27.16	9.05	3	14872.90	5.05	.002*
Trial^2:Epoch	3.08	1.03	3	14868.60	0.57	.633
Prior Radial Error:Trial:Epoch	11.59	3.86	3	14884.60	2.15	.091
Prior Radial Error:Trial^2:Epoch	9.58	3.19	3	14883.20	1.78	.148
Prior Radial Error^2:Trial:Epoch	12.35	4.12	3	14880.90	2.29	.076
Prior Radial Error^2:Trial^2:Epoch	3.19	1.06	3	14884.00	0.59	.620
Prior Radial Error^3:Trial:Epoch	36.47	12.16	3	14879.10	6.78	<.001*
Prior Radial Error^3:Trial^2:Epoch	16.60	5.53	3	14877.80	3.09	.026*

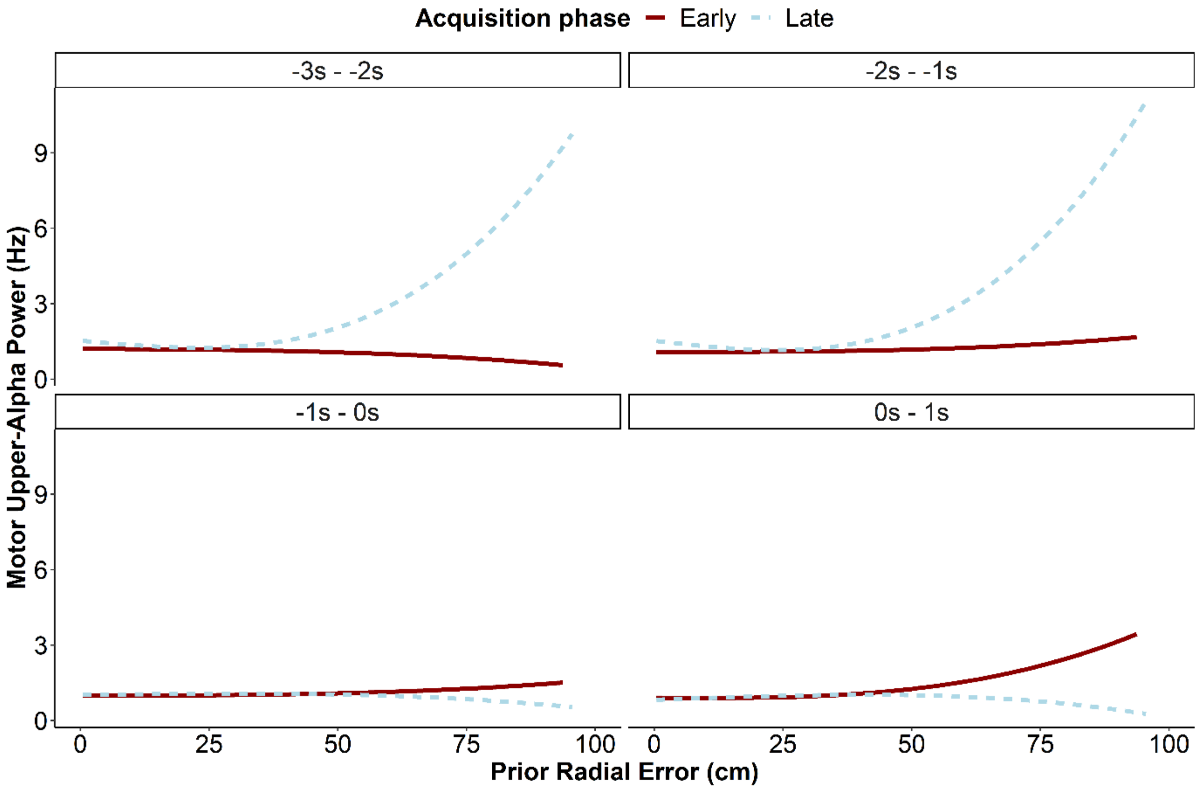
Note. Number of observations: 15081, participants: 42. * indicates significant differences. Sum Sq = sum of squares. Mean Sq = mean square. NumDF = numerator degrees of freedom. DenDF = denominator degrees of freedom.

The analysis revealed a significant main effect of trial (omnibus test $p = .008$) and a significant main effect of epoch (omnibus test $p < .001$). These effects were superseded by significant interactions of Prior Radial Error^2 by Trial (omnibus test $p = .027$), Prior Radial Error^3 by Trial (omnibus test $p = .002$), and Trial by Epoch (omnibus test $p = .002$). Significant interactions of Prior Radial Error^2 by Trial^2 (omnibus test $p = .002$), and Prior Radial Error^3 by Trial^2 (omnibus test $p < .001$) were also found. Finally, we detected three-way interactions of Prior Radial Error^3 by Trial by Epoch (omnibus test $p < .001$), and Prior Radial Error^3 by Trial^2 by Epoch (omnibus test $p = .026$).

No other main effects ($ps \geq .272$) or interactions ($ps \geq .076$) were found. Figure 14 depicts the effect of prior radial error and trial number on motor upper-alpha power for each epoch.

Figure 14

Motor Upper-Alpha Power as a Function of Prior Radial Error, Acquisition Phase, and Epoch



Note. Figure shows the relationship between motor upper-alpha power and prior radial error in each epoch as a function of acquisition phase, where the early phase is in burgundy and represents trials 2 to 50, and the late phase is in light blue and represents trials 51-100.

Comparing Model 1 and Model 2

The Wald likelihood ratio tests indicated that Model 2, but not Model 1, was a significant improvement above the Reference (trial and epoch-only) Model, and that Model 2's likelihood ratio was significantly different from Model 1's likelihood ratio (Table 8). It is important to notice, however, that Model 2 is more complex than both the Reference Model and Model 1, which can result in overfitting.

Table 8

Wald Likelihood Ratio Test for the Reference Model, Model 1, and Model 2 of Motor Upper-Alpha Power

Model	*npar*	*AIC*	*logLik*	*Test*	*L.Ratio*	*p-value*
Reference	20	52432	-26196			
Model 1	29	52444	-26193	Reference vs. Model 1	5.77	.763
Model 2	65	52263	-26067	Model 1 vs. Model 2	252.73	<.001*

Note. * indicates significant differences. npar= number of parameters. AIC= Akaike Information Criterion. logLik= log-likelihood. L.Ratio= Absolute likelihood-ratio.

The conditional pseudo-R^2 revealed that, overall, Model 2 had a higher effect size than Model 1, suggesting that error magnitude explained more of the motor upper-alpha power variance than perceptions of success (Table 9).

Table 9

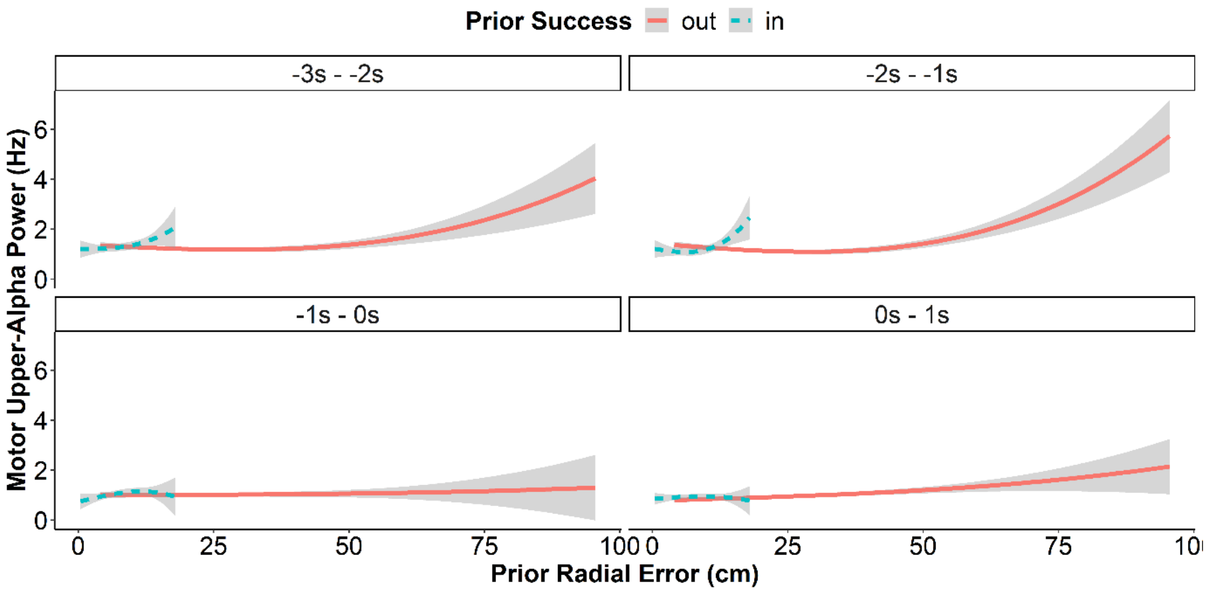
Conditional Pseudo-R^2 Test for the Reference Model, Model 1, and Model 2 on Motor Upper-Alpha Power

Model	*Conditional R^2*
Reference	.29
Model 1	.29
Model 2	.31

Figure 15 depicts the relationship between prior success, prior radial error, and motor upper-alpha power.

