# The Effects of Implicit Learning, Practicing with the Expectation of Teaching, and Anxiety Training on Motor Performance Under Psychological Pressure

by

Daniel Aranha Rego Cabral

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
December 10, 2022

Keywords: expecting to teach, motor learning, choking under pressure, implicit learning, reinvestment theory

Approved by

Dr. Matthew Miller, Chair, Professor of Kinesiology
Dr. Mary Rudisill, Professor of Kinesiology
Dr. Danielle Wadsworth, Professor of Kinesiology
Dr. Jennifer Robinson, Professor of Psychological Sciences

Abstract

This dissertation describes a research program focused on investigating learning strategies that could prevent individuals from showing inferior motor performance when performing under psychological pressure, a phenomenon known as *choking under pressure*. Choking has been studied from the perspective of different theories. Each theory attempts to explain the choking phenomenon by different mechanisms, such as distraction, self-focus, and perceived threat. Collectively, this dissertation investigates strategies based on these mechanisms that could potentially prevent choking. The first two chapters focus on the reinvestment theory, which suggests that anxiety leads to the disruption of the skill's performance due to the individual reinvesting attentional resources in skill execution by using declarative knowledge acquired earlier during practice, which reduces movement automaticity and hinders performance. Theoretically, by limiting the amount of declarative knowledge acquired during skill acquisition (e.g., implicit learning), individuals may be able to maintain performance under pressure. In the first chapter, a meta-analytic approach was used to investigate whether employing an implicit learning strategy, as opposed to explicit learning, during the practice of a motor skill would prevent choking under pressure on subsequent low- and high-pressure post-tests. Results showed that participants who learned a motor skill implicitly performed better under a high-pressure condition than a low-pressure post-test, whereas participants in the comparison group performed similarly between conditions.

The second chapter is an attempt to test whether the benefits of practicing a motor skill with the expectation of teaching it to another person are preserved when the skill is performed under pressure if the skill is also practiced with an implicit learning strategy. This research follows prior findings that expecting to teach enhances skill learning, but also leads to choking

under pressure. My results revealed that participants who practiced with the expectation of teaching and used an implicit learning technique (i.e., analogy instructions) did not choke under pressure from low- to high-pressure post-test, whereas those who had the expectation of teaching and learned through a set of explicit instructions choked from low- to high-pressure.

The third chapter investigates the effect of anxiety training (AT; i.e., practicing a skill under enhanced anxiety levels) on subsequent performance during low-, mild- and high-pressure post-test. This chapter was based on three different theories of choking, namely the reinvestment theory, attentional control theory, and biopsychosocial mode challenge and threat. All of them are speculated to explain the benefit of AT to performance under pressure. Accordingly, I measured mechanistic variables related to each theory (i.e., movement reinvestment, mental effort, and perceived challenge/threat). Participants were assigned to either an AT or control group. The former practiced a motor skill under mild levels of pressure, whereas the latter practiced the same motor skill but with no pressure. Results showed that the AT group preserved performance across the post-tests, whereas the control group choked from low- to mild-, and from mild- to high-pressure post-tests. However, I did not find evidence that the mechanistic variables explained the relationship between AT and choking.

Taken together, the studies described in this dissertation show two potential interventions to prevent choking under pressure. First, as shown in the first two chapters, implicit learning seems to be an effective intervention to prevent choking (chapter 1) and to maintain the benefits of expecting to teach when performing under pressure (chapter 2). Second, practicing a motor skill under mild levels of anxiety is advantageous not only during a post-test with similar anxiety levels as the practice phase but also under a post-test with enhanced levels of anxiety, suggesting that the benefits of AT are transferrable to higher-stakes environments (chapter 3).

Acknowledgments

I would like to start by thanking my committee for their support and suggestions to improve my dissertation project. I would like to thank Dr. Miller for being an outstanding mentor, supervisor, and friend. I have no words to thank you enough. You can make sure I will do my best to be like you when mentoring my students in the future. You are an inspiration for how to treat students and how to extract the best of them. Not only that, but the way you taught us to do science will be with me for as long as I conduct research. Thank you! I would like to thank Dr. Rudisill for always having my back when I needed it (you know what I am talking about). Also, I want to say that I will miss our pre-pandemic lunches in the break room. Thank you for everything, Dr. Rudisill.

To my research assistants, thank you for helping me with tons and tons of data processing throughout the years, and thank you for the opportunity of mentoring you and I hope I have taught you something that will be useful in the future. To my labmates, thank you for making me a better researcher. I certainly learned a lot from you. To the whole Kinesiology family, students, faculty, and staff thank you for making my journey so pleasurable and much easier than what I was expecting when I first got to Auburn.

To my Brazilian friends, Paulo Mesquita and Juliana Parma, thank you for going through this journey with me. It was not easy, but you guys made it more smooth and more fun! To Powei and Young, the first friends I made in Auburn, thank you for all the moments we spent together playing basketball, watching movies, and hanging out. I will miss you guys. To Brady Rush, and the whole Rush family, I don’t think I have words to thank you enough. You have token me under your wing and made me feel like part of the family. I hope we meet again someday.

Now, going overseas, I would like to thank my parents, Aparecida and Junior. First, thank you for enrolling me in English classes when I was a kid. It totally paid off and I will surely do the same for my kids one day. Second, thank you for your amazing support throughout the years. I could not have done this without your help and you know that. Now I can see how hard you both worked in the past 30 years to give me the best education you could afford. I know that was a huge sacrifice and I know you do not regret doing it. You can be sure I did my best, so thank you and I love you!

To my best friend and brother, Matheus, I cannot thank you enough for your emotional support in the past years. Thank you for calling me several times per week; that meant a lot to me. Also, thank you for coming all the way from Northeast Brazil to Alabama to spend some time with me earlier this year. It was revitalizing and we had a lot of fun together. Love you, bro!

Finally, I would like to thank my wife, Malu, for her unconditional love and support, especially when we were separated during the pandemic. Only we know how hard it was, but we managed to stay together no matter what. Thank you for always having my back, making sure I was okay, and supporting me during this whole process. I know it was as hard on you as it was on me, but now we leave this chapter of our life stronger as a family and unbreakable as a couple. I hope to do the same for you in your new adventure as a graduate student. We have always been proud of how we push each other toward important accomplishments, rather than staying comfortable where we are. We have always pushed each other to pursue new challenges and new adventures, so I will never be able to thank you enough and I dedicate this victory to you. I love you!

Table of Contents

List of Tables

List of Figures

## Chapter 1: The effect of implicit learning on motor performance under psychological pressure: A systematic review and meta-analysis

### Introduction

Imagine your basketball team is down by one point with one second left in the game and you must shoot two free throws. Now imagine having to kick a penalty in a soccer match with your team down by one goal. Certainly, the ability to perform a motor skill under psychological pressure is crucial, yet people often choke under pressure, which has been defined as "the occurrence of inferior performance despite individual striving and situational demands for superior performance" [Baumeister, 1984, p. 610]). Choking occurs due to increased anxiety caused by individuals' perception that their capacity to perform the task falls short of the demands for superior performance (Baumeister, 1984; Baumeister & Showers, 1986; Mesagno & Hill, 2013). This increased anxiety triggers a chain of psychobiological responses that negatively impact performance, especially when the consequences of failing are believed to be large (Vine et al., 2016). Since almost all motor skills learned could be performed under pressure, it is important to search for learning strategies that maintain training benefits when the skill is performed under pressure.

Traditionally, motor skill learning begins with a verbal-cognitive stage, wherein the learner acquires a substantial amount of declarative knowledge about the skill and uses this knowledge to consciously control movement (Janacsek & Nemeth, 2013). Only after this phase would the learner be able to move to an autonomous stage and perform the skill with automaticity, largely reducing the use of declarative knowledge and conscious control. This two-stage process is known as *explicit* learning and reflects how motor skills are often taught. For example, an instructor may give a learner detailed step-by-step instructions about how to control their movements to perform a skill. Conversely, a person could learn a motor skill such that they bypass the verbal-cognitive stage, limiting the accrual of declarative knowledge and its use in conscious control of motor skills (Masters, 1992; Maxwell, Masters, & Eves, 2000). This type of learning is known as *implicit* learning, and an example is when an instructor uses an analogy to teach a learner how to perform a skill. For example, a basketball coach could instruct a learner to shoot a free throw as though they are trying to put cookies into a jar on a high shelf (Lam, Maxwell, & Masters, 2009). In this case, learners would accrue less declarative knowledge because the multiple task-related instructions would be reduced to a single comprehensive rule. Importantly, studies suggest that when a skill is learned implicitly, participants are less likely to choke under pressure compared to participants who learned the skill explicitly (Hardy, Mullen, & Jones, 1996; Liao & Masters, 2001; Masters, 1992; Vine, Moore, Cooke, Ring, & Wilson, 2013).

There are at least two explanations for why implicit learning has been shown to decrease the probability of choking under pressure. First, implicitly learned skills cause performers to lack declarative knowledge to reinvest in conscious control of movement when executing the skill, thus maintaining performance under pressure. This explanation is a direct corollary of reinvestment theory, which is related to a broader class of explicit monitoring theories (DeCaro et al., 2011). Reinvestment theory contends that dispositional and situational factors, such as psychological pressure, encourage individuals to use declarative knowledge acquired earlier in learning to attempt to consciously monitor and control practiced movements (Masters & Maxwell, 2008). This internal focus of attention paradoxically hinders movement execution due to inefficient muscle activation as well as rigid and uncorrelated effector movement (Lohse et al., 2011, 2014; Lohse & Sherwood, 2012). Second, implicitly learned skills may be performed with greater automaticity (fewer attentional resources) (Kal, Prosée, Winters, & van der Kamp, 2018), thereby allowing performance to be maintained when pressure causes attentional resources to be misallocated to irrelevant external stimuli (e.g., fans) and/or internal cues (e.g., worries about the task outcome), as predicted by distraction theories of choking under pressure (e.g., attentional control theory) (Eysenck et al., 2007; Jackson et al., 2006; Mesagno et al., 2015; Oudejans et al., 2011).

Studies have attempted to test the benefits of learning a skill implicitly vs. explicitly on high-pressure performance. However, results are inconsistent, thereby making conclusions hard to draw. For instance, Lam, Maxwell, and Masters (2009) had participants learn a free-throw basketball shot implicitly by an analogy instruction or explicitly by following a list of eight specific rules related to the biomechanics of the shot. The authors reported participants who practiced in the implicit group performed similarly in low-pressure and high-pressure post-test, whereas participants who practiced in the explicit group performed worse in the high-pressure compared to the low-pressure post-test (i.e., choked under pressure). Importantly, Lam et al. (2009) found participants in the explicit group reported more task-relevant rules than participants in the implicit group, thus revealing an association between declarative knowledge and choking under pressure. Conversely, Schücker, Hagemann, and Strauss (2013) had participants learn how to perform a golf putt either implicitly by an analogy instruction or explicitly by following a list of six specific putting rules. After the acquisition phase, participants performed the learned skill in a low-pressure post-test, followed by a high-pressure post-test and second low-pressure post-test. The authors reported no differences between the groups that learned implicitly vs. explicitly on the post-test, despite the explicit group reporting more task-relevant rules, thus failing to reveal an association between declarative knowledge and choking under pressure. This is in line with a recent systematic review that investigated possible strategies to prevent choking under pressure in sports, where the authors found inconsistent effects of implicit learning on choking under pressure (Gröpel & Mesagno, 2019).