Figure 15

Relationship of Prior Success and Prior Radial Error on Motor Upper-Alpha Power



Note. Figure shows, for each epoch, the relationship between motor upper-alpha power and prior radial error for previous successful (in) and unsuccessful (out) trials in traced blue and solid pink lines, respectively. The shaded areas represent 95% confidence intervals.

Model 3 - Effects of Zone of Success Assignment and Error Magnitude on Motor Upper-Alpha Power

Results of the analysis of the effect of assigned zone of success, error magnitude on the previous trial, and trial number on motor upper-alpha power are presented in Table 10.

Table 10

Model 3 Type III Analysis of Variance Table with Satterthwaite's method

Effects	*Sum Sq*	*Mean Sq*	*NumDF*	*DenDF*	*F value*	*p-value*
Zone	<0.01	<0.01	1	41.30	<0.01	.983
Prior Radial Error	0.84	0.84	1	661.80	0.47	.494
Prior Radial Error²	3.41	3.41	1	13622.70	1.90	.169
Prior Radial Error³	0.80	0.80	1	8485.20	0.45	.504
Trial	12.94	12.94	1	47.30	7.19	.010*
Trial²	1.12	1.12	1	165.90	0.62	.432

Epoch	39.20	13.07	3	248.20	7.26	<.001*
Zone:Prior Radial Error	0.98	0.98	1	661.80	0.54	.462
Zone:Prior Radial Error²	0.15	0.15	1	13622.70	0.08	.773
Zone:Prior Radial Error³	0.27	0.27	1	8485.20	0.15	.701
Zone:Trial	0.30	0.30	1	47.30	0.17	.685
Zone:Trial²	1.31	1.31	1	165.90	0.73	.394
Zone:Epoch	8.48	2.83	3	248.20	1.57	.197
Prior Radial Error:Trial	2.67	2.67	1	14512.30	1.48	.224
Prior Radial Error:Trial²	0.23	0.23	1	13628.60	0.13	.720
Prior Radial Error:Epoch	3.08	1.03	3	14794.10	0.57	.634
Prior Radial Error²:Trial	7.26	7.26	1	11507.40	4.03	.045*
Prior Radial Error²:Trial²	14.70	14.70	1	9859.40	8.17	.004*
Prior Radial Error²:Epoch	3.14	1.05	3	14852.50	0.58	.627
Prior Radial Error³:Trial	4.99	4.99	1	8917.10	2.78	.096
Prior Radial Error³:Trial²	10.00	10.00	1	7954.00	5.56	.018*
Prior Radial Error³:Epoch	3.39	1.13	3	14780.30	0.63	.596
Trial:Epoch	23.20	7.73	3	14787.70	4.30	.005*
Trial²:Epoch	3.26	1.09	3	14779.20	0.60	.613
Zone:Prior Radial Error:Trial	0.18	0.18	1	14512.30	0.10	.750
Zone:Prior Radial Error:Trial²	4.86	4.86	1	13628.60	2.70	.100
Zone:Prior Radial Error:Epoch	6.55	2.18	3	14794.10	1.21	.303
Zone:Prior Radial Error²:Trial	0.03	0.03	1	11507.40	0.02	.894
Zone:Prior Radial Error²:Trial²	0.30	0.30	1	9859.40	0.17	.681
Zone:Prior Radial Error²:Epoch	9.43	3.14	3	14852.50	1.75	.155
Zone:Prior Radial Error³:Trial	0.02	0.02	1	8917.10	0.01	.921
Zone:Prior Radial Error³:Trial²	0.60	0.60	1	7954.00	0.33	.564
Zone:Prior Radial Error³:Epoch	9.57	3.19	3	14780.30	1.77	.150
Prior Radial Error:Trial:Epoch	3.55	1.19	3	14814.50	0.66	.578
Prior Radial Error:Trial²:Epoch	2.82	0.94	3	14791.00	0.52	.667
Prior Radial Error²:Trial:Epoch	11.32	3.78	3	14803.20	2.10	.098
Prior Radial Error²:Trial²:Epoch	1.77	0.59	3	14797.20	0.33	.806
Prior Radial Error³:Trial:Epoch	3.59	1.20	3	14822.10	0.67	.574
Prior Radial Error³:Trial²:Epoch	0.61	0.20	3	14801.40	0.11	.952
Zone:Prior Radial Error:Trial:Epoch	15.17	5.06	3	14814.50	2.81	.038*
Zone:Prior Radial Error:Trial²:Epoch	8.41	2.80	3	14791.00	1.56	.197
Zone:Prior Radial Error²:Trial:Epoch	1.58	0.53	3	14803.20	0.29	.831
Zone:Prior Radial Error²:Trial²:Epoch	7.35	2.45	3	14797.20	1.36	.253
Zone:Prior Radial Error³:Trial:Epoch	6.61	3.20	3	14822.10	1.78	.148
Zone:Prior Radial Error³:Trial²:Epoch	9.63	3.21	3	14801.40	1.79	.148

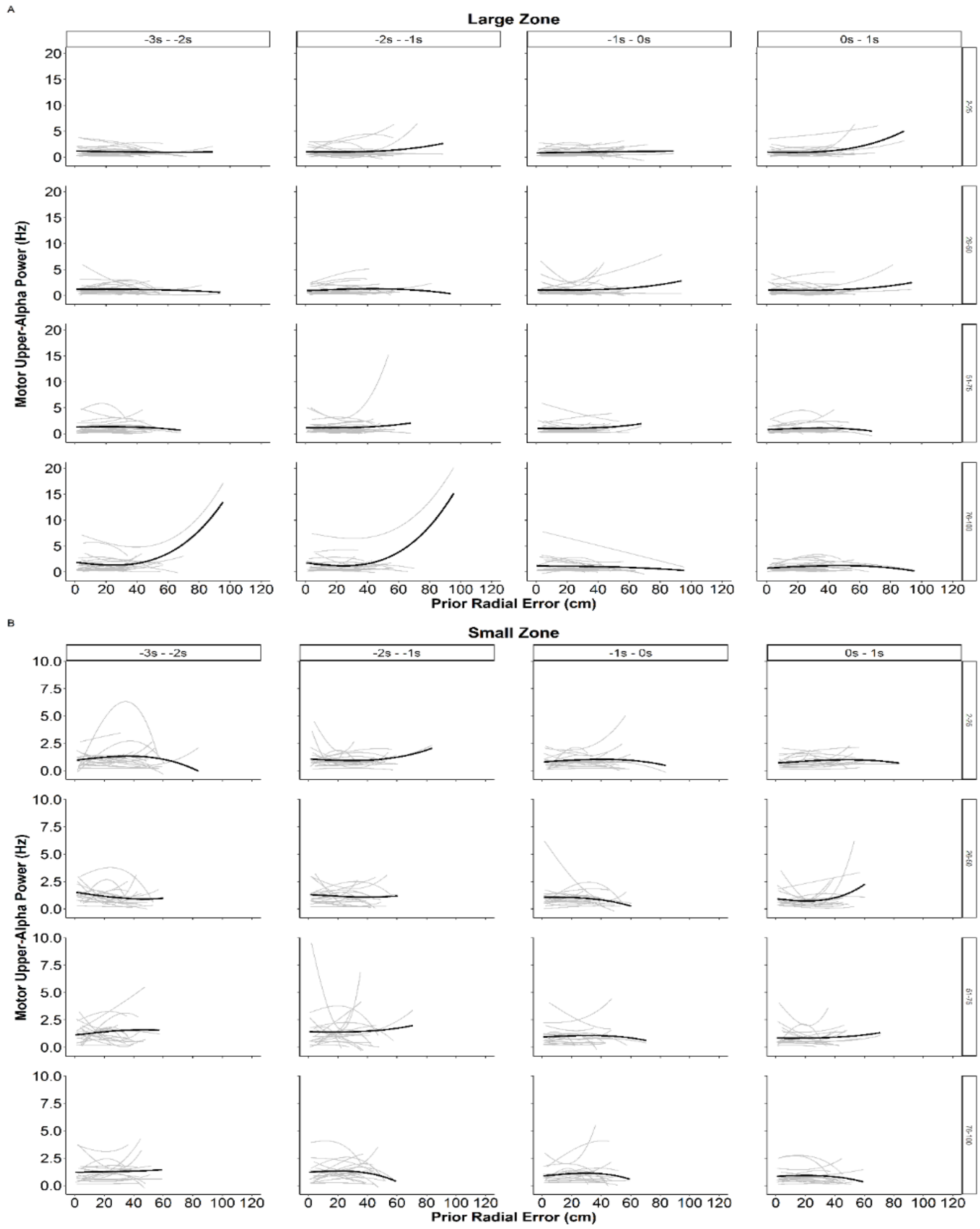
Note. Number of observations: 15081, participants: 42. * indicates significant differences. Sum Sq = sum of squares. Mean Sq = mean square. NumDF = numerator degrees of freedom. DenDF = denominator degrees of freedom.

The analysis revealed a positive significant main effect of trial (omnibus test $p = .010$) and a main effect of epoch (omnibus test $p < .001$). These effects were superseded by significant interactions of Trial by Epoch (omnibus test $p = .005$), Prior Radial Error^3 by Trial (omnibus test $p = .045$), and Zone by Prior Radial Error by Trial by Epoch (omnibus test $p = .038$). Significant interactions of Prior Radial Error^2 by Trial^2 (omnibus test $p = .004$), and Prior Radial Error^3 by Trial^2 (omnibus test $p = .018$) were also found. No other main effects (ps $\geq .169$) or interactions (ps $\geq .096$) were found. To unpack the Zone by Prior Radial Error by Trial by Epoch interaction, we conducted post-hoc tests where we ran the same model (without zone as a predictor) for each zone of success[5]. A significant Prior Radial Error by Trial by Epoch interaction was only found for the large zone group (omnibus test $p = .011$). For the small zone group, only a significant Trial by Epoch interaction (omnibus test $p = .035$) and significant main effect of trial (omnibus test $p = .013$) were found (Figure 16).

Figure 16

Relationship Between Zone of Success, Prior Radial Error, Trial Number, Epoch, and Motor-Upper Alpha Power

[5] Because the model did not converge with trial^2 in the random-effects of the large zone group model, we removed it from the random-effects of both the large zone group and the small zone group models.



Note: Figure A shows, for the large zone group, the relationship between motor upper-alpha power and prior radial error in trials 2 to 25, 26 to 50, 51-75, and 76-100 (represented by different rows) for every epoch (represented by different columns). The black line represents the cubic relationship between motor upper-alpha and prior radial error for the sample, whereas the gray lines represent the cubic slope for each participant. Figure B shows, for the small zone group, the relationship between motor upper-alpha power and prior radial error in trials 2 to 25, 26 to 50, 51-75, and 76-100 (represented by different rows) for every epoch (represented by different columns). The black line represents the cubic relationship between motor upper-alpha and prior radial error for the sample, whereas the gray lines represent the cubic slope for each participant.

Discussion

Performance success has a bidirectional relationship with feedback-related and motor-preparatory brain activity. Feedback about a successful movement serves as a positive RPE that reinforces the precipitating action, so that it can be repeated in future trials. On the contrary, feedback about an unsuccessful action signals that neural resources should be allocated to reprogramming the movement in future trials, so that the unsuccessful action is avoided and the goal can be achieved (Margraf et al., 2022). Importantly, an objective outcome can lead to feedback that is subjectively perceived as successful or unsuccessful. Therefore, it is important to investigate how objective and subjective outcomes affect feedback-related and motor-preparatory brain activity. Thus, we recorded participants' EEG while having them practice a motor skill with an objective outcome (how far from the center of a target a puck that they shot landed [error magnitude]) and a subjective outcome (whether the puck landed in a zone of success surrounding the center of the target). Some participants had a relatively large zone of success and others had a relatively small zone of success. We explored how the objective and subjective success of trials as well as their interaction influenced the EEG measures of RewP amplitude, a proxy of RPE, and motor upper-alpha power, an index of motor programming. We also examined how experience (practice trial number) and its interaction with objective and subjective success influenced RewP and motor upper-alpha power.

Based on previous literature showing that feedback with a positive valence elicits larger RewPs than feedback with a negative valence (Margraf et al., 2022; Weinberg et al., 2014), we hypothesized that trials that stopped within participants' zone of success would elicit larger RewPs than unsuccessful trials. Results from our first model confirm this hypothesis and are consistent with the assumption that feedback with positive valence is interpreted as rewarding. Crucially, the zone of success assigned to participants was arbitrary, indicating that giving subjective meaning to learners' performance outcomes can be an effective and simple way to manipulate the brain reward system during practice. This result is consistent with past research showing that a simple manipulation of performers' perceptions is capable of affecting RPEs (RewPs) during practice. In Wilhelm et al. (2019), fictitious information about the performance of other participants in the task led performers to believe some blocks of trials were more difficult than others. Even though task difficulty remained the same in each block and participants had about the same number of successful trials in each block, positive feedback during "difficult" blocks led to larger RewPs. We add to these results by showing that not only the perceived likelihood of success but also the actual perception of success during practice can be modulated by task instructions to affect feedback processing.

Crucially, RewP seems to be responsive not only to binary and subjective feedback (e.g., in or out of zone of success) but also to graded, objective feedback (e.g., error magnitude [radial error]), as revealed by Frömer et al. (2016). In that study, successful trials with smaller error magnitudes produced more positive RewPs than successful trials with larger error magnitudes. Similar conclusions can be drawn from Model 2 in our study, which showed that smaller errors resulted in larger RewPs, supporting our hypothesis that error magnitude would show a negative relationship with RewP amplitude.

RPEs are expected to be affected by the valence and value of the reward, but are also expected to be negatively affected by reward likelihood. For instance, Williams et al. (2017) reported smaller RewPs following good trials in easier conditions wherein good performance was expected, as compared to difficult conditions in which successful trials were rare. Thus, we hypothesized that there would be a decrease in RewP amplitude throughout practice, given that performers should expect more positive outcomes later in practice as they improve. However, although trial had a negative effect on RewP when alone in the model, no effect of trial was found when controlling for radial error and success. The mini-shuffleboard task used in this experiment was difficult, as evidenced by an average success rate below 50% for participants, even for those with the easier criterion of success (Table 2 of Chapter 2), and by a non-significant performance improvement during the acquisition phase (Figure 4 of Chapter 2). Given the small performance improvement during acquisition, it is possible that participants' expectations for positive outcomes did not change considerably from the beginning to the end of acquisition, explaining the lack of a trial number effect on RewP. It is possible that a main effect of trial number could still be found in lengthier practice sessions or for an easier motor task, similar to the effects of practice found on a cognitive task in Williams et al. (2018).

After observing that both perception of success and radial error explain RewP variance, we compared models to identify which one explained more variance. Interestingly, we observed that, although the radial error model (Model 2) was more complex (i.e., had more parameters due to the quadratic terms of radial error added to both the fixed and random effects), the success model (Model 1) had a larger effect on feedback processing. To the best of our knowledge, this is the first study that compared the effects of subjective and objective success on RewP during skill acquisition. Our results indicate that, to the RewP, the subjective interpretation of the outcome matters more than its objective level of accuracy.

Beyond the main effects of perception of success, error magnitude, and experience on feedback processing, we also expected these predictors to interact with each other. However, we did not observe interactions involving experience. These interaction hypotheses were largely based on the premise that participants would perform better from the beginning to the end of acquisition. Specifically, we expected that unsuccessful trials and inaccurate trials would result in small RewPs throughout practice, but successful and accurate trials would initially result in high RewP amplitudes that would progressively decrease as participants came to expect these outcomes later in practice. However, that participants probably did not change their expectations for success and accuracy during acquisition likely precluded trial number from moderating the effects of error magnitude or perception of success. We also expected an interaction between perception of success and error magnitude and an interaction between perception of success, error magnitude, and experience. Unfortunately, we were unable to test these hypotheses directly, due to multicollinearity between perception of success and error magnitude. Instead, we tested whether the criterion of success assigned to the performer moderated the effect of radial error on RewP. Interestingly, our results revealed an interaction between the quadratic term of radial error and zone of success. Specifically, for trials with the smallest errors, participants with a small zone of success exhibited larger RewPs than participants with a large zone of success. Conversely, for trials with more moderate errors, participants with a large zone of success exhibited larger RewPs than participants with a small zone of success. For trials with the largest errors, participants with a small zone of success exhibited larger RewPs, but this result is based on relatively few data points and is likely forced by our quadratic model; thus, it should be interpreted with caution. Our results are consistent with the reinforcement learning theory prediction that when reward is infrequent, as was the case for participants with the small zone of success, positive feedback is surprising and results in large RPEs. For participants with the large zone of success, moderately accurate trials were often successful, which explains why these participants had larger RewPs for these trials than participants with a small zone of success.