Given these inconsistent findings, we conducted a systematic review and meta-analyses. Our systematic review and meta-analyses expand upon the systematic review by Gröpel and Mesagno (2019) in several ways. First, we provide a quantitative synthesis (meta-analyses) that estimates the magnitude of the effect of implicit learning to performance under pressure. Second, we explore the nature of the effect (e.g., whether it is due to explicit groups choking under pressure and/or implicit groups choking less, maintaining performance, or improving under pressure). Third, we statistically test for bias in the effect size estimate (e.g., using Egger et al. (1997)'s test). Finally, we focused on studies that verified implicit learning and psychological pressure occurred. We used random-effects models to estimate the magnitude of the effects of performing a motor skill under a low-pressure vs. high-pressure condition when the skill was learned either implicitly or more explicitly. We searched for articles that had human subjects learn a motor skill implicitly or explicitly and perform the skill under low- and high-pressure conditions. Consistent with theories of choking under pressure, particularly reinvestment theory, we predicted participants who learned a motor skill implicitly would demonstrate a change in skill performance from a low-pressure to a high-pressure condition that is less negative/more positive in comparison to participants who learn a skill more explicitly.

## Methods

This study followed the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines (Moher et al., 2009). The meta-analysis pre-registration, data, code and supplementary material are publicly available at https: https://osf.io/q8v6m/.

### Study Eligibility

Studies using human subjects were included with no restrictions regarding participants/population. Since there is no consensus on what interventions promote implicit learning, interventions were those the study authors referred to as “implicit” learning strategies and comparison groups were those not designated as being given “implicit” learning strategies (Kleynen et al., 2014). However, the authors were required to verify that the implicit learning group did indeed learn more implicitly than the comparison group. Specifically, authors needed to report the implicit learning group accrued significantly ($p \leq .05$) less declarative knowledge than the comparison group (e.g., in a free recall test of skill techniques). If a study had more than one implicit learning or comparison group, we used the implicit learning group showing the least amount of declarative knowledge accrual and the comparison group showing the most. We adopted this strategy to select intervention and comparison groups because implicit and explicit learning strategies are believed to occur along a continuum, with certain strategies leading to less or more accrual of declarative knowledge (Kleynen et al., 2015). The intervention and comparison groups must have practiced a task requiring movement to accomplish a goal that is increasingly likely to be achieved with practice (Schmidt & Lee, 2020). Outcomes were objective behavioral performance under low-pressure (these could include the last practice blocks and/or post-test) and high-pressure conditions, and the authors were required to demonstrate significantly ($p \leq .05$) increased pressure (e.g., self-reported anxiety, heart rate) in the high-pressure condition. If a study had multiple low- or high-pressure conditions, the first low- and high-pressure conditions were used. This was done to make the low- and high-pressure conditions more homogenous between studies, since most had only one low- and high-pressure condition. If multiple behavioral measures of performance were reported, the one most closely associated with accuracy (e.g., radial error as opposed to bivariate variable error [Hancock et al., 1995]) was used, since accuracy is most reflective of meeting the objective of a task (e.g., hitting a target).

**Search Strategy**

The search was conducted on (1) MEDLINE, (2) Web of Science, and (3) PsycINFO. The search strategy can be found in the supplementary material. In addition to these databases, we searched EBSCO open dissertations for unpublished research.[1] We only considered articles published in English. We also manually searched citations of the retrieved articles. The search was conducted until 29 June 2020. We did not specify a date range.

**Risk of Bias Assessment**

To assess risk of bias, we used the Cochrane's risk of bias tool 2.0 (Sterne et al., 2019, p. 2). Two authors (DARC and MWM) independently assessed five main categories of bias: randomization process, deviations from intended interventions, missing outcome data, measurements of the outcome, and selections of reported results. Each item was classified as high, low, or unclear risk of bias according to the Cochrane suggested algorithm for each individual study. However, we did not score one question of the deviations from intended interventions. Specifically, question 2.1 ("Were participants aware of their assigned group during the experiment?") is not particularly useful when it comes to the interventions often seen in the motor learning field because participants from different groups receive group-specific instructions from the experimenters, which makes participants aware of their group assignment and is an inherent risk of bias. Additionally, it was expected that most studies would be classified as having an unclear or high risk of bias in multiple categories due to standard practices in motor learning studies. For example, it is rare for studies to report detailed randomization procedures or the blinding of researchers from participants' group assignment, which would be necessary for a study to be considered low risk in the randomization process and deviations from intended interventions categories, respectively. Thus, the risk of bias assessment aimed to compare the

[1] We planned to only include studies that had been published in a peer-review journal, and we ended up doing so after not returning any unpublished studies that met our criteria.

studies with each other, rather than classify each study as having an overall high, low, or unclear risk of bias. Specifically, we were interested in identifying any bias that could categorize studies within the review and serve as a moderator variable in the meta-analyses. For example, if approximately half the studies were considered low risk in the selections of reported results category due to pre-registering their analyses, we could examine whether being low risk in this category moderated meta-analytic results. This strategy was inspired by a recent meta-analysis examining the effect of training with enhanced anxiety levels for performance under pressure (Low et al., 2021).

**Data Extraction**

Data were extracted and organized into an Excel sheet. To characterize each study, we extracted sample size, population (e.g., healthy adults, athletes), participant mean age, gender, the motor skill used, implicit (e.g., errorless, analogy, dual-task) and comparison (e.g., rule-based or biomechanical instructions, internal focus of attention, errorful learning) group learning interventions, pressure manipulation (e.g., monetary incentive, video recording, social evaluation), number of trials and blocks in the pretest (if the article reported it), number of trials and blocks in the acquisition phase, if the high-pressure condition was ≥24h after the training phase or if it was in the same day as the acquisition phase, performance variables (e.g., radial error, accuracy, distance, time), and variables related to the pressure manipulation (e.g., heart rate, anxiety levels). For the meta-analyses, we extracted means and standard deviations of motor performance, and sample size. If data were missing, we contacted authors to obtain it. If authors were unable to provide data, they were extracted from figures using ImageJ (Rueden et al., 2017). When an experiment failed to show data necessary to determine standard deviations, we used the function "*impute_SD*" from the R package "*metagear*" (Lajeunesse, 2016), with the "*Bracken1992*" method (Bracken, 1992). This method fills missing information using the coefficient of variation from all studies with complete information (Bracken, 1992). We used this method for three studies (Hardy et al., 1996; Liao & Masters, 2001; Masters, 1992).

**Data Synthesis and Analysis**

To calculate effect sizes, we used the sample sizes of the implicit and comparison groups as well as their means and standard deviations for the low- and high-pressure conditions. Consistent with Lakens (2013), we calculated Cohen's $d_s$, Hedges' $g_{av}$, and Hedges' $g_{rm}$, to measure the change in performance from the low- to high-pressure condition for the implicit learning groups and comparison groups and chose to use Hedges' $g_{av}$ because it was closer to Cohen's $d_s$ than Hedges' $g_{rm}$ was.[2]

Three main meta-analyses were conducted. First, we performed a multivariate random-effects meta-analysis (Berkey, Hoaglin, Antczak-Bouckoms, Mosteller, & Colditz, 1998) with group (implicit vs. comparison) serving as a moderator, study serving as a random effect, and Hedges' $g_{av}$ serving as the outcome variable. Comparison group Hedges' $g_{av}$ was set as the intercept. (As sensitivity analyses, we repeated this meta-analysis with four additional studies that partially met our inclusion criteria [see supplementary material; Table S2]). We followed this meta-analysis with two univariate (traditional) random-effects meta-analyses, one for the implicit learning group and one for the comparison group, with Hedges' $g_{av}$ serving as the outcome variable in both analyses. For each meta-analysis, the $Q$-test was used to assess heterogeneity across studies (Cochran, 1954). For the univariate random-effects meta-analyses, $H^2$ and $I^2$ was used to assess consistency across studies (Cochran, 1954), and bias was assessed

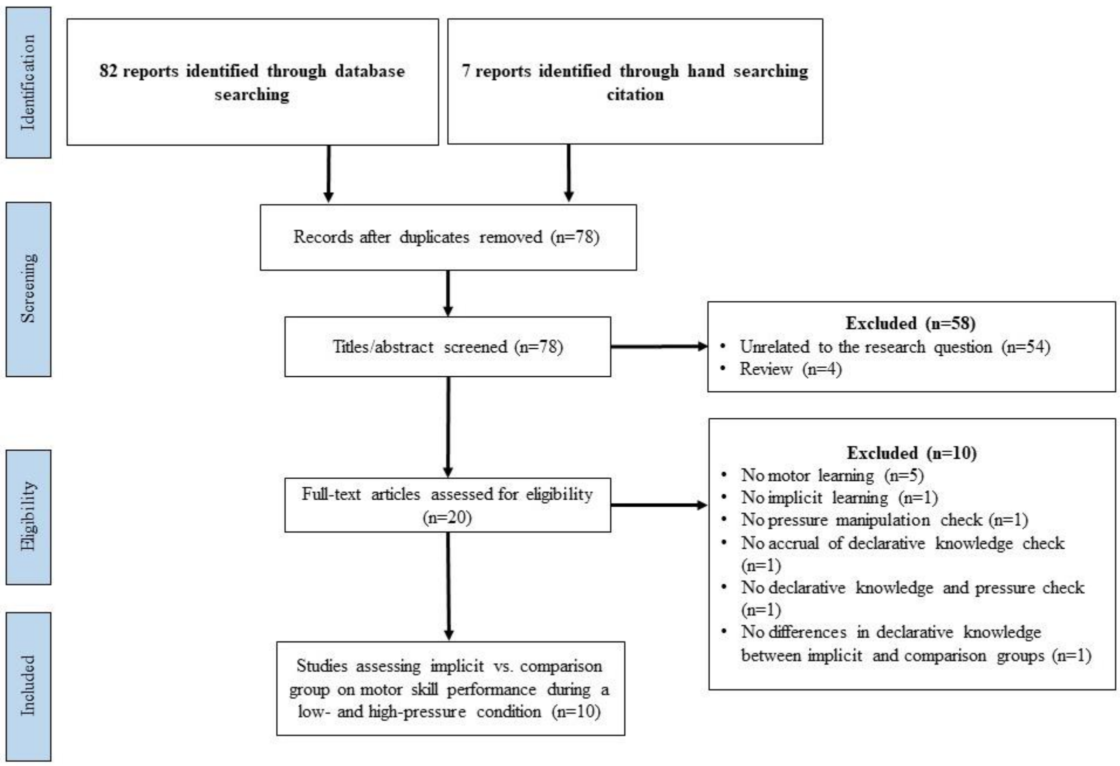
[2] In the pre-registration, we mistakenly implied that correlation coefficients between low- and high-pressure post-test are needed to calculate Hedges' $g_{av}$, but they are only needed to calculate Hedges' $g_{rm}$. We also departed from our pre-registration by using Lakens (2013)'s strategy to determine whether to use Hedges' $g_{av}$ or Hedges' $g_{rm}$. (Our method for determining correlation coefficients to calculate Hedges'$g_{rm}$ is described in the pre-registration.) This strategy led us to choose Hedges' $g_{av}$, which is what we had pre-registered

by visually inspecting funnel plots of effect sizes and standard errors as well as conducting Egger's regressions, where the standard error serves as a predictor of the effect size across studies (Egger et al., 1997). An asymmetrical funnel plot and significant association between standard error and effect size in Egger's regression suggests that there is bias, whereas a symmetrical plot and a non-significant relationship suggests that there is no bias. To check for outliers and influential cases, we examined Cook's distances ($D_i$). Specifically, if $D_i$ of a given study was ≥ .5 and visually distinguished from the $D_i$ of the other studies, it was considered an influential case (Glen, 2016). If any influential cases were detected, we removed them from analyses and reported sensitivity analyses with the case(s) included in supplementary material. Analyses were conducted using the "*metafor*" package for R (Viechtbauer, 2010). The alpha level was set at .05.