Overall, our results suggest that simple task instructions provided by an instructor, coach, or physical therapist can shape the way learners interpret performance outcomes, and consequently their rewarding value. A task that has a difficult criterion for success will make successful outcomes rare but highly rewarding. A task that has an easier criterion for success will make successful outcomes more frequent but less rewarding. Increasing the value of an action is expected to result in increased dopaminergic activity in the midbrain and to be associated with motor memory consolidation, increasing the likelihood that the action is selected in the future (Lohse et al., 2019). Importantly, however, frequent activation of the neural reward system may promote long-term retention. Therefore, in practice, a moderately challenging criterion of success should be adopted, such that learners have their reward system activated when they achieve successful outcomes, and these outcomes are somewhat frequent. This conclusion is aligned with the optimal challenge point framework, which posits that the functional level of task difficulty has an inverted-U relationship with the potential learning benefit of a practice session, such that the optimal challenge point does not coincide with the highest nor the lowest level of success on the task (Guadagnoli & Lee, 2004).

In addition to mechanisms of feedback processing, motor-preparatory brain activity was also shown to be affected by perceptions of success. We expected that higher motor upper-alpha power would be observed on trials following those perceived as successful in comparison to those perceived as unsuccessful, given that participants would be less inclined to reprogram successful movements, thus requiring fewer neural resources for movement preparation. Although we did not directly confirm this hypothesis, a prior success by trial number by epoch interaction in Model 1 showed that this effect is dependent on task experience, being mostly observed later in practice. This result is likely due to the fact that in initial stages of learning, learners have a weak internal model (Callan et al., 2014), making it difficult to understand how to repeat the movement that resulted in success. Additionally, movement preparation generally is consciously controlled and inefficient in that stage (Fitts & Posner, 1967), resulting in a large allocation of neural resources regardless of the previous outcome. As the learner gains control over their movement and performs it with more automaticity, successful outcomes progressively result in less motor programming activity, while unsuccessful outcomes still compel learners to allocate substantial neural resources to movement preparation, so that the previous movement can be corrected, and the goal can be achieved. Importantly, the change in motor upper-alpha power across practice for trials following successful outcomes was only observed in the seconds preceding movement onset, not in the second following the start of the movement. This result indicates that the decreased allocation of resources throughout practice is strictly reserved for movement preparation, rather than for movement execution. Along the same lines, Cooke et al. (2015) found that experts showed higher upper-alpha power following successful trials than unsuccessful trials in the three seconds that preceded movement initiation (but not in the second after movement initiation), while no effect of outcome was found for novices. We expand upon these results by showing that the effect of objective performance outcomes can be created with subjective performance outcomes, using simple and arbitrary task instructions, and that the effect of experience can be observed on a trial-by-trial basis, over a single practice session.

Interestingly, when determining the quantity of resources to allocate to motor programming, the neural system seems to consider the performance outcome of the prior trial in a graded fashion in addition to a binary one. Although we did not find a main effect of radial error on motor upper-alpha power as we hypothesized, our Model 2 results showed that error magnitude affected movement preparation in a non-linear fashion, when experience and epoch were also considered. Crucially, this model had a larger effect size than the success model (Model 1), suggesting that prior radial error explains more of motor upper-alpha power variance than prior success. Although this result should be taken with caution, given the difference in complexity between these two models, we provide some evidence that, different from feedback processing, objective success seems to matter more than subjective success for movement preparation. As far as we know, this is the first study to compare these effects and the first study to show that motor upper-alpha power is responsive to graded feedback. The strict use of binary measurements (e.g., made vs. missed shots) confounded the effects of outcome success and error magnitude in previous studies.

Although we wanted to test, in an exploratory way, whether perception of success and error magnitude interact to affect movement preparation, we were unable to do so, due to multicollinearity between these variables. Instead, we tested whether the criterion of success assigned to the performer moderated the effect of prior radial error on motor upper-alpha power. We found that providing learners with a criterion of success further complicates the relationship between error magnitude, trial number, and epoch, such that this interaction was present for participants with the large zone, but not for participants with the small zone. Therefore, for learners with an easy criterion of success, motor-preparatory brain activity seems to be modulated by interactions of objective success, experience, and epoch, while learners with a more difficult criterion appear to have their motor-preparatory brain activity influenced by experience and epoch. Once again, these results confirm that task instructions that affect learners' perceptions of success can affect the way movement is programmed during practice.

Finally, based on previous studies that showed effects of expertise on motor upper-alpha power (Cooke et al., 2014, 2015; Percio et al., 2010), we predicted a positive main effect of trial number, such that the more experienced the participants became, the fewer resources would be allocated for motor programming. The effect of trial was indeed consistent across models, but it was also shown to be moderated by other variables (prior success, prior radial error, epoch, and zone), sometimes assuming a quadratic relationship in these interactions. It can be concluded that, overall, motor upper-alpha power tends to increase throughout practice, but this increase can be larger or smaller depending on the outcome of the previous trial, the determined criterion of success, and epoch. Interestingly, even though the performance of the participants did not improve significantly across acquisition blocks (see Chapter 2), fewer resources were used to achieve similar outcomes, indicating higher neural efficiency as a function practice (Neubauer & Fink, 2009).

In conclusion, the results of this study make it clear that both subjective success and objective success have significant and important effects on mechanisms of feedback processing and movement preparation that underlie motor skill acquisition, such as modulating the value and frequency of RPEs received in a practice session, and the quantity of neural resources dedicated to motor-preparatory brain activity. Importantly, we demonstrated that these effects can be achieved by the establishment of arbitrary criteria for success with simple task instructions that give subjective meaning to learners' outcomes.

References

Abbas, Z. A., & North, J. S. (2018). Good-vs. poor-trial feedback in motor learning: The role of self-efficacy and intrinsic motivation across levels of task difficulty. *Learning and Instruction*, *55*, 105–112. https://doi.org/10.1016/j.learninstruc.2017.09.009

Abe, M., Schambra, H., Wassermann, E. M., Luckenbaugh, D., Schweighofer, N., & Cohen, L. G. (2011). Reward improves long-term retention of a motor memory through induction of offline memory gains. *Current Biology*, *21*(7), 557–562. https://doi.org/10.1016/j.cub.2011.02.030

Ávila, L. T. G., Chiviacowsky, S., Wulf, G., & Lewthwaite, R. (2012). Positive social-comparative feedback enhances motor learning in children. *Psychology of Sport and Exercise*, *13*(6), 849–853. https://doi.org/10.1016/j.psychsport.2012.07.001

Babiloni, C., Percio, C. D., Iacoboni, M., Infarinato, F., Lizio, R., Marzano, N., Crespi, G., Dassù, F., Pirritano, M., Gallamini, M., & Eusebi, F. (2008). Golf putt outcomes are predicted by sensorimotor cerebral EEG rhythms. *The Journal of Physiology*, *586*(Pt 1), 131. https://doi.org/10.1113/jphysiol.2007.141630

Bacelar, M. F. B., Lohse, K. R., & Miller, M. W. (2020). The Effect of Rewards and Punishments on Learning Action Selection and Execution Components of a Motor Skill. *Journal of Motor Learning and Development*, *1*(aop), 1–22. https://doi.org/10.1123/jmld.2019-0039

Bacelar, M. F. B., Parma, J. O., Cabral, D., Daou, M., Lohse, K. R., & Miller, M. W. (2022). Dissociating the contributions of motivational and information processing factors to the self-controlled feedback learning benefit. *Psychology of Sport and Exercise*, *59*, 102119. https://doi.org/10.1016/j.psychsport.2021.102119

Bacelar, M. F. B., Parma, J. O., Murrah, W. M., & Miller, M. W. (2022). Meta-analyzing enhanced expectancies on motor learning: Positive effects but methodological concerns. *International Review of Sport and Exercise Psychology*. https://doi.org/10.1080/1750984X.2022.2042839

Bahmani, M., Diekfuss, J. A., Rostami, R., Ataee, N., & Ghadiri, F. (2018). Visual illusions affect motor performance, but not learning in highly skilled shooters. *Journal of Motor Learning and Development*, *6*(2), 220–233. https://doi.org/10.1123/jmld.2017-0011

Bahmani, M., Wulf, G., Ghadiri, F., Karimi, S., & Lewthwaite, R. (2017). Enhancing performance expectancies through visual illusions facilitates motor learning in children. *Human movement science*, 55, 1-7. https://doi.org/10.1016/j.humov.2017.07.001

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents*, 5, pp. 307-337. Greenwich, CT: Information Age Publishing.