## Results

Ten studies met the criteria to be included in the primary meta-analyses. Figure 1 shows the process of article screening. Eighty-nine articles were identified through the systematic search and hand searching citations. After duplicates were removed, 78 articles were screened, and 58 of those were excluded because they were not related to the research question or were review articles, leaving a total of 20 articles. From those, five were excluded for not involving a motor skill, one for not involving implicit learning, and four for not confirming implicit learning and/or changes in psychological pressure (see Figure 1 for detail). Table 1 depicts the descriptive statistics and other relevant information, such as whether the study reported a significant Group (implicit learning vs. comparison) x Condition (low-pressure vs. high-pressure) interaction (6 of 10 studies reported a significant interaction and all significant interactions were large [Cohen, 1962]). Of the 10 studies meeting the criteria to be included in the primary meta-analyses, one (Bright & Freedman, 1998) was identified as an influential case ($D_i$ = 0.56 in the univariate random-effects meta-analysis for the implicit learning group; see Figures S1 and S3 in supplementary materials) and was removed from further analyses. All meta-analyses were repeated with this study included, and results from these analyses can be found in supplementary materials.

**Figure 1.** PRISMA Flow Chart.




## Risk of Bias

Table 1 provides an overview of risk of bias for each included study in the main and exploratory analyses. All studies had an unclear risk of bias in the randomization domain due to an absence of providing detailed descriptions of randomization procedures. Similarly, an unclear risk of bias was attributed to all studies in the deviations from intended intervention domain due to not indicating whether experimenters were aware of participants' group assignment. Lastly, nine of ten studies were classified as having an unclear risk of bias (one had a high risk of bias) in the selection of reported result domain due to not reporting the presence of pre-specified analysis plans. Overall, studies had low concerns in the missing outcome data and measurement of the outcome domains due to no indications of missing data and the use of objective outcome measures, respectively. Since the studies had similar risk of bias in each domain, we did not create any risk of bias factors to serve as moderating variables in the meta-analyses.

**Table 1.**

*Summary of Studies in the Main and Exploratory Analyses*

**Studies included in the main analysis**

| **Study** | **Motor skill** | **Comparison of motor learning strategies** | **Number of trials in acquisition phase** | **Population** | **Sample size per group** | **Age (years)** | **Number of trials and blocks in the low-pressure and high-pressure conditions** | **Pressure manipulation in the high-pressure post-tes** | **Reported a statistically significant Group (implicit and comparison) x condition (low- and high-pressure) interaction (Y/N)** | **Effect size of Group x condition interaction** |
|---|---|---|---|---|---|---|---|---|---|---|
| Masters (1992) | Golf putting | Dual task vs. technical instructions | 3 sessions of 100 trials | Novice | 8 | 27.22 (18-46) | High-pressure - 2 blocks of 50 trials; low pressure - 2 blocks of 50 trials* | Chance to win or lose money and evaluation by a golf expert | Y | Not reported |
| Hardy et al. (1996) | Golf putting | Dual task vs. technical instructions | 3 sessions of 100 trials | College students | 8 | 21.13 | High-pressure - 2 blocks of 50 trials; low-pressure - 2 blocks of 50 trials* | Chance to win or lose money and evaluation by a golf expert | Y | Not reported |
| Bright & Freedman (1998) | Golf putting | Dual-task vs. technical instructions | 3 sessions of 40 trials | Novice College students | 16 | Not reported | High pressure - 1 block of 40 trials; low-pressure: 1 block of 40 trials* | Performance evaluation by a professional golfer; video recording | Y | Not reported |
| Liao & Masters (2001) | Table tennis | Analogy vs. technical instructions | 5 blocks of 50 trials | Novice College students | 9 | 21.5±4.1 | High-pressure - 1 block of 50 trials; low-pressure – 1 block of 50 trials | Ego-treatment feedback | N | $\eta^2_p$ = .43 |
| Koedijker et al. (2007) | Table tennis | Analogy vs. technical instructions | 9 blocks of 50 trials | Novice | 8 | 21.80±3.58 | High-pressure - 1 block of 50 trials; low-pressure – 1 block of 50 trials | Video recording and chance to win a prize | Y | $f$ = .56 |
| Koedijker et al. (2008) | Table tennis | Analogy vs. technical instructions | 14 blocks of 100 trials | Novice | Implicit ($n$ = 7); comparison ($n$ = 8) | 19.6±3.4 | High-pressure - 1 block of 50 trials; low-pressure – 1 block of 50 trials | Chance to win a prize, experimenter counted score out loud and video recording | N | $f$ = .43 |
| Lam et al. (2009) | Basketball shots | Analogy vs. technical instructions | 2 days – 6 blocks of 30 trials per day | Novice College students | 12 | Implicit (21.08±1.16); comparison (21.92±1,73) | High-pressure - 1 block of 40 trials; low-pressure – 1 block of 40 trials | Performance evaluation by a basketball expert; chance to win money based on performance | Y | $\eta^2_p$ = .16 |
| Zhu et al. (2011) | Golf putting | Errorless vs. errorful | 6 blocks of 50 trials | Novice College students | 9 | 22.05±2.26 | High-pressure - 1 block of 10 trials; low-pressure – 1 block of 10 trials | Video recording | N | Not reported |
| Vine et al. (2013) | Golf putting | Analogy vs. technical instructions | 8 blocks of 40 trials in three sessions during a 7-day period (two blocks in session 1, and 3 blocks in sessions 2 and 3) | Novice College students | 15 | 21.22±4.41 | High-pressure - 1 block of 20 trials; low-pressure – 1 block of 20 trials | Chance to win money; comparison to other participants | Y | $\eta^2_p$ = .14 |
| Schücker et al. (2013) | Golf putting | Analogy vs. technical instructions | 6 blocks of 50 trials | Novice College students | Implicit ($n$ = 20); comparison ($n$ = 21) | 21.44±2.98 | High-pressure - 1 block of 20 trials; low-pressure – 1 block of 20 trials | Chance to win money | N | Not reported |

Note: *Low-pressure was considered as the last block of the acquisition phase.

## Descriptive Synthesis

### *Sample Characteristics*

The sample size of the 10 included studies was 226 (112 individuals in the implicit group and 114 in the comparison group). Seven studies reported that the sample was composed of college students (Bright & Freedman, 1998; Hardy et al., 1996; Lam et al., 2009; Liao & Masters, 2001; Schücker et al., 2013; Vine et al., 2013; Zhu et al., 2011), whereas the remaining three did not report any particular characteristic of their sample besides age (Koedijker et al., 2007, 2008; Masters, 1992) (see Table 2). The study samples were fairly similar regarding age, with the lowest mean age (19.6±3.4) reported by Koedijker et al. (2008), and the highest (27.22; min = 18; max = 46) reported by Masters (1992). The majority of participants was female ($n$ = 157, relative to $n$ = 81 males), although three studies did not report participants' gender (Masters, 1992; Vine et al., 2013; Zhu et al., 2011). As far as participants' prior experience with the skill, only one study did not report participants' skill level prior to the acquisition phase (Hardy et al., 1996), although the authors checked for performance differences in the first five putts of the practice phase between groups and did not find any, thus suggesting the groups were at the same skill level prior to practice. One study conducted the same analysis, although it stated that participants had no prior experience with the skill (Masters, 1992). Despite the remaining eight studies reporting that participants were novice to the motor skill, only two tested whether performance prior to the acquisition phase was different between groups (Lam et al., 2009; Vine et al., 2013), and both found no differences in baseline performance.

**Table 2.**
*Risk of Bias Assessment Results*

| Studies | Randomization | Deviations from intended interventions | Missing data | Outcome measurement | Selection of reported results |
|---|---|---|---|---|---|
| Masters, 1992 | Unclear | Unclear | Low | Low | Unclear |

| | | | | | |
|---|---|---|---|---|---|
| Hardy, 1996 | Unclear | Unclear | Low | Low | Unclear |
| Bright & Freedman, 1998 | Unclear | Unclear | Low | Low | Unclear |
| Liao & Masters, 2001 | Unclear | Unclear | Low | Low | Unclear |
| Koedjiker et al., 2007 | Unclear | Unclear | Low | Low | High |
| Koedjiker et al., 2008 | Unclear | Unclear | Low | Low | Unclear |
| Lam et al., 2009 | Unclear | Unclear | Low | Unclear | Unclear |
| Zhu et al., 2011 | Unclear | Unclear | Low | Low | Unclear |
| Schucker et al., 2013 | Unclear | Unclear | Low | Low | Unclear |
| Vine et al., 2013 | Unclear | Unclear | Low | Low | Unclear |

### *Studies Characteristics*

Across the included studies, three motor skills were used, namely golf ($k$ = 6) (Bright & Freedman, 1998; Hardy et al., 1996; Masters, 1992; Schücker et al., 2013; Vine et al., 2013; Zhu et al., 2011), table tennis ($k$ = 3) (Koedijker et al., 2007, 2008; Liao & Masters, 2001), and basketball free throws ($k$ = 1) (Lam et al., 2009).

Eight studies analyzed the number of reported explicit rules as a measure of declarative knowledge (Bright & Freedman, 1998; Hardy et al., 1996; Lam et al., 2009; Liao & Masters, 2001; Masters, 1992; Schücker et al., 2013; Vine et al., 2013; Zhu et al., 2011), and two analyzed participants' detailed written description of how they executed the skill. In particular, the authors scored the description as internal (e.g., related to the movement) or external (e.g., related to the environment) acquired rules (Koedijker et al., 2007, 2008). While both rule types reflect explicit knowledge, studies have found that an external focus of attention increases skill automaticity, whereas an internal focus of attention is associated with greater explicit processing and conscious control of movement (Kal, van der Kamp, & Houdijk, 2013; Poolton, Maxwell, Masters, & Raab, 2006).

As far as pressure manipulation, studies induced pressure through reward contingency (*k* = 7) (Hardy et al., 1996; Koedijker et al., 2007, 2008; Lam et al., 2009; Masters, 1992; Schücker et al., 2013), videotaping (*k* = 5) (Bright & Freedman, 1998; Koedijker et al., 2007, 2008; Lam et al., 2009; Zhu et al., 2011), social comparison (*k* = 2) (Schücker et al., 2013; Vine et al., 2013), performance evaluation (*k* = 4) (Bright & Freedman, 1998; Hardy et al., 1996; Masters, 1992; Zhu et al., 2011), and/or ego-threating feedback (*k* = 2) (Liao & Masters, 2001; Vine et al., 2013). (Note that some studies used multiple manipulations.) To check effectiveness of the pressure manipulation, studies relied on a combination of heart rate and self-reported anxiety (*k* = 5) (Bright & Freedman, 1998; Hardy et al., 1996; Lam et al., 2009; Masters, 1992; Zhu et al., 2011) or only self-reported anxiety (*k* = 5) (Koedijker et al., 2007, 2008; Liao & Masters, 2001; Schücker et al., 2013; Vine et al., 2013).

***Performance Outcome***

As far as performance outcome under low- and high-pressure conditions, four studies revealed that the implicit group maintained performance from low- to high-pressure, whereas the comparison group reduced performance (Lam et al., 2009; Liao & Masters, 2001; Masters, 1992; Vine et al., 2013); two reported that the implicit group improved from low- to high-pressure, while the comparison group maintained performance (Bright & Freedman, 1998; Hardy et al., 1996); one revealed that the implicit group performance increased from low- to high-pressure, while the comparison group decreased (Koedijker et al., 2007); one showed that the implicit group had superior performance in both low- and high-pressure, relative to the comparison group, but neither group's performance changed from low- to high-pressure (Zhu et al., 2011); and two revealed that both groups performed equally in both conditions, with neither changing from low-to high-pressure (Koedijker et al., 2007; Schücker et al., 2013). Overall, seven of ten studies showed that learning implicitly is beneficial for performance under pressure, whereas three studies revealed no performance changes from low- to high-pressure for the implicit or comparison group.

**Multivariate Random-Effects Meta-Analysis**

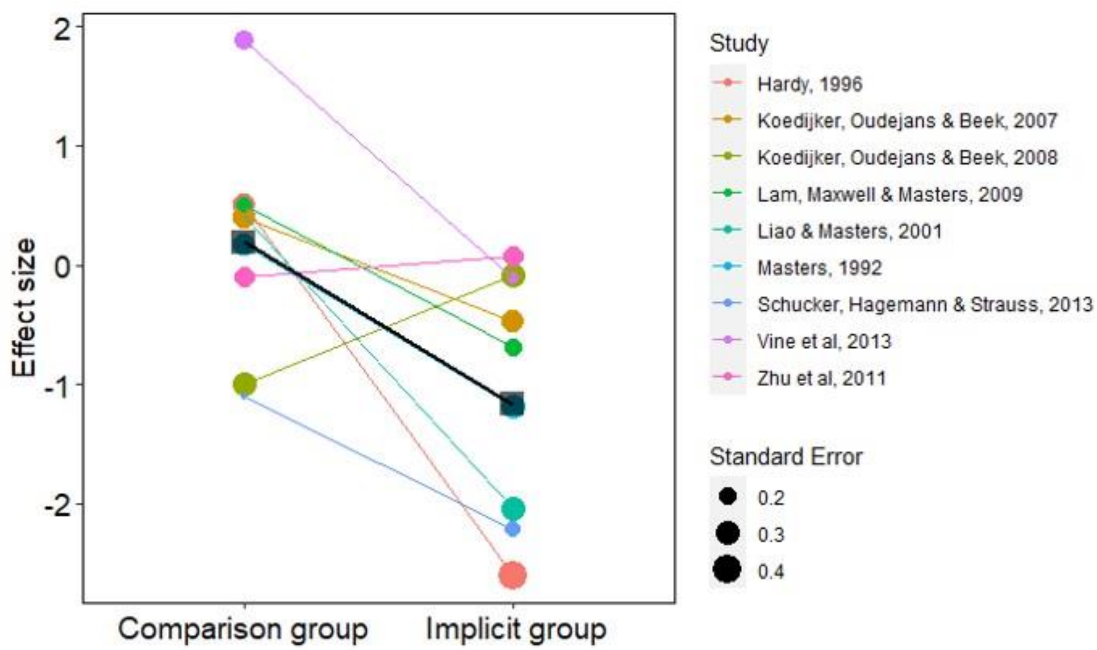
Since Hedges' $g_{av}$ was calculated as low-pressure performance minus high-pressure performance, negative values indicate better performance under high-pressure. Results revealed that group was a significant moderator ($QM(1) = 27.98$, $p < .001$), with the implicit group demonstrating improved performance from low- to high-pressure (Hedges' $g_{av} = -1.17$, 95% lower $CI = -1.61$, upper $CI = -0.74$, $p < .001$) relative to the comparison group which did not significantly change from low- to high-pressure (Hedges' $g_{av} = 0.19$, 95% lower $CI = -0.33$, upper $CI = 0.71$, $p = .483$) (see Figure 2). Results also revealed the presence of heterogeneity ($QE$ (16) $= 68.92$, $p < .001$).

It is worth considering whether the implicit group's significantly greater improvement from the low- to high-pressure condition relative to the comparison group is driven by the implicit group being significantly worse under low-pressure than the comparison group (the case in four of nine studies), which would give the implicit group significantly more room for improvement under high-pressure. To address this question, we conducted a sensitivity analysis with the five studies that did *not* show a significant group difference.[3] Results demonstrated that group remained a significant moderator ($QM(1) = 5.15$, $p = .023$), with the implicit group still exhibiting improved performance from low- to high-pressure (Hedges' $g_{av} = -0.70$, 95% lower $CI = -1.30$, upper $CI = -0.10$, $p = .023$) relative to the comparison group which did not significantly

[3] We thank an anonymous reviewer for suggesting this sensitivity analysis. See Figure S2 for a forest plot depicting the results of a meta-analysis on low-pressure performance as a function of group.

change from low- to high-pressure (Hedges' $g_{av}$ = 0.11, 95% lower *CI* = -0.32, upper *CI* = 0.53, *p* = .624). Results also revealed the presence of heterogeneity (*QE* (8) = 15.80, *p* = .045). For comparison, we repeated this analysis with studies that *did* show a significant group difference. Results demonstrated that group was a significant moderator (*QM*(1) = 28.19, *p* < .001), with the implicit group exhibiting improved performance from low- to high-pressure (Hedges' $g_{av}$ = -1.69, 95% lower *CI* = -2.32, upper *CI* = -1.07, *p* < .001) relative to the comparison group which did not significantly change from low- to high-pressure (Hedges' $g_{av}$ = 0.28, 95% lower *CI* = -0.84, upper *CI* = 1.40, *p* = .624). Results also revealed the presence of significant heterogeneity (*QE* (6) = 48.51, *p* < .001).

**Figure 2.**
*Univariate Random-Effects Meta-Analyses*

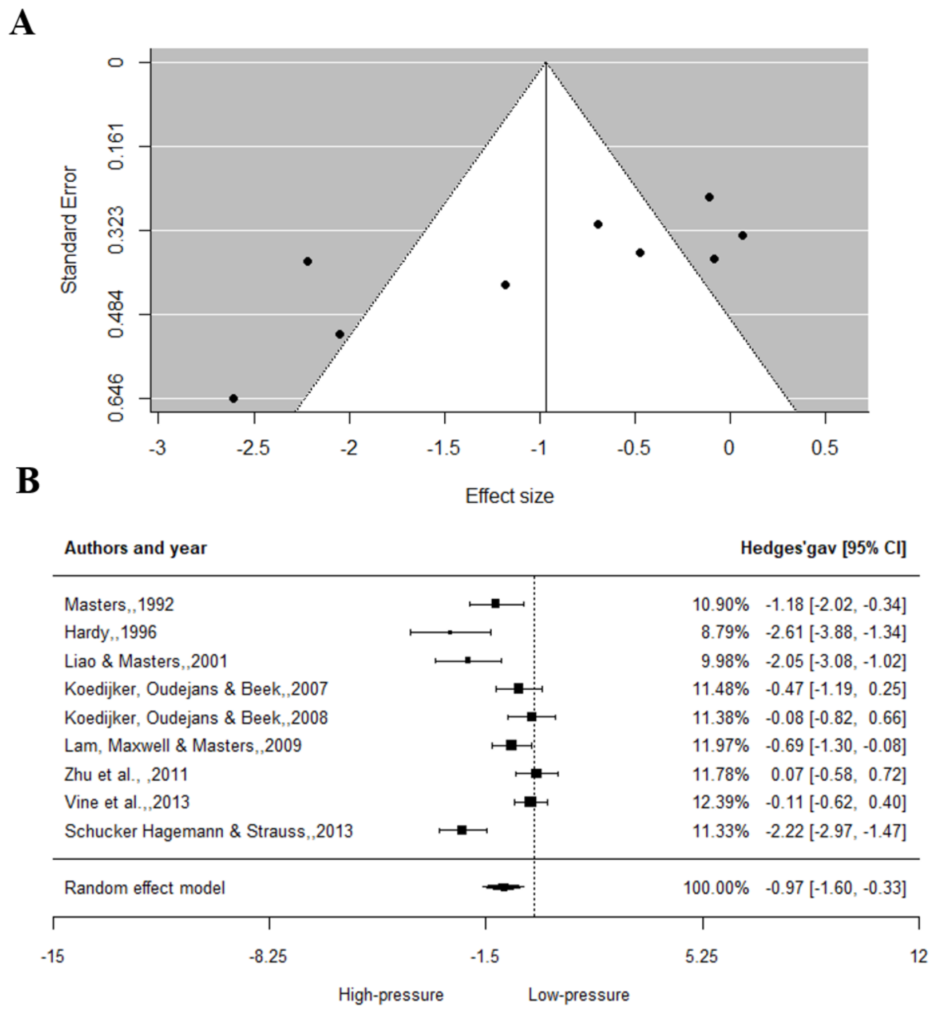



*Note:* Plot contrasting implicit and comparison group effect sizes (Hedges' $g_{av}$) across studies. Negative values on the y-axis indicate better performance under high- vs. low-pressure, and positive values indicate better performance under low- vs. high-pressure. The square markers represent the average effect size for each group. The circles are colored by study and weighted by standard error with larger circles depicting greater standard error.

### *Implicit Learning Group*

Figure 3 shows the forest and funnel plots for the implicit group univariate random-effects meta-analysis. Consistent with the multivariate model, the univariate model revealed that participants who learned a motor skill implicitly performed better in a high-pressure condition compared to a low-pressure condition (Hedges' $g_{av}$ = -0.97, 95% lower $CI$ = -1.60, upper $CI$ = -0.33, $p$ = .003). In addition, results revealed the presence of high heterogeneity ($H^2$ = 6.69, $I^2$ = 85.06%; test for heterogeneity ($Q(8)$ = 46.00, $p$ < .001). The funnel plot (Figure 3A) suggests an asymmetrical distribution, and the Egger's regression test was significant ($Z$ = -3.13, $p$ = .002), indicating bias.

**Figure 3.**
*Funnel and Forest Plots for the Implicit Learning Group Performance During Low- vs. High-Pressure*



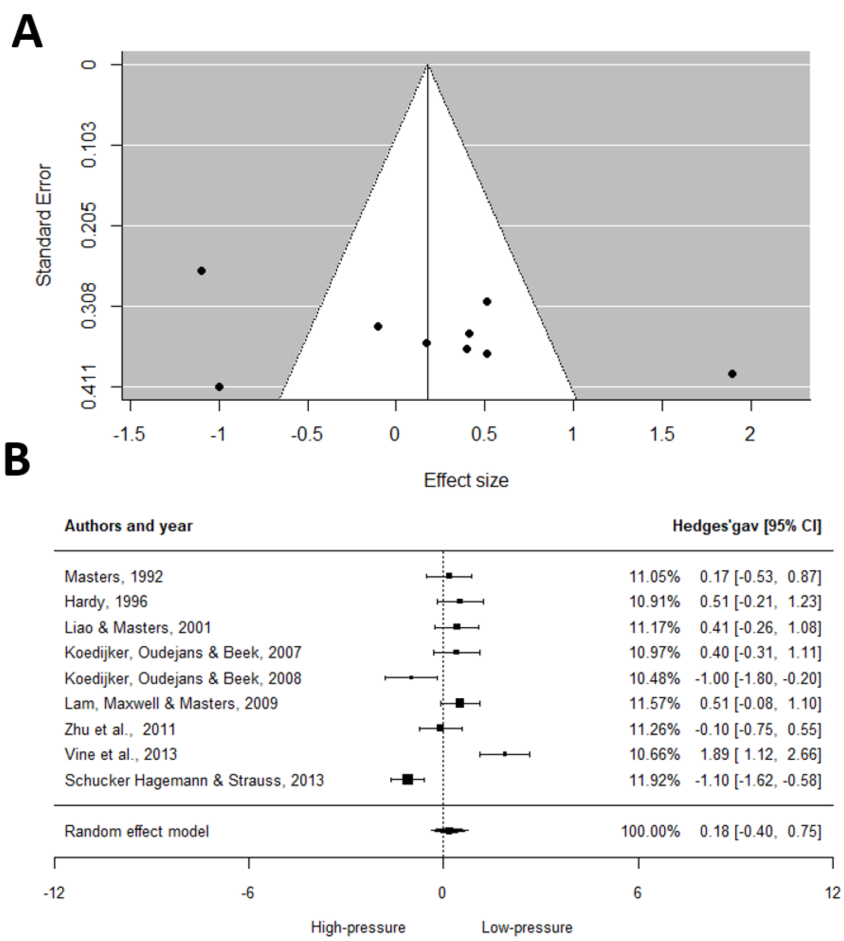

*Note:* (A) Funnel plot for the implicit learning group performance during low- vs. high-pressure conditions relating effect size with the standard error in each study. A negative effect size (left side) means that the implicit group performed better under the high-pressure condition, whereas a positive effect size (right side) means that the implicit group performed better in a low-pressure condition. (B) Forest plot showing the effect sizes and 95% confidence intervals for each study and the summary effect size from the univariate random-effects model. The data markers are weighted by standard error and the percentages represent the weight of each study in the model.