Barker, J., Jones, M., & Greenlees, I. (2010). Assessing the immediate and maintained effects of hypnosis on self-efficacy and soccer wall-volley performance. Journal of Sport and Exercise Psychology, 32(2), 243–252. https://doi.org/10.1123/jsep.32.2.243

Barton, K. (2018). *MuMIn*: Multi-model inference. R package version 1.42.1. https://CRAN.R-project.org/package=MuMIn

Barzouka, K., Bergeles, N., & Hatziharistos, D. (2007). Effect of simultaneous model observation and self-modeling of volleyball skill acquisition. *Perceptual and Motor Skills*, 104(1), 32–42. https://doi. org/10.2466/pms.104.1.32-42

Baumeister, R. F. (1984). Choking under pressure: Self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of Personality and Social Psychology*, *46*, 610–620. https://doi.org/10.1037/0022-3514.46.3.610

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference*: A practical information-theoretic approach, 2 ed.

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, 14(5), 365–376. https://doi.org/10.1038/nrn3475

Caldwell, A. R.. *TOSTER*: An R package and jamovi module for equivalence testing. (2022). https://aaroncaldwell.us/TOSTERpkg

Caldwell, A. R., Vigotsky, A. D., Tenan, M. S., Radel, R., Mellor, D. T., Kreutzer, A., Lahart, I. M., Mills, J. P., Boisgontier, M. P., & Consortium for Transparency in Exercise Science (COTES) Collaborators. (2020). Moving sport and exercise science forward: A call for the adoption of more transparent research practices. *Sports Medicine*, 50(3), 449–459. https://doi.org/10.1007/s40279-019-01227-1

Callan, D. E., & Naito, E. (2014). Neural processes distinguishing elite from expert and novice athletes. Cognitive and Behavioral Neurology, 27(4), 183-188. https://doi.org/10.1097/WNN.0000000000000043

Carter, E. C., Schönbrodt, F. D., Gervais, W. M., & Hilgard, J. (2019). Correcting for bias in psychology: A comparison of meta-analytic methods. *Advances in Methods and Practices in Psychological Science*, 2(2), 115–144. https://doi.org/10.1177/2515245919847196

Carter, M. J., Smith, V., & Ste-Marie, D. M. (2016). Judgments of learning are significantly higher following feedback on relatively good versus relatively poor trials despite no actual learning differences. *Human Movement Science*, 45, 63–70. https://doi.org/10.1016/j.humov.2015.11.006

Chauvel, G., Wulf, G., & Maquestiaux, F. (2015). Visual illusions can facilitate sport skill learning. *Psychonomic Bulletin & Review*, 22(3), 717–721. https://doi.org/10.3758/s13423-014-0744-9

Chiviacowsky, S., Ávila, L., & Kaefer, A. (2010). Efeitos do feedback após boas tentativas de prática na aprendizagem de uma habilidade motora complexa em crianças. *Revista Da Educação Física/UEM*, 21(2), 183–190. https://doi.org/10.4025/reveducfis.v21i2.7745

Chiviacowsky, S., Cardozo, P. L., & Chalabaev, A. (2018). Age stereotypes' effects on motor learning in older adults: The impact may not be immediate, but instead delayed. *Psychology of Sport and Exercise*, 36, 209–212. https://doi.org/10.1016/j.psychsport.2018.02.012

Chiviacowsky, S., & Drews, R. (2014). Effects of generic versus non-generic feedback on motor learning in children. *Plos One*, 9(2), e88989. https://doi.org/10.1371/journal.pone.0088989

Chiviacowsky, S., & Drews, R. (2016). Temporal-comparative feedback affects motor learning. *Journal of Motor Learning and Development*, 4(2), 208–218. https://doi.org/10.1123/jmld.2015-0034

Chiviacowsky, S., & Harter, N. M. (2015). Perceptions of competence and motor learning: Performance criterion resulting in low success experience degrades learning. *Brazilian Journal of Motor Behavior*, 9(1). https://doi.org/10.20338/bjmb.2015-0002

Chiviacowsky, S., Harter, N. M., Gonçalves, G. S., & Cardozo, P. L. (2019). Temporal-comparative feedback facilitates golf putting. *Frontiers in Psychology*, 9, 2691. https://doi.org/10.3389/fpsyg.2018.02691

Chiviacowsky, S., & Wulf, G. (2007). Feedback after good trials enhances learning. *Research Quarterly for Exercise and Sport*, 78(2), 40–47. https://doi.org/10.1080/02701367.2007.10599402

Chiviacowsky, S., Wulf, G., & Lewthwaite, R. (2012). Self-controlled learning: The importance of protecting perceptions of competence. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00458

Chiviacowsky, S., Wulf, G., Wally, R., & Borges, T. (2009). Knowledge of results after good trials enhances learning in older adults. *Research Quarterly for Exercise and Sport*, 80(3), 663–668. https://doi.org/10.1080/02701367.2009.10599606

Chung, Y., Lewthwaite, R., Winstein, C. J., Monterosso, J. R., & Fisher, B. E. (2020). Expectancy and affective response to challenging balance practice conditions in individuals with parkinson's disease. *European Journal of Neuroscience*. psyh. https://doi.org/10.1111/ejn.14723

Cooke, A., Gallicchio, G., Kavussanu, M., Willoughby, A., McIntyre, D., & Ring, C. (2015). Premovement high-alpha power is modulated by previous movement errors: Indirect evidence to endorse high-alpha power as a marker of resource allocation during motor programming. *Psychophysiology*, *52*(7), 977–981. https://doi.org/10.1111/psyp.12414

Cooke, A., Kavussanu, M., Gallicchio, G., Willoughby, A., McIntyre, D., & Ring, C. (2014). Preparation for action: Psychophysiological activity preceding a motor skill as a function of expertise, performance outcome, and psychological pressure. *Psychophysiology*, *51*(4), 374–384. https://doi.org/10.1111/psyp.12182

Daou, M., Lohse, K. R., & Miller, M. W. (2018). Does practicing a skill with the expectation of teaching alter motor preparatory cortical dynamics? *International Journal of Psychophysiology*, *127*, 1–10. https://doi.org/10.1016/j.ijpsycho.2018.02.013

DeCaro, M. S., Thomas, R. D., Albert, N. B., & Beilock, S. L. (2011). Choking under pressure: Multiple routes to skill failure. *Journal of Experimental Psychology: General*, *140*, 390–406. https://doi.org/10.1037/a0023466

Deci, E. L., & Ryan, R. M. (2000). The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychological Inquiry*, *11*(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01

Del Percio, C., Infarinato, F., Iacoboni, M., Marzano, N., Soricelli, A., Aschieri, P., ... & Babiloni, C. (2010). Movement-related desynchronization of alpha rhythms is lower in athletes than non-athletes: a high-resolution EEG study. *Clinical Neurophysiology*, 121(4), 482-491.

Diekfuss, J. A., & Raisbeck, L. D. (2016). Focus of attention and instructional feedback from NCAA division 1 collegiate coaches. *Journal of Motor Learning and Development*, 4(2), 262–273. https://doi.org/10.1123/jmld.2015-0026

Drews, R., Chiviacowsky, S., & Wulf, G. (2013). Children's motor skill learning is influenced by their conceptions of ability. *Journal of Motor Learning and Development,* 1(2), 38–44. https://doi.org/10.1123/jmld.1.2.38

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, 56(2), 455–463. https://doi.org/10.1111/j.0006-341X.2000.00455.x

Dyke, F., Godwin, M. M., Goel, P., Rehm, J., Rietschel, J. C., Hunt, C. A., & Miller, M. W. (2014). Cerebral cortical activity associated with non-experts' most accurate motor performance. *Human Movement Science*, *37*, 21–31. https://doi.org/10.1016/j.humov.2014.06.008

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, 315(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629

Englund, G., Sarnelle, O., & Cooper, S. D. (1999). The importance of data-selection criteria: Meta-analyses of stream predation experiments. *Ecology*, 80(4), 1132–1141. https://doi.org/10.1890/0012-9658(1999)080[1132:TIODSC]2.0.CO;2

Fanelli, D. (2012). Negative results are disappearing from most disciplines and countries. *Scientometrics*, 90(3), 891–904. https://doi.org/10.1007/s11192-011-0494-7

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. https://doi.org/10.3758/BF03193146

Fitts, P. M., & Posner, M. I. (1967). *Human performance*. Belmont, CA: Brooks/Cole.