### *Comparison Group*

Figure 4 shows the forest and funnel plots for the comparison group univariate random-effects meta-analysis. Consistent with the multivariate model, the univariate model revealed that participants who learned a motor skill more explicitly performed similar in low- and high-pressure conditions (Hedges' $g_{av}$ = 0.18, 95% lower *CI* = -0.40, upper *CI* = 0.75, *p* = .546). In addition, results revealed the presence of high heterogeneity ($H^2$ = 6.63, $I^2$ = 84.92%; test for heterogeneity ($Q(8) = 53.66$, $p < .001$). The funnel plot (Figure 3A) suggests a symmetrical distribution, and the Egger's regression test was not significant ($Z = 1.07$, $p = .284$), indicating no evidence bias.

**Figure 4.**

*Funnel and Forest Plots for the Comparison Group Performance During Low- vs. High-Pressure*




*Note:* (A) Funnel plot for the comparison group performance during low- vs. high-pressure conditions relating effect size with the standard error in each study. A negative effect size (left side) means that the comparison group performed better under the high-pressure condition, whereas a positive effect size (right side) means that the comparison group performed better in a low-pressure condition. (B) Forest plot showing the effect sizes and 95% confidence intervals for each study and the summary effect size from the univariate random-effects model. The data markers are weighted by standard error and the percentages represent the weight of each study in the model.

**Multivariate Random-Effects Meta-Analysis with Number of Trials During Practice as a Moderator**

Finally, we conducted another multivariate, random-effects meta-analysis with group, number of trials during the practice phase of each study, and the interaction of these two variables serving as moderators, study serving as a random effect, and Hedges' $g_{av}$ serving as the outcome variable. Comparison group Hedges' $g_{av}$ was set as the intercept. We included number of trials during practice because it is possible that more practice trials could lead to greater accrual of declarative knowledge, particularly in the comparison groups, due to participants having more opportunities to test hypotheses about errors. In this case, comparison groups with more practice trials should exhibit greater performance decrement under high-pressure, consequently increasing the benefit of implicit learning. Conversely, it is possible more practice trials could lead to a greater accrual of declarative knowledge in the implicit group, as evident when comparing Masters (1992) with Maxwell, Masters, and Eves (2000). Both studies had participants practice golf putting in an implicit learning condition wherein they performed a secondary task, but Masters had participants practice 400 putts whereas Maxwell et al. had participants practice 3000 putts. Maxwell et al.'s implicit learning participants acquired two explicit rules, whereas Masters' acquired one, suggesting that more practice trials may cause implicitly trained participants to accrue more explicit knowledge.

The test of the moderators was significant ($QM(3) = 36.69$, $p < .001$, with the implicit group exhibiting improved performance from low- to high-pressure (Hedges' $g_{av}$ = -2.08 95% lower $CI$ = -2.84, upper $CI$ = -1.33, $p < .001$) relative to the comparison group which did not significantly change from low- to high-pressure (Hedges' $g_{av}$ = 0.60 95% lower $CI$ = -0.36upper $CI$ = 1.54, $p = .169$). Importantly, the Group x Number of Trials During Practice interaction was significant (Hedges' $g_{av}$ = 0.002, 95% lower $CI$ = 0.001, upper $CI$ = 0.004, $p = .003$, indicating the benefit of implicit learning relative to the comparison group was significantly smaller in studies with more practice trials. Results also revealed the presence of heterogeneity ($QE$ (14) = 61.67, $p < .001$).

## Discussion

The aim of the present systematic review and meta-analysis was to investigate the prediction that individuals learning a motor skill implicitly demonstrate a change in skill performance from a low-pressure to a high-pressure condition that is less negative/more positive in comparison to individuals who learn a skill more explicitly. Results from the systematic review revealed an advantage of learning a motor skill implicitly when performing under psychological pressure for seven of ten studies. In general, the studies had unclear risks of bias in the randomization process, deviations from intended interventions, and selection of reported results domains, and low risks of bias in the missing outcome data and measurements of the outcome domains. Importantly, the unclear risks of bias may be due to authors not reporting procedures rather than not undertaking them (The Cochrane Collaboration, 2013). For example, authors may have conducted the randomization process consistent with a low risk of bias, but only reported that participants were 'randomly assigned' to different groups, which is not enough information for the process to be considered low risk. Such brief reporting may be due to factors such as limited word space or unawareness of the value of reporting detailed information. The increased opportunities to share methodological details such as randomization procedures and pre-registered analysis plans as online supplements and calls for researchers to do so (Lohse et al., 2016) may encourage them to report procedures in more detail moving forward. However, other aspects of motor learning research are inherent to the discipline and unlikely to change, including participant awareness of group assignment, which increases risk, and using objective outcome measures, which lowers risk.

Results from the multivariate random-effects meta-analysis revealed that participants who learned a motor skill implicitly improve their skill performance from a low-pressure to a high-pressure condition relative to participants who learned a skill more explicitly and perform similarly under low- and high-pressure (Figure 2). These effects are robust to sensitivity analyses when an outlying/influential study is included or four studies that failed to meet some of the inclusion criteria are added (Bellomo et al., 2018; Bobrownicki et al., 2015; Lam et al., 2009; Schücker et al., 2010) supplementary material). Additionally, these effects are moderated by the number of practice trials such that the implicit learning group advantage is significantly attenuated in studies with more practice trials.

It is worth noting the meta-analyses reveal the implicit learning group benefit is due to their improvement under pressure, not the comparison groups choking under pressure, and the systematic review shows that five of ten studies report the implicit learning group maintained/improved performance under pressure while the comparison group choked. The result makes sense considering the methods used to investigate the implicit learning advantage. Specifically, most studies (Bright & Freedman, 1998; Koedijker et al., 2007, 2008; Lam et al., 2009; Liao & Masters, 2001; Masters et al., 2020; Schücker et al., 2013; Zhu et al., 2011) used novices and had them complete a few hundred practice trials (median = 300), probably leaving them ample room for improvement during low- and high-pressure conditions totaling nearly 100 trials on average (median = 90). As the low-pressure condition preceded the high-pressure condition in all studies, implicit group participants would be expected to improve from the low- to high-pressure condition, as they continued to increase their performance while being shielded from the negative effects of psychological pressure. Conversely, comparison group participants should have this improvement thwarted by psychological pressure, thus showing little change in performance. Interestingly, the implicit group improvement from low- to high-pressure is descriptively stronger in studies where the implicit group was significantly worse than the comparison group during low-pressure, suggesting the implicit groups in these studies may have had their learning constrained, thereby affording them even more room for improvement in the high-pressure condition. Notably, the study wherein the implicit group exhibited the largest improvement (Hardy et al., 1996) had a low-pressure condition wherein participants were practicing a skill while performing a secondary task, which is thought to restrict learning (Masters & Poolton, 2012).

The implicit learning group's attenuated advantage in studies with more practice trials may be due to increased practice trials providing them more opportunities to make errors and test hypotheses about how to correct errors, resulting in accrual of declarative knowledge and, thus, the attenuation of their performance advantage under pressure. This explanation is consistent with the comparison between Maxwell et al. (2000) and Masters (1992), which shows that implicitly trained participants who had more practice trials accrued more declarative knowledge (Maxwell et al.) relative to participants who had fewer trials (Masters). Alternatively, the implicit group's fading benefit as a function of practice trials may be due to the comparison group achieving greater automaticity with more trials (Kal et al., 2018). It is also worth noting that, with more practice trials, participants may become better able to report declarative knowledge use while not necessarily using this knowledge for motor control. Irrespective of the explanation for the moderating effect of practice trials on implicit learning's benefit under high-pressure, there is a need for more studies with varying numbers of practice trials to determine whether it is indeed the case that the implicit learning advantage under pressure is attenuated when skills are acquired with many practice trials, as is the case in real-world training settings.

Although systematic review and meta-analysis results confirm implicit learning's advantage for pressurized performance, meta-analysis results indicate bias with respect to the implicit learning groups. Specifically, a funnel plot reveals an asymmetrical distribution of the effect of performing a skill in low- vs. high-pressure conditions (Figure 3A) and a significant negative relationship between effect size and precision for the implicit learning groups. The funnel plot's asymmetry seems to be due to an absence of studies with low precision that fail to show a benefit for high- vs. low-pressure (i.e., a lack of studies in the lower-right corner of Figure 3A), which suggests reporting bias (e.g., publication bias, selective outcome and/or analysis reporting) (Sterne et al., 2011). For the explicit groups, the funnel plot is symmetrical, resembling a horizontal cylinder (Figure 4A) reflecting large heterogeneity (Sterne et al., 2011), and does not show a significant relationship between effect size and precision. Taken together, meta-analysis results suggest findings from relatively imprecise studies where the implicit group exhibits large performance improvement under high-pressure are preferentially reported.

Results also reveal median sample size per group for studies included in the systematic review and meta-analysis is 9, which is below the median sample size per group (11) reported in a recent meta-analysis of motor learning studies (Lohse, Buchanan, & Miller, 2016) where the authors critiqued the sample sizes as being small and, thus, the studies as being underpowered. Small sample sizes are associated with poor study quality and large effect sizes (Sterne et al., 2011), and underpowered studies have a lower probability of finding a real effect, lower positive predictive value (i.e., lower probability that a positive result is real), and a high likelihood of overestimating the magnitude of a real effect (Ioannidis, 2005). Thus, future studies should improve power by increasing sample size. (Notably, power can also be increased by strengthening experimental manipulations, thus providing an incentive to bolster pressure manipulations.) To estimate the sample size required to test the implicit learning advantage under pressure, we used G*Power 3.1.9.4 (Faul et al., 2009) to conduct an *a priori* power analysis for a hypothetical experiment where researchers want to test for a Group (implicit learning vs.

comparison) x Condition (low- vs. high-pressure) interaction. Specifically, we selected the ANOVA: Repeated measures, within-between interaction test; assumed common alpha (.05) and beta (.20) values; set number of groups and measurements equal to 2 (2 groups measured under low- and high-pressure); set the correlation among repeated measures to .50; and set the nonsphericity epsilon to 1 (since sphericity cannot be violated in the case of two repeated measures). Assuming a medium effect size ($f = .25$), using $f$ as in Cohen (1988) as recommended by G*Power 3.1.9.4., the required sample size per group is 64, more than seven times the median sample size used in the studies included in the primary meta-analysis. Thus, studies in the present meta-analysis are underpowered, unless the real Group x Condition interaction effect is large ($f \geq .71$) (Cohen, 1962). Although the present meta-analysis indicates a large effect (a large improvement from low- to high-pressure for the implicit learning group and a negligible decrease for the comparison group), results also suggest preferential reporting of findings from studies demonstrating large improvements for the implicit group. Therefore, it may be more reasonable to assume a medium effect, and, of course, the real effect size may be $f < .25$, so researchers may want to power their studies to detect smaller effects. Since the calculated sample size will be a large increase 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 evidence of bias and underpowered studies found in the implicit learning-performance under pressure literature reflects these problems in the broader motor learning literature (Lohse et al., 2016), where some hypotheses from highly cited theories have not been confirmed by pre-registered and well-powered studies (Grand et al., 2017; McKay & Ste-Marie, 2020). Like these hypotheses, the prediction that implicit learning is advantageous for performance under pressure requires more rigorous testing before implicit learning should be recommended to practitioners (e.g., athletes, coaches). To this point, a recent systematic review found inconsistent evidence that analogy and other implicit learning strategies prevent choking under pressure (Gröpel & Mesagno, 2019). Thus, the hypothesis that implicit learning is beneficial for performance under pressure requires more thorough testing. Specifically, we recommend the hypothesis be tested with well-powered studies and pre-registered experiments or registered reports containing a priori hypotheses specifying outcome variables and analysis plans, which will reduce reporting bias (Caldwell et al., 2020).

Besides the evidence suggesting the reviewed studies were underpowered and may have been selectively reported, some had other limitations. Specifically, a few studies did not indicate participants' skill level (Masters, 1992), did not use a criterion to define it (Hardy et al., 1996), or determined it based on a lax criterion (Bright & Freedman, 1998). Since the nature of the implicit learning benefit under pressure may be shaped by participants' skill level (see earlier discussion), it is important to specify it. The fact that no studies counterbalanced the order of the pressure conditions is also relevant to the nature of the implicit learning advantage, because always having the low-pressure condition precede the high-pressure condition may have allowed the implicit learning group participants to improve performance under the latter. Counterbalancing the order of conditions would have avoided this issue, but it could have raised others, such as having participants' anxiety in a high-pressure condition carry over to the low-pressure condition.

The present study had some limitations. First, it included a small number of studies ($k$ = 10), partially due to the exclusion of those that did not report successfully manipulating psychological pressure and/or declarative knowledge. Notably, the addition of these studies did not substantially influence the effect of pressure condition on performance of the implicit learning or comparison groups. Nonetheless, the small $k$ limits the precision with which we can estimate group differences in the effect of pressure condition on performance. Additionally, the small $k$ precludes examining between-study factors that may moderate the effect of implicit learning on performance under pressure or explain variance in performance under pressure unexplained by group assignment. For example, we were unable to investigate whether effects were moderated by the percentage of males/females in a study, which would have been valuable to assess given that gender affects performance under pressure (Englert & Seiler, 2020). More generally, the individual zones of optimal functioning model predicts that some athletes perform worse under pressure, but some perform better, meaning that there are potential individual differences that may moderate the effect of implicit learning on performance under pressure and that these differences should be considered in future research (Hanin & Syrjä, 1995a, 1995b). The small *k* also means the asymmetrical funnel plot and negative relationship between effect size and precision for the implicit learning groups could be due to chance rather than bias (Sterne et al., 2011). The second limitation is that we had to impute missing data for some of the studies. Specifically, 3 of 10 studies did not report performance standard deviations (Hardy et al., 1996; Liao & Masters, 2001; Masters, 1992), so we imputed this data with the coefficient of variation from the other 7 studies. This technique assumes the standard deviation is associated with the other studies' coefficients of variation (Bracken, 1992). We do acknowledge, though, that all studies in the primary analysis were more than 5 years old and the authors may have discarded study data. Nonetheless, we highlight the need for future studies to provide their data set and/or correlation matrix in open platforms to facilitate future meta-analyses. Taking the present study's limitations together, meta-analyses' effect size and evidence of bias results are limited by the small *k*, but we believe assumptions about standard deviations coefficients used to calculate individual study effect sizes are reasonable and unlikely to have been very influential in the meta-analyses' results.

## Conclusion

In conclusion, results confirm the prediction that implicit motor learning benefits performance under pressure, with the benefit due to implicit learning improving performance under pressure. Crucially, sample size assessment suggests studies are likely underpowered, and bias evaluation indicates findings from relatively imprecise studies where the implicit group exhibits large performance improvement under high-pressure are preferentially reported. Taken together, the present study's findings reveal some evidence for the prediction that implicit learning benefits performance under pressure. To provide stronger evidence, researchers are encouraged to conduct more rigorous pre-registered studies with larger sample sizes and stronger pressure manipulations as well as report all data needed for future meta-analyses.

## Chapter 2: Does learning a skill with the expectation of teaching it impair the skill's execution under psychological pressure if the skill is learned with analogy instructions?

### Introduction

Determining practice conditions that enhance motor learning is important to facilitate motor behavior. The value of practice conditions that enhance motor learning depends on whether the learning benefits are transferred to novel contexts (Schmidt & Lee, 2019), particularly those likely to be encountered while performing the skill and those with high importance. For example, a practice condition may improve a learner's encoding and consolidation of a skill, however the practice condition's efficacy is limited if the skill cannot be successfully retrieved and performed in high-stakes environments, under psychological pressure. As many skills must be performed in high-stakes environments, such as sports competition, it is crucial to determine practice conditions that enhance learning and preserve learning benefits under psychological pressure. Recently, Daou, Hutchison et al. (2019) revealed that practicing a motor skill with the expectation of teaching it to another person loses its benefit when the learned skill is performed under psychological pressure. Therefore, the purpose of the present study was to determine whether the expecting to teach approach can be modified to preserve its learning advantage, and, in so doing, shed light on the mechanisms underlying the loss of the benefit under psychological pressure.

Some initial research of the expecting to teach approach showed that when participants study academic information with the expectation of teaching it, they exhibit augmented learning (Bargh & Schul, 1980; Benware & Deci, 1984; Nestojko, Bui, Kornell, & Bjork; 2014). However, other studies failed to reveal this effect (Renkl, 1995; Ross & Di Vesta, 1976) or demonstrated ambiguous learning effects (enhancements on short-term, but not long-term, test performance; Fiorella & Mayer, 2013; Fiorella & Mayer, 2014). Daou, Buchanan, Lindsey, Lohse, and Miller (2016) conducted the first investigation into whether expecting to teach enhances learning of motor skills, which rely more heavily on procedural knowledge than academic information does (Rosenbaum, Carlson, & Gilmore, 2001). Daou, Buchanan et al. observed having learners practice and study a motor skill with the expectation of teaching it to another person enhanced skill learning in comparison to having learners practice and study a skill with the expectation of being tested, and this effect has been replicated several times (Daou, Hutchison et al., 2019; Daou, Lohse, & Miller, 2016; Daou, Lohse, & Miller, 2018; Daou, Rhoads, Jacobs, Lohse, & Miller, 2019). Although research has failed to reveal the mechanisms underlying the learning benefit of expecting to teach, studies have consistently shown that the learning advantage occurs concomitant to large gains in declarative knowledge about the learned skill (Daou et al., 2018; Daou, Buchanan et al., 2016; Daou, Hutchison et al., 2019; Daou, Lohse et al., 2016; Daou, Rhoads et al., 2019). As motor skills acquired with large gains in declarative knowledge are highly susceptible to decrement under psychological pressure (Hardy, Mullen, & Jones, 1996; Koedijker, Oudejans, & Beek, 2007; Lam, Maxwell, & Masters, 2009a, 2009b; Liao & Masters, 2001; Masters, 1992), it was unsurprising that Daou, Hutchison et al. (2019) revealed that the expecting to teach benefit vanished under psychological pressure, due to participants who practiced with the expectation of teaching 'choking' in a high-pressure post-tes. Daou, Hutchison et al. concluded that participants who practiced with the expectation of teaching choked likely due to their accrual of declarative knowledge while practicing, however the authors were unable to provide evidence to support this conclusion. Nonetheless, their conclusion is consistent with reinvestment theory (Masters & Maxwell, 2008), which contends that dispositional and situational factors, such as psychological pressure, trigger individuals to use declarative knowledge acquired earlier in learning to attempt to consciously monitor and control practiced movements. This focus of attention on movement, ironically, impairs performance (Wulf, 2013).

Critically, learners who accrue more declarative knowledge during skill practice are more likely to exhibit performance decrement under pressure, because they have more declarative knowledge to 'reinvest' in motor control.

A corollary of reinvestment theory is that motor skills learned relatively implicitly, with minimal gains in declarative knowledge, should be resilient to psychological pressure (Maxwell & Masters, 2008), and research generally supports this proposition (Hardy et al., 1996; Koedijker et al., 2007; Lam et al., 2009a, 2009b; Liao & Masters, 2001; Masters, 1992). An effective strategy to encourage implicit motor learning is to provide learners with an analogy about how to perform the skill rather than explicit rules, strategies, and techniques regarding skill performance (Lam et al., 2009a, 2009b; Liao & Masters, 2001). With an analogy, declarative knowledge about multiple rules is reduced into a single, comprehensive rule. For example, Lam et al. (2009a) instructed participants in an analogy practice condition to "shoot as if you are trying to put cookies into a cookie jar on a high shelf" (p.344) while practicing a basketball free throw, whereas participants in an explicit practice condition group were instructed to follow a list of eight specific rules while practicing. Participants in both conditions performed low- and high-pressure post-test and were asked to recall free-throw shooting rules. Participants who practiced in the analogy condition reported fewer rules, indicative of more implicit learning, and performed equally well under low- and high-pressure post-test, whereas the explicit condition group performed worse under the high-pressure than the low-pressure post-test (i.e., they choked under pressure).

Since the expecting to teach approach is a practical way to enhance motor learning, it would be beneficial to determine a way to maintain the learning advantage under psychological pressure. As Daou, Hutchison et al. (2019) attributed the choking effect exhibited by participants who practiced with the expectation of teaching to the accrual of declarative knowledge, a promising means to prevent the choking effect is to promote implicit learning by instructing learners to use an analogy to practice a motor skill rather than a list of rules. Indeed, Daou, Hutchison et al. asked participants to study an instruction booklet containing a list of rules to follow while practicing the skill, likely prompting learners who expected to teach to attend to the rules so that they could disseminate them to another person; an analogy instruction would reduce this attention to rules. Importantly, it is unlikely that minimizing the accrual of declarative knowledge by learners who expect to teach will reduce their learning advantage, as declarative knowledge has been found to not significantly relate to motor learning in an expecting to teach paradigm (Daou, Buchanan et al., 2016). Even with analogy instructions, it is possible that learners who expect to teach could accrue greater declarative knowledge by engaging in more learning activities, such as discovery learning and hypothesis testing, than those who expect to test. Nonetheless, the practical question of whether the choking effect associated with the expecting to teach approach is prevented by using analogy instructions can still be answered.

The present study investigated whether having learners practice a motor skill with the expectation of teaching it and using an analogy to practice it preserves the learning advantage of expecting to teach under psychological pressure. Specifically, participants were assigned to four groups. One group practiced with the expectation of teaching a motor skill and received an analogy instruction (teach/analogy); one group practiced with the expectation of teaching the skill and received specific explicit rules related to the skill (teach/explicit); one group practiced with the expectation of being tested on the skill and received an analogy instruction (test/analogy); and one group practiced with the expectation of being tested on the skill and received explicit rules about the skill (test/explicit). One day after skill practice (6 blocks of 10 putts on a single day), all groups performed low- and high-pressure post-test. With this 2 (Expectation: teach/test) x 2 (Instruction: analogy/explicit) x 2 (Post-test: low-pressure/high-pressure) design, we predicted a

3-way interaction. In particular, we predicted participants in the teach groups would exhibit superior post-test performance on the low-pressure post-test relative to their test group counterparts, but the effect of expecting to teach on performance in the high-pressure post-test would be moderated by instruction. Specifically, expecting to teach would be advantageous for participants who trained with analogy instructions, but not for participants who trained with explicit instructions. This result would indicate that practicing a motor skill with the expectation of teaching and an analogy imparts a learning advantage that can be manifested in a high-stakes environment. Crucially, this result would also strongly suggest that the reason learners who practice with the expectation of teaching choke under pressure is due to their accrual of declarative knowledge while practicing, thus addressing a shortcoming of Daou, Hutchison et al. (2019). However, if the choking effect associated with the expecting to teach approach is not prevented by employing analogy instructions, this would not eliminate the possibility that the choking effect is caused by the accrual of declarative knowledge. Specifically, learners who practice with the expectation of teaching may accrue a relatively large amount of declarative knowledge despite receiving analogy instructions and use this knowledge during the high-pressure post-test. Importantly, we can use free recall tests of declarative knowledge use during the high-pressure post-test to shed light on this possibility.