Frömer, R., Lin, H., Wolf, C. K. D., Inzlicht, M., & Shenhav, A. (2021). Expectations of reward and efficacy guide cognitive control allocation. *Nature Communications*, *12*(1), 1030. https://doi.org/10.1038/s41467-021-21315-z

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

Ghorbani, S. (2019). Motivational effects of enhancing expectancies and autonomy for motor learning: An examination of the OPTIMAL theory. *The Journal of General Psychology*, 146(1), 79–92. https://doi.org/10.1080/00221309.2018.1535486

Ghorbani, S., & Bund, A. (2020). Motivational effects of enhanced expectancies for motor learning in individuals with high and low self-efficacy. *Perceptual and Motor Skills*, 127(1), 263–274. https://doi.org/10.1177/0031512519892390

Goudini, R., Saemi, E., Ashrafpoornavaee, S., & Abdoli, B. (2018). The effect of feedback after good and poor trials on the continuous motor tasks learning. *Acta Gymnica*, 48(1), 3–8. https://doi.org/10.5507/ag.2018.001

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), Art. 2. psyh. https://doi.org/10.1123/jmld.2016-0041

Grealy, M. A., Cummings, J., & Quinn, K. (2019). The effect of false positive feedback on learning an inhibitory-action task in older adults. *Experimental Aging Research*, 45(4), 346–356. https://doi.org/10.1080/0361073X.2019.1627494

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), Art. 2. https://doi.org/10.3200/JMBR.36.2.212-224

Hancock, G. R., Butler, M. S., & Fischman, M. G. (1995). On the Problem of Two-Dimensional Error Scores: Measures and Analyses of Accuracy, Bias, and Consistency. *Journal of Motor Behavior*, *27*(3), 241–250. https://doi.org/10.1080/00222895.1995.9941714

Harrer, M., Cuijpers, P., Furukawa, T. A., & Ebert, D. D. (2019). Doing meta-analysis in R: A hands-on guide. https://bookdown.org/MathiasHarrer/Doing_Meta_Analysis_in_R/

Harter, N. M., Cardozo, P. L., & Chiviacowsky, S. (2019). Conceptions of ability influence the learning of a dance pirouette in children. *Journal of Dance Medicine & Science*, 23(4), 167–172. https://doi.org/10.12678/1089-313X.23.4.167

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ*, 327(7414), 557–560. https://doi.org/10.1136/bmj.327.7414.557

Hodges, N. J., & Lohse, K. R. (2022). An extended challenge-based framework for practice design in sports coaching. *Journal of Sports Sciences*, *40*(7), 754–768. https://doi.org/10.1080/02640414.2021.2015917

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

Iwatsuki, T., & Regis, C. J. (2021). Relatively Easy Criteria for Success Enhances Motor Learning by Altering Perceived Competence. *Perceptual and Motor Skills*, *128*(2), 900–911. https://doi.org/10.1177/0031512520981237

Jennings, C. T., Reaburn, P., & Rynne, S. B. (2013). The effect of a self-modelling video intervention on motor skill acquisition and retention of a novice track cyclist's standing start performance. *International Journal of Sports Science & Coaching*, 8(3), 467–480. https://doi.org/10.1260/1747-9541.8.3.467

Jimenez-Diaz, J., Chaves-Castro, K., & Morera-Castro, M. (2020). Effect of self-controlled and regulated feedback on motor skill performance and learning: A meta-analytic study. *Journal of Motor Behavior*, 53(2), 385–398. https://doi.org/10.1080/00222895.2020.1782825

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

Kim, T., Jimenez-Diaz, J., & Chen, J. (2017).The effect of attentional focus in balancing tasks: A systematicreview with meta-analysis. *Journal of Human Sport and Exercise*, 12(2). https://doi.org/10.14198/jhse.2017.122.22

Klimesch, W. (1996). Memory processes, brain oscillations and EEG synchronization. *International Journal of Psychophysiology*, *24*(1), 61–100. https://doi.org/10.1016/S0167-8760(96)00057-8

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, 4, 863. https://doi.org/10.3389/fpsyg.2013.00863

Lakens, D. (2014). Performing high-powered studies efficiently with sequential analyses. *European Journal of Social Psychology*, 44(7), 701–710. https://doi.org/10.1002/ejsp.2023

Lakens, D., Pahlke, F., & Wassmer, G. (2021). *Group Sequential Designs: A Tutorial*. PsyArXiv. https://doi.org/10.31234/osf.io/x4azm

Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence Testing for Psychological Research: A Tutorial. *Advances in Methods and Practices in Psychological Science*, *1*(2), 259–269. https://doi.org/10.1177/2515245918770963

Lee, T. D., Swinnen, S. P., & Serrien, D. J. (1994). Cognitive effort and motor learning. *Quest (grand Rapids, Mich)*, 46(3), 328–344. https://doi.org/10.1080/00336297.1994.10484130

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. https://doi.org/10.1016/j.humov.2016.08.005

Leiker, A. M., Pathania, A., Miller, M. W., & Lohse, K. R. (2019). Exploring the Neurophysiological Effects of Self-Controlled Practice in Motor Skill Learning. *Journal of Motor Learning and Development*, *7*(1), 13–34. https://doi.org/10.1123/jmld.2017-0051

Lessa, H. T., Tani, G., & Chiviacowsky, S. (2018). Benefits of enhanced expectancies through temporal-comparative feedback for motor learning in older adults. *International Journal of Sport Psychology*, 49(6), 521–530. https://doi.org/10.7352/IJSP.2018.49.521

Lewthwaite, R., & Wulf, G. (2010). Social-comparative feedback affects motor skill learning. *Quarterly Journal of Experimental Psychology*, 63(4), 738–749. https://doi.org/10.1080/17470210903111839

Lohse, K., Buchanan, T., & Miller, M. (2016). Underpowered and overworked: Problems with data analysis in motor learning studies. *Journal of Motor Learning and Development*, 4(1), 37–58. https://doi.org/10.1123/jmld.2015-0010

Lohse, K., Miller, M., Faria Braga Bacelar, M., & Krigolson, O. (2019). *Errors, rewards, and reinforcement in motor skill learning* (p. 39–60). https://doi.org/10.4324/9781351189750-3

Lohse, K. R., Jones, M., Healy, A. F., & Sherwood, D. E. (2014). The role of attention in motor control. *Journal of Experimental Psychology: General*, *143*, 930–948. https://doi.org/10.1037/a0032817

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

Lohse, K. R., & Sherwood, D. E. (2012). Thinking about muscles: The neuromuscular effects of attentional focus on accuracy and fatigue. *Acta Psychologica*, *140*(3), 236–245. https://doi.org/10.1016/j.actpsy.2012.05.009

Lohse, K. R., Sherwood, D. E., & Healy, A. F. (2010). How changing the focus of attention affects performance, kinematics, and electromyography in dart throwing. *Human Movement Science*, *29*(4), 542–555. https://doi.org/10.1016/j.humov.2010.05.001

Lutz, K., Pedroni, A., Nadig, K., Luechinger, R., & Jäncke, L. (2012). The rewarding value of good motor performance in the context of monetary incentives. *Neuropsychologia*, *50*(8), 1739–1747. https://doi.org/10.1016/j.neuropsychologia.2012.03.030

Margraf, L., Krause, D., & Weigelt, M. (2022). Valence-dependent Neural Correlates of Augmented Feedback Processing in Extensive Motor Sequence Learning - Part I: Practice-related Changes of Feedback Processing. *Neuroscience*, *486*, 4–19. https://doi.org/10.1016/j.neuroscience.2021.04.016

Makaruk, H., Starzak, M., & Marak Porter, J. (2020). Influence of attentional manipulation on jumping performance: A systematic review and meta-analysis. *Journal of Human Kinetics*, 75(1), 65–75. https://doi.org/10.2478/hukin-2020-0037

Masters, R., & Maxwell, J. (2008). The theory of reinvestment. *International Review of Sport and Exercise Psychology*, *1*(2), 160–183. https://doi.org/10.1080/17509840802287218

Maxwell, J. P., Masters, R. S. W., Kerr, E., & Weedon, E. (2001). The implicit benefit of learning without errors. *The Quarterly Journal of Experimental Psychology Section A*, *54*(4), 1049–1068. https://doi.org/10.1080/713756014

McAuley, E., Duncan, T., & Tammen, V. V. (1989). Psychometric Properties of the Intrinsic Motivation Inventory in a Competitive Sport Setting: A Confirmatory Factor Analysis. *Research Quarterly for Exercise and Sport*, *60*(1), 48–58. https://doi.org/10.1080/02701367.1989.10607413

McGuinness, L. A., & Higgins, J. P. T. (2021). Risk-of-bias VISualization (robvis): An R package and Shiny web app for visualizing risk-of-bias assessments. *Research Synthesis Methods*, *12*(1), 55–61. https://doi.org/10.1002/jrsm.1411

Meadows, C. C., Gable, P. A., Lohse, K. R., & Miller, 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

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & for the PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *BMJ*, 339(jul21 1), b2535–b2535. https://doi.org/10.1136/bmj.b2535

Moskowitz, J. B., Gale, D. J., Gallivan, J. P., Wolpert, D. M., & Flanagan, J. R. (2020). Human decision making anticipates future performance in motor learning. *PLOS Computational Biology*, *16*(2), e1007632. https://doi.org/10.1371/journal.pcbi.1007632

Navaee, S. A., Abedanzadeh, R., Salar, S., & Sharif, M. R. (2018). The Effects of Positive Normative Feedback on Learning a Throwing Task among Children with Autism Spectrum Disorder. *Nursing and Midwifery Studies*, *7*(2), Art. 2. https://doi.org/10.4103/nms.nms_71_17

Navaee, S. A., Farsi, A., & Abdoli, B. (2016). The effect of normative feedback on stability and efficacy of some selected muscles in a balancing task. *International Journal of Applied Exercise Physiology*, 5 (1), 43–52. Retrieved June 19, 2020, from http://www.ijaep.com/index.php/IJAE/article/view/46

Neubauer, A. C., & Fink, A. (2009). Intelligence and neural efficiency. *Neuroscience & Biobehavioral Reviews*, *33*(7), 1004–1023. https://doi.org/10.1016/j.neubiorev.2009.04.001

Neumann, D. L., & Thomas, P. R. (2008). A camera-based scoring system for evaluating performance accuracy during a golf putting task. *Behavior Research Methods*, *40*(3), 892–897. https://doi.org/10.3758/BRM.40.3.892

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

Ong, N. T., Hawke, J., & Hodges, N. J. (2019). Target Size Manipulations Affect Error-Processing Duration and Success Perceptions but not Behavioural Indices of Learning. *Brain Sciences*, *9*(5). https://doi.org/10.3390/brainsci9050119

Ong, N. T., & Hodges, N. J. (2018). Balancing our perceptions of the efficacy of success-based feedback manipulations on motor learning. *Journal of Motor Behavior*, 50(6), 614–630. https://doi.org/10.1080/00222895.2017.1383227

Ong, N. T., Lohse, K. R., & Hodges, N. J. (2015). Manipulating target size influences perceptions of success when learning a dart-throwing skill but does not impact retention. *Frontiers in Psychology*, *6*.

Oostenveld, R., & Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology*, *112*(4), 713–719. https://doi.org/10.1016/S1388-2457(00)00527-7

Palmer, K., Chiviacowsky, S., & Wulf, G. (2016). Enhanced expectancies facilitate golf putting. *Psychology of Sport and Exercise*, *22*, 229–232. https://doi.org/10.1016/j.psychsport.2015.08.009

Parma, J. O., Bacelar, M. F. B., Cabral, D. A. R., Lohse, K. R., Hodges, N. J., & Miller, M. W. (2023). That looks easy! Evidence against the benefits of an easier criterion of success for enhancing motor learning. *Psychology of Sport and Exercise*, 102394. https://doi.org/10.1016/j.psychsport.2023.102394

Parma, J. O., Bacelar, M. F. B., Cabral, D. A. R., Recker, R. S., Orsholits, D., Renaud, O., Sander, D., Krigolson, O. E., & Miller, M. W., Cheval, B., & Boisgontier, M. P. (in press). Relationship between reward-related brain activity and opportunities to sit. *Cortex*.

Pascua, L. A. M., Wulf, G., & Lewthwaite, R. (2015). Additive benefits of external focus and enhanced performance expectancy for motor learning. *Journal of Sports Sciences*, 33(1), 58–66. https://doi.org/10.1080/02640414.2014.922693

Percio, C. D., Infarinato, F., Iacoboni, M., Marzano, N., Soricelli, A., Aschieri, P., Eusebi, F., & Babiloni, C. (2010). Movement-related desynchronization of alpha rhythms is lower in athletes than non-athletes: A high-resolution EEG study. *Clinical Neurophysiology*, *121*(4), 482–491. https://doi.org/10.1016/j.clinph.2009.12.004

Pick, J. L., Nakagawa, S., & Noble, D. W. A. (2018). Reproducible, flexible and high-throughput data extraction from primary literature: The metaDigitise R package. *Evolutionary Biology*. https://doi.org/10.1101/247775

Porter, J. M., Wu, W. F. W., & Partridge, J. A. (2010). Focus of attention and verbal instructions: Strategies of elite track and field coaches and athletes. *Sport Science Review*, 19(3-4), 77–89. https://doi.org/10.2478/v10237-011-0018-7

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

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 and Exercise*, 13(4), 378–382. https://doi.org/10.1016/j.psychsport.2011.12.008

Saemi, E., Wulf, G., Varzaneh, A. G., & Zarghami, M. (2011). "Feedback" após boas versus más tentativas melhora a aprendizagem motora em crianças. *Revista Brasileira de Educação Física e Esporte*, 25(4), 673–681. https://doi.org/10.1590/S1807-55092011000400011

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

Sanli, E. A., Patterson, J. T., Bray, S. R., & Lee, T. D. (2013). Understanding self-controlled motor learning protocols through the self-determination theory. *Frontiers in Psychology*, 3, 611. https://doi.org/10.3389/fpsyg.2012.00611

Schmidt, L., Braun, E. K., Wager, T. D., & Shohamy, D. (2014). Mind matters: Placebo enhances reward learning in Parkinson's disease. *Nature Neuroscience*, *17*(12), Art. 12. https://doi.org/10.1038/nn.3842

Schmidt, R. A., & Lee, T. D. (2020). *Motor learning and performance: From principles to application* (6th ed.). Human Kinetics.

Schmidt, R. A., Lee, T. D., Winstein, C. J., Wulf, G., & Zelaznik, H. N. (2018). *Motor control and learning: A behavioral emphasis* (6th ed.). Human Kinetics.

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). p-curve and effect size: Correcting for publication bias using only significant results. *Perspectives on Psychological Science*, 9(6), 666–681. https://doi.org/10.1177/1745691614553988

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). https://doi.org/10.1038/srep36056

Ste-Marie, D. M., Vertes, K., Rymal, A. M., & Martini, R. (2011). Feedforward self-modeling enhances skill acquisition in children learning trampoline skills. *Frontiers in Psychology*, 2, 155. https://doi.org/10.3389/fpsyg.2011.00155

Sterne, J. A. C., Savović, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., Cates, C. J., Cheng, H.-Y., Corbett, M. S., Eldridge, S. M., Emberson, J. R., Hernán, M. A., Hopewell, S., Hróbjartsson, A., Junqueira, D. R., Jüni, P., Kirkham, J. J., Lasserson, T., Li, T.,…Higgins, J. P. T. (2019). Rob 2: A revised tool for assessing risk of bias in randomised trials. *BMJ*, l4898. https://doi.org/10.1136/ bmj.l4898

Sterne, J. A. C., Sutton, A. J., Ioannidis, J. P. A., Terrin, N., Jones, D. R., Lau, J., Carpenter, J., Rucker, G., Harbord, R. M., Schmid, C. H., Tetzlaff, J., Deeks, J. J., Peters, J., Macaskill, P., Schwarzer, G., Duval, S., Altman, D. G., Moher, D., & Higgins, J. P. T. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ*, 343, d4002–d4002. https://doi.org/10.1136/bmj.d4002

The Cochrane Collaboration. (2013). *Supplementary guidance for authors undertaking reviews with the cochrane consumers and communication review group*. http://cccrg.cochrane.org/sites/cccrg.cochrane.org/files/public/uploads/Study_design_guide2013.pdf

Trempe, M., Sabourin, M., & Proteau, L. (2012). Success modulates consolidation of a visuomotor adaptation task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(1), 52–60. https://doi.org/10.1037/a0024883

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36(1), 1–48. https://doi.org/10.18637/jss.v036.i03

Viechtbauer, W., & Cheung, M. W. L. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods*, 1(2), 112–125. https://doi.org/10.1002/jrsm.11

Vine, S., Moore, L., Cooke, A., Ring, C., & Wilson, M. (2013). Quiet eye training: A means to implicit motor learning Quiet eye training: A means to implicit motor learning. *International journal of sport psychology*, *44*, 367. https://doi.org/10.7352/IJSP2013.44.367

Wang, K.-P., Cheng, M.-Y., Chen, T.-T., Huang, C.-J., Schack, T., & Hung, T.-M. (2020). Elite golfers are characterized by psychomotor refinement in cognitive-motor processes. *Psychology of Sport and Exercise*, *50*, 101739. https://doi.org/10.1016/j.psychsport.2020.101739

Weinberg, A., Riesel, A., & Proudfit, G. H. (2014). Show me the Money: The impact of actual rewards and losses on the feedback negativity. *Brain and Cognition*, *87*, 134–139. https://doi.org/10.1016/j.bandc.2014.03.015

Wilhelm, R. A., Miller, M. W., & Gable, P. A. (2019). Neural and Attentional Correlates of Intrinsic Motivation Resulting from Social Performance Expectancy. *Neuroscience*, *416*, 137–146. https://doi.org/10.1016/j.neuroscience.2019.07.039

Williams, C. C., Hassall, C. D., Trska, R., Holroyd, C. B., & Krigolson, O. E. (2017). When theory and biology differ: The relationship between reward prediction errors and expectancy. *Biological Psychology*, *129*, 265–272. https://doi.org/10.1016/j.biopsycho.2017.09.007