## Methods

The study's data, code, questionnaires, putting instructions and supplementary material are publicly available at https://osf.io/vpr92/.

### Sample

Men and women between the ages of 18 and 30 years participated in the study and could receive course credit for participation. This demographic was convenient to the investigators and has been used in similar past studies (e.g., Daou, Hutchison et al., 2019). Participants must had putted (anything from playing miniature golf to playing 18 holes on a standard golf course) between one and thirty times in their lifetime and not more than twenty times in the past year. Participants with this amount of experience were most sensitive to the expecting to teach and pressure manipulations in past experiments (Daou, Hutchison et al., 2019). Since these participants had at least minimal putting experience, the instructions and practice likely afforded them an opportunity to improve their skill by internalizing the analogy/explicit rules rather than guiding them through a completely novel movement, and participants who expected to teach may especially have taken advantage of this opportunity in preparation for their teaching episode. Participants must have been free from physical illness, injury, or disability that could make putting difficult. Participants were asked to refrain from alcohol/drug consumption within 24 hr of both days of the study, caffeine consumption within 3 hr of both days of the study, and to get a good amount of sleep the night before each day of the study while also trying to get the same amount of sleep each night.

**Sample Size Calculation**

Since our study was novel, we were unable to estimate the effect size for an Expectation x Instructions x Post-test interaction. Thus, we powered our study to detect an Expectation x Post-test interaction, which we estimated to be medium ($\eta^2_p$ = .093) in our sample, based on the effect size observed in our past research among participants who met the inclusion and outlier criteria for the present study (Daou, Hutchison et al., 2019). We powered the study to detect this interaction because the given instructions to the participants should moderate it to a relatively large degree, based on past research investigating the effects of instructions on post-test performance (Lam et al., 2009a; Liao & Masters, 2001). To do the power analyses, we used G*Power 3.1.9.2 (Faul, Erdfelder, Lang, & Buchner, 2007) and entered the aforementioned effect size (as in SPSS) along with the following parameters: $\alpha = .05$, power = .9, number of groups = 4, number of measurements = 2, and nonsphericity correction $\varepsilon = 1$ into an ANOVA with repeated-measures testing for a within-between interaction. The required sample size was determined to be 148, but we decided to collect data from 164 participants to account for data loss (e.g., participant dropout, problems with data entry). We determined the final sample submitted to statistical analysis would include at least 148 participants, and we would ensure equal *n* in each group (by recruiting additional participants, if necessary). In terms of data exclusion, we decided to only exclude participants if there was a technical error in recording their putts at the post-test or if one of their average low- or high-pressure post-test radial error values had a *z*-score > 3.00. In the latter case, we decided to report the primary statistical results with and without the inclusion of the participant.

**Task**

All participants used a standard (88.9 cm) golf putter to putt a standard golf ball from a starting position indicated by a 5 cm line painted in white washable paint on an artificial grass surface to a target cross (+) comprised of two 10.8 cm lines painted in white washable paint and located 300 cm away from the starting position (Daou, Hutchison et al., 2019). Participants' objective was to make the ball stop as close to the center of the target as possible.

**Procedure**

All participants completed the experiment individually. After consenting to the experiment, participants completed a demographics questionnaire asking their sex, age, putting experience, any illness, injury, or disability that could make putting difficult, whether they consumed alcohol/drugs within the last 24 hr, whether they consumed caffeine within the last 3 hr, and how long they slept the previous night. (All questionnaires can be found on the Open Science Framework (https://osf.io/vpr92/), see Questionnaires and Instructions component.) Once the experimenter confirmed that the participants met the inclusion criteria (see Sample section), participants completed the Movement Specific Reinvestment Scale (Masters, Eves, & Maxwell, 2005). The Movement Specific Reinvestment Scale is frequently used to examine individual tendencies to reinvest in motor control (Huffman, Horslen, Carpenter, & Adkin, 2009; Kal, van der Kamp, Houdijk, Groet, Bennekom, & Sherder, 2015; Klämpfl, Lobinger, & Raab, 2013; Malhorta, Poolton, Wilson, Ngo, & Masters, 2012; Vine, Moore, Cooke, Ring, & Wilson, 2013) and possesses good psychometric properties (Masters et al., 2005). The Movement Specific Reinvestment Scale consists of the conscious motor processing and movement self-consciousness subscales, which ask participants to indicate how strongly they agree with statements related to their tendency to attempt to control their movements and monitor their movements, respectively. Participants responded on a 6-point scale anchored by "strongly disagree" and "strongly agree". The Movement Specific Reinvestment Scale data were collected because it was possible that they would be used to explore whether individual tendencies toward reinvestment explained residual variance in the model, thus increasing the amount of variance explained by the other factors in the model, as was the case in Daou, Hutchison et al. (2019). Next, participants put a physiological monitoring device around their chest (BioHarness 3.0, Zephyr Technology, Annapolis, MD) to get used to wearing it, which they were asked to do the following day as well. Physiological data such as heart rate and heart rate variability were collected because it was possible that they would be extracted from the device for supplemental/exploratory analyses.

***Pretest***

After completing the demographics questionnaire, participants performed the pretest phase, which consisted of one block of ten putts.

***Practice***

After pretest, participants were quasi-randomly assigned (based on sex) to the teach/analogy, teach/explicit, test/analogy, or test/explicit groups, and the corresponding expectation manipulation occurred. Participants in the teach groups were told, “Tomorrow you will teach another participant how to putt,” and participants in the test group were told, “Tomorrow you will be tested on your putting skills.” Next, the instruction manipulation occurred. Participants in the analogy groups read the following: “Keep your body still like a grandfather clock and use your arms the same way that the pendulum of the clock operates. (A pendulum is a weight hung from a fixed point so that it can swing freely backward and forward. [See diagram on the right].)” (Vine et al., 2013). Participants in the explicit groups read the following:

> "1. Take your stance with your legs shoulder-width apart.
>
> 2. Set your position so that your head is directly above the ball looking down.
>
> 3. Keep your clubhead square to the ball.
>
> 4. Allow your arms and shoulders to remain loose.
>
> 5. In the putting action, your arms should swing freely backward and forward from your body, which should be still. Make sure that you accelerate through the ball.
>
> 6. After contact, follow through but keep your head still and facing down" (adapted from Vine et al., 2013). (For instruction sheets participants read, see https://osf.io/vpr92/, Questionnaires and Instructions.)

Participants in all groups had 2 min to read and study the analogy or explicit instructions (Daou, Hutchison et al., 2019). Next, participants completed the practice phase by performing six blocks of ten putts, taking a 1 min break between each block (participants sat in a chair during the breaks). When participants stopped practicing, they completed the Intrinsic Motivation Inventory

(Ryan, 1982). Intrinsic Motivation Inventory data were collected because it was possible that they would be used for exploratory analyses.

***Post-test***

Twenty-two to twenty-six hours after completing pretest and practice, participants returned to complete the experiment. Participants responded to the demographics questionnaire questions about drug/alcohol use, caffeine use, and previous night sleep. Participants in the teach groups were told, “The participant who you were going to teach did not show up today, so you will actually be tested on your putting instead.” Then, participants put on the physiological recording device. Next, they completed low-pressure and high-pressure tests in counterbalanced order. For the low-pressure test, the experimenter told participants, “In this set of ten putts, your goal is to make the ball stop as close to the center of the target as possible. Please, try to do the best you can.” For the high-pressure test, the experimenter told participants, “In the next set of ten putts, you will be recorded and critically analyzed by a golf expert who will give you a grade.” The experimenter affixed an iPad to the edge of a table, approximately 45° to the right and 225 cm in front of participants. The iPad’s screen faced participants so that they could see themselves being recorded. After the iPad was set-up, the experimenter told participants, “The combination of the golf expert grade and your performance during this set will allow you to compete against the rest of the participants for the 1$^{st}$ prize of $50, 2$^{nd}$ prize of $40, 3$^{rd}$ prize of $30, 4$^{th}$ prize of $20, and 5$^{th}$ prize of $10. In summary, you will be putting for money.” As the experimenter explained the rewards, he took an envelope from a cabinet, pulled money from it, and displayed the potential monetary rewards to participants, after which he placed the money on a 91-cm high countertop, approximately 30° to the left and 100 cm in front of participants. Our pressure manipulation involved two types of pressure revealed to elicit choking in previous studies: performance-contingent outcomes and monitoring by others (e.g., DeCaro et al., 2011).

After each post-test, participants completed the Revised Competitive State Anxiety Inventory-2 (Cox, Martens, & Russell, 2003) to determine manipulation efficacy. The Revised Competitive State Anxiety Inventory-2 is frequently used to assess anxiety in motor skill studies (Allsop & Gray, 2014; Elliot, Polman, & Taylor, 2014; Kinrade, Jackson, & Ashford, 2015; Kuan, Morris, Kueh, & Terry, 2018; Mullen, Jones, Oliver, & Hardy, 2016) and possesses good psychometric properties (Cox et al., 2003). The cognitive and somatic anxiety subscales were of interest since the pressure manipulation was intended to modulate anxiety. (Nonetheless, participants completed the self-confidence subscale as well) (Jackson, Ashford, & Norsworthy, 2006.) The cognitive and somatic anxiety subscale items ask participants to report how much they are currently feeling various indicators of anxiety. All responses were made by reporting a number between 0 and 100 on a scale with "not at all" corresponding to 0, followed by "somewhat", then "moderately so", and finally "very much so", which corresponds to 100.

After finishing post-test, participants completed a free recall test to measure declarative knowledge use. Specifically, participants were asked to report, in as much detail as possible, any rules, methods, or techniques they recall using to putt during the *high*-pressure post-tes. This type of free recall test is frequently used to assess declarative knowledge in motor skill studies (Daou, Buchanan et al., 2016; Daou, Lohse et al., 2016; Daou et al., 2018; Maxwell, Masters, & Eves, 2000; Maxwell, Masters, Kerr, & Weedon, 2001; Zhu, Poolton, Wilson, Maxwell, & Masters, 2011).

## Data Processing

### *Putting*

Putts were recorded with an iPad mounted to the ceiling above the target cross. We measured the ball's location relative to the target using a custom-developed program written in the National Instruments LabVIEW graphical programming language by Neumann and Thomas

(2008). Putting accuracy was indexed by recording radial error as recommended by Hancock, Butler, and Fischman (1995): $Radial\ Error = (x^2 + y^2)^{1/2}$, where x and y represent the magnitude of error along the respective axes (i.e., how far away from the target cross the ball stopped in the horizontal and vertical directions). Precision was indexed by recording bivariate variable error as recommended by Hancock et al.: $Bivariate\ Variable\ Error = \{(\frac{1}{k}) \sum_{i=1}^{k}[(x_i - x_c)^2 + (y_i - y_c)^2]\}^{1/2}$, where $k$ = trials in a block and $_c$ = centroid along the given axis (x or y) for that block. Radial error and bivariate variable error were calculated over pretest (10 putts) with the possibility that they could be used as covariates in exploratory analyses. Crucially, we did *not* plan to use either as an a priori covariate. Radial error and bivariate variable error were calculated for the first, third, and sixth blocks of the practice phase to get a glimpse into improvement during performance without overly slowing data processing. To assess motor learning and choking under pressure, radial error and bivariate variable error were calculated for the low- and high-pressure post-test. (For Neumann and Thomas's LabVIEW program to measure the distance of the ball to the target, and our Python code to import these distances, calculate radial error and bivariate variable error, and create a data frame with the error scores, see https://osf.io/vpr92/, Code to Measure Putt Location and Calculate Error Scores.)