Williams, C. C., Hecker, K. G., Paget, M. K., Coderre, S. P., Burak, K. W., Wright, B., & Krigolson, O. E. (2018). The application of reward learning in the real world: Changes in the reward positivity amplitude reflect learning in a medical education context. *International Journal of Psychophysiology*, *132*(Part B), 236–242. https://doi.org/10.1016/j.ijpsycho.2017.10.010

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

Wulf, G., Chiviacowsky, S., & Lewthwaite, R. (2010). Normative feedback effects on learning a timing task. *Research Quarterly for Exercise and Sport*, 81(4), 425–431. https://doi.org/10.1080/02701367.2010.10599703

Wulf, G., Chiviacowsky, S., & Lewthwaite, R. (2012). Altering mindset can enhance motor learning in older adults. *Psychology and Aging*, 27(1), 14–21. https://doi.org/10.1037/a0025718

Wulf, G., Freitas, H. E., & Tandy, R. D. (2014). Choosing to exercise more: Small choices increase exercise engagement. *Psychology of Sport and Exercise*, *15*(3), 268–271. https://doi.org/10.1016/j.psychsport.2014.01.007

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

Wulf, G., Lewthwaite, R., & Hooyman, A. (2013). Can ability conceptualizations alter the impact of social comparison in motor learning? *Journal of Motor Learning and Development*, 1(1), 20–30. https://doi.org/10.1123/jmld.1.1.20

Yamada, M., Diekfuss, J. A., & Raisbeck, L. D. (2020). Motor behavior literature fails to translate: A preliminary investigation into coaching and focus of attention in recreational distance runners. *International Journal of Exercise Science*, 3, 789–801.

Ziv, G., & Lidor, R. (2021). Different task success criteria affect expectancies of success but do not improve golf putting performance. *Psychology of Sport and Exercise*, *54*, 101887. https://doi.org/10.1016/j.psychsport.2021.101887

Ziv, G., Lidor, R., & Lavie, M. (2019). Enhanced expectancies in golf putting – A replication study with increased ecological validity. *International Journal of Sport and Exercise Psychology*, 19(1), 1–12. https://doi.org/10.1080/1612197X.2019.1637362

Ziv, G., Lidor, R., & Lavie, M. (2021). Enhanced expectancies in golf putting – a replication study with increased ecological validity. *International Journal of Sport and Exercise Psychology*, *19*(1), 1–12. https://doi.org/10.1080/1612197X.2019.1637362

Ziv, G., Ochayon, M., & Lidor, R. (2019). Enhanced or diminished expectancies in golf putting—Which actually affects performance? *Psychology of Sport and Exercise*, *40*, 82–86. https://doi.org/10.1016/j.psychsport.2018.10.003

Zobe, C., Krause, D., & Blischke, K. (2019). Dissociative effects of normative feedback on motor automaticity and motor accuracy in learning an arm movement sequence. *Human Movement Science*, 66, 529–540. https://doi.org/10.1016/j.humov.2019.06.004

APPENDIX A

Feedback Processing Models

Variables:

→ RewP = single-trial RewP amplitude

→ Success.c = contrast-coded success (in/out) in the current trial

→ Trial.c = linear contrast-coded trial (1-unit change = 10 trials)

→ Trial.c.sq = $(\text{Trial.c})^2$

→ Radial.error.c = group-mean-centered single-trial radial error

→ Radial.error.c.sq = $(\text{Radial.error.c})^2$

→ Zone.c = contrast-coded criterion of success (large zone/small zone)

→ SubID = participant identification

Model 1 - Effects of Perception of Success on the RewP

```
Model.1 = lme ( RewP ~

    # fixed-effects
        Success.c * Trial.c +
        Success.c * Trial.c.sq,

    # random-effects
        random = ~ 1 + Success.c + Trial.c + Trial.c.sq | SubID,
        method='ML', data = dat1)
```

Model 2 - Effects of Error Magnitude on the RewP

```
Model.2 = lme (RewP ~

    # fixed-effects
        Radial.error.c * Trial.c +
        Radial.error.c.sq * Trial.c +
        Radial.error.c * Trial.c.sq +
        Radial.error.c.sq * Trial.c.sq,

    # random-effects
        random = ~ 1 + Radial.error.c + Radial.error.c.sq + Trial.c + Trial.c.sq | SubID,
    method='ML', data = dat1)
```

Model 3 - Effects of Zone of Success Assignment and Error Magnitude on the RewP

```
Model.3 = lme (RewP ~

	# fixed-effects
		Trial.c * Radial.error.c * Zone.c +
		Trial.c.sq * Radial.error.c * Zone.c +
		Trial.c * Radial.error.c.sq * Zone.c +
		Trial.c.sq * Radial.error.c.sq * Zone.c,

	# random-effects
		random = ~ 1 + Radial.error.c + Radial.error.c.sq + Trial.c + Trial.c.sq | SubID,
		method='REML', data = dat1a)
```

APPENDIX B

Movement Preparation Models

Variables:

→ Alpha = single-trial motor upper-alpha power

→ Prior.Success.c = contrast-coded success (in/out) in the previous trial

→ Trial.c = linear contrast-coded trial (1-unit change = 10 trials)

→ Trial.c.sq = $(\text{Trial.c})^2$

→ Prior Radial.error.c = group-mean-centered radial error in the previous trial

→ Prior.Radial.error.c.sq = $(\text{Prior Radial.error.c})^2$

→ Prior.Radial.error.c.cb = $(\text{Prior Radial.error.c})^3$

→ Epoch = epoch (-3s - -2s, -2s - -1s, -1s - 0s, 0s - 1s)

→ Zone.c = contrast-coded criterion of success (large zone/small zone)

→ SubID = participant identification

Model 1 - Effects of Perception of Success on Motor-Upper Alpha

```
Model.1 = lmer (Alpha ~

        # fixed-effects
                Prior.Success.c * Trial.c * Epoch +
                Prior.Success.c * Trial.c.sq * Epoch +

        # random-effects
                (1 + Trial.c | SubID) + (1 | SubID:Epoch),
                REML= FALSE, dat2, optCtrl=list(maxfun=5e5)))
```

Model 2 - Effects of Error Magnitude on Motor-Upper Alpha

```
Model.2 = lmer (Alpha ~

        # fixed-effects
                Prior.Radial.Error.c * Trial.c * Epoch +
                Prior.Radial.Error.c * Trial.c.sq * Epoch +
                Prior.Radial.Error.c² * Trial.c * Epoch +
                Prior.Radial.Error.c² * Trial.c.sq * Epoch +
                Prior.Radial.Error.c³ * Trial.c * Epoch +
                Prior.Radial.Error.c³ * Trial.c.sq * Epoch +
```

```
    # random-effects
        (1 + Trial.c + Trial.c.sq + Prior.RE.c + Prior.RE.c.sq | SubID) +
        (1 | SubID:Epoch),
        REML= FALSE, dat2, optCtrl=list(maxfun=5e5)))
```

Model 3 - Effects of Zone of Success Assignment and Error Magnitude on Motor Upper Alpha

```
Model.3 = lmer (Alpha ~

    # fixed-effects
        Zone.c*Prior.Radial.Error.c * Trial.c * Epoch +
        Zone.c*Prior.Radial.Error.c * Trial.c.sq * Epoch +
        Zone.c*Prior.Radial.Error.c² * Trial.c * Epoch +
        Zone.c*Prior.Radial.Error.c² * Trial.c.sq * Epoch +
        Zone.c*Prior.Radial.Error.c³ * Trial.c * Epoch +
        Zone.c*Prior.Radial.Error.c³ * Trial.c.sq * Epoch +

    # random-effects
        (1 + Trial.c + Trial.c.sq + Prior.RE.c | SubID) +
        (1 | SubID:Epoch),
        REML= TRUE, dat2, optCtrl=list(maxfun=5e5)))
```

Post-Hoc Models - Effects of Error Magnitude, Trial, and Epoch on Motor Upper Alpha for each Zone

```
Model.3.Zone = lmer (Alpha ~

    # fixed-effects
        Prior.Radial.Error.c * Trial.c * Epoch +
        Prior.Radial.Error.c * Trial.c.sq * Epoch +
        Prior.Radial.Error.c² * Trial.c * Epoch +
        Prior.Radial.Error.c² * Trial.c.sq * Epoch +
        Prior.Radial.Error.c³ * Trial.c * Epoch +
        Prior.Radial.Error.c³ * Trial.c.sq * Epoch +

    # random-effects
        (1 + Trial.c + Prior.RE.c | SubID) +
        (1 | SubID:Epoch),
        REML= TRUE, dat2_Zone, optCtrl=list(maxfun=5e5)))
```