***Self-Reported Anxiety***

Cronbach's α was calculated to determine the reliability of the Competitive State Anxiety Inventory-2 cognitive and somatic anxiety subscales for the low- and high-pressure post-test. If reliability was good (αs ≥ .700), then items would be averaged within the subscales. Next, a Pearson's correlation coefficient was calculated between the cognitive and somatic anxiety subscales for each post-test, and if $r \geq .500$, the subscales would be averaged together for each post-test. Otherwise, the subscales would not be combined for statistical analysis.

If the subscales did not exhibit good reliability, then physiological data would serve as the primary measure of anxiety. Specifically, Bioharness data would be extracted and analyzed using Omnisense software (Zephyr Technology, Annapolis, MD). Heart rate would be averaged from the time participants were read test instructions until they completed the test for the low- and high-pressure post-test. Heart rate variability (root mean square of successive differences and high frequency [0.150 – 0.400 Hz]) would also be assessed for these same periods.

### *Free Recall*

Two indices of declarative knowledge use were extracted from participants' responses on the free recall test. First, 'all concepts' referred to the number of statements about a concept (rule) (e.g., "I held my left hand over above my right"), ignoring statements irrelevant to technical performance (e.g., "I was told to putt ten times to the target"). Second, hypothesis testing referred to statements indicating that the participant had tested hypotheses related to their putting stroke (e.g., "I adjusted the swing path of the putter after each missed ball" or "I tried to keep my head still throughout my putting stroke"). That is, hypothesis testing statements were those that indicated the participant made a prediction about the relationship between their putting movement and putt outcome (Maxwell et al., 2001). We ignored retrospective statements (e.g., "I held my left hand above my right" or "My feet were shoulder-width apart") that may not have been used or thought about while putting, and we also ignored statements irrelevant to technical performance.

### *Movement Specific Reinvestment Scale*

Cronbach's α was calculated to determine the reliability of the Movement Specific Reinvestment Scale. Daou, Hutchison et al. (2019) found the Movement Specific Reinvestment

Scale had good reliability when all items were considered as one scale rather than dividing the Movement Specific Reinvestment Scale into its movement self-consciousness and conscious motor processing subscales. Further, we did not expect that either subscale should account for more residual variance in our data than the other subscale. Thus, we assessed the reliability across all items and summed them into a single scale if $\alpha \geq .700$. If the Movement Specific Reinvestment Scale had an $\alpha < .700$, then we would not consider conducting exploratory analyses with it.

**Statistical Analysis**

We conducted a 2 (Expectation) x 2 (Instructions) ANOVA with pretest radial error serving as the dependent variable. If the $\eta^2_p$s of the main effects or interaction $\geq .0099$ (Richardson, 2011), then pretest putting performance would be included as a covariate in all subsequent analyses involving putting performance.

Our primary analysis of interest was a 2 (Expectation) x 2 (Instructions) x 2 (Post-test) mixed-factor ANOVA with repeated-measures on the last factor, and radial error serving as the dependent variable. This follows because radial error (accuracy) was more sensitive to the Expectation x Post-test interaction observed by Daou, Hutchison et al. (2019) and reflects the objective of the putting task (accuracy with respect to target). Nonetheless, we conducted a secondary analysis using the same model with bivariate variable error serving as the dependent variable. We predicted an Expectation x Instructions x Post-test interaction, which would be followed up with separate 2 (Expectation) x 2 (Instructions) ANOVAs for the low- and high-pressure post-test. For the low-pressure post-test, we predicted a main effect of expectation, such that the teach groups would exhibit lower radial error (greater accuracy) than the test groups. For the high-pressure post-test, we predicted an Expectation x Instructions interaction. We would follow up this interaction with separate one-tailed *t*-tests (expectation) for the analogy and explicit groups. For the analogy groups, we predicted a significant effect of expectation, such that the teach/analogy group would exhibit lower radial error than the test/analogy group. For the explicit groups, we did not predict a significant effect of expectation. Movement Specific Reinvestment Scale score and/or pretest error scores would be considered for use as covariates/between-subjects factors in exploratory analyses of putting data.

To assess practice performance, we conducted a 2 (Expectation) x 2 (Instructions) x 3 (Block: 1/3/6) mixed-factor ANOVA with repeated-measures on the last factor separately for radial error and bivariate variable error. We predicted a main effect of block, such that participants would exhibit a linear decrease in radial error and bivariate variable error as a function of block (Daou, Hutchison et al., 2019).

To assess anxiety, we conducted a 2 (Expectation) x 2 (Instructions) x 2 (Post-test) mixed-factor ANOVA, with repeated-measures on the last factor and the total Revised Competitive State Anxiety Inventory-2 score serving as the dependent variable. If cognitive and somatic anxiety were not strongly correlated (see Self-Reported Anxiety section), then we would conduct a MANOVA instead, with the Revised Competitive State Anxiety Inventory-2 cognitive and somatic anxiety subscales serving as dependent variables. We predicted a main effect of post-test, with higher anxiety occurring on the high-pressure post-test. If heart rate and heart rate variability needed to be used to assess anxiety, they would be submitted to a 2 (Expectation) x 2 (Instructions) x 2 (Post-test) MANOVA.

To assess declarative knowledge use during high-pressure post-test, a 2 (Expectation) x 2 (Instructions) MANOVA was conducted with all concepts and hypothesis testing free recall scores serving as the dependent variables. We predicted an Expectation x Instructions interaction for all concepts (Daou, Hutchison et al., 2019). We would then conduct separate one-tailed *t*-tests (expectation) for the analogy and explicit groups. For the explicit groups, we predicted that teach participants would recall using more concepts than test participants. We do not predict an effect for the analogy groups.

The mixed-factor ANOVAs should be robust to violations of homogeneity of variance since we ensured equal *n*s. For the practice performance ANOVA, which has three levels of the repeated-measure, we would apply the Greenhouse-Geisser correction if sphericity was violated. Although ANOVAs should be robust to violations of normality, the one-tailed *t*-tests that follow them may not be (Field, Miles, & Field, 2012). Thus, we tested for violations of normality using the Shapiro-Wilk test and Q-Q plots, both of which we would consider when determining whether the data were non-normal. Since data from similar past research (Daou, Hutchison et al., 2019) suggests a positive skew is possible, a natural log transformation would be applied to data exhibiting a non-normal distribution. To test MANOVA assumptions, a Box test was used to assess homogeneity of covariance matrices, and a Shapiro test was used to determine multivariate normality. If these tests failed, then multiple ANOVAs would be conducted instead of MANOVAs.

We would conduct sensitivity analyses excluding participants who consumed alcohol/drugs within 24 hr of the first or second day of the experiment, caffeine within 3 hr of the first or second day of the experiment, or reported differences in sleep duration >2 hr between the night before the first and second day of the experiment. If the statistical significance of results of the primary analysis did not change when excluding these participants, then they would remain in the dataset. If the statistical significance of results of the primary analysis did change when excluding these participants, then they will be removed, and we would recruit additional participants to ensure $N = 148$ with equal *n*s per group.

The Auburn University Institutional Review Board approved the study (Protocol #20-519 EP 2011) on December 12, 2020. Data collection began on March 8, 2021 and finished on April 14, 2022.

## Results

### Descriptive Data

We collected data from 164 participants. However, 11 were excluded due to dropping out (*n* = 3), playing golf more than 30 times in their lifetime (*n* = 3), having missing post-test performance data (*n* = 1), showing up for post-test more than 26 hr after pretest and practice (*n* = 2), or not being able to complete data collection (*n* = 2). We then collected data from 3 extra participants to ensure equal sample sizes per group, resulting in 39 participants per group and a total sample size of 156 participants. Data for statistical analysis and trial-by-trial putting data[4] can be found at https://osf.io/vpr92/, Data. R Markdown files showing all statistical analyses and results can be found at https://osf.io/vpr92/, R Markdown files for Statistical Analyses and Results. After computing *z*-scores for low- and high-pressure post-test, four participants were identified as outliers and excluded. Thus, our results are based on analyses of 152 participants. Results of the primary statistical analyses with the four outliers included can be found at https://osf.io/vpr92/, Supplemental Results. Table 1 shows age, sex, and putting experience data by group.

[4] Please note the following about missing data. For pretest, four participants (IDs 45, 74, 114, and 142) are missing a datapoint due to a missing photograph of a putt. One participant (ID 82) shot all putts in practice block 1 out of the frame of the iPad and we did not have time to measure the putts by hand, so all datapoints for this participant for this block are missing. For practice block 1, five participants (IDs 10, 76, 98, 135, and 157) are missing a datapoint due to a missing photograph of a putt. For practice block 6, seven participants (IDs 75, 76, 98, 111, 116, 134, and 144) are missing a datapoint due to a missing photograph of a putt. For high-pressure post-test, one participant (ID 11) is missing a datapoint due to a missing photograph of a putt.

**Table 1**
*Age, Sex, and Putting Experience by Group*

| | Teach/Analogy ($n$ = 38; 21 females) | | Teach/Explicit ($n$ = 39; 24 females) | | Test/Analogy ($n$ = 38; 21 females) | | Test/Explicit ($n$ = 37; 24 females) | |
|---|---|---|---|---|---|---|---|---|
| | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Age (Years) | 20.82 | 1.78 | 20.69 | 1.84 | 21.47 | 2.69 | 20.59 | 2.13 |
| Lifetime Putting Experience[a] | 1.55 | 0.76 | 1.62 | 0.75 | 1.63 | 0.75 | 1.43 | 0.65 |
| Past-Year Putting Experience[a] | 0.74 | 0.45 | 0.82 | 0.51 | 0.74 | 0.60 | 0.62 | 0.49 |

*Note*: [a]0 = Never putted; 1 = Putted 1 – 10 times; 2 = Putted 11 – 20 times; 3 = Putted 21 – 30 times.

## Putting Accuracy

### *Post-Test*

Although the Shapiro-Wilk tests for low- and high-pressure post-test radial error data were significant (low-pressure post-test: $W(152) = .96$, $p < .001$; high-pressure post-test: $W(152) = .97$, $p = .006$), the Q-Q plots were fairly linear (Figure S1). To further investigate normality, we considered skewness and kurtosis values, which were modest (low-pressure post-test: skewness = 0.83, kurtosis = 0.88; high-pressure post-test: skewness = 0.63, kurtosis = 0.81). Thus, we concluded that data did not require a transformation to approximate normality. The 2 (Expectation) x 2 (Instructions) ANOVA with pretest radial error serving as the dependent variable revealed the main effect of expectation to be $\eta^2_p = .02$ ($p = .141$), the main effect of instructions to be $\eta^2_p < .01$ ($p = .327$), and the Expectation x Instructions interaction to be $\eta^2_p < .01$ ($p = .961$). Since the main effect of expectation was $\eta^2_p \geq .0099$, we included pretest radial error as a covariate in the analyses.[5]

[5] In hindsight, our pre-registered plan to include pretest radial error as a covariate if the size of the group main effects or interaction exceeded a particular value was not ideal because ANCOVA assumes the covariate is independent from the treatment effect (Field et al., 2012). However, we decided to follow our pre-registered plan since pretest radial error did not significantly differ as a function of treatment. Additionally, since our main statistical analyses were ANCOVAs, we tested the homogeneity of regression slopes assumption for each model to determine whether the covariate (pretest radial error) significantly interacted with any of the independent variables (expectation, instructions, and post-test) or their interaction terms. Specifically, for each ANCOVA with post-test radial error serving as the dependent variable, we added an interaction term for pretest radial error and each other independent variable, and, when applicable, an interaction term for pretest radial error and each interaction between the other independent variables, then reran the model. Importantly, results showed that pretest radial error did not significantly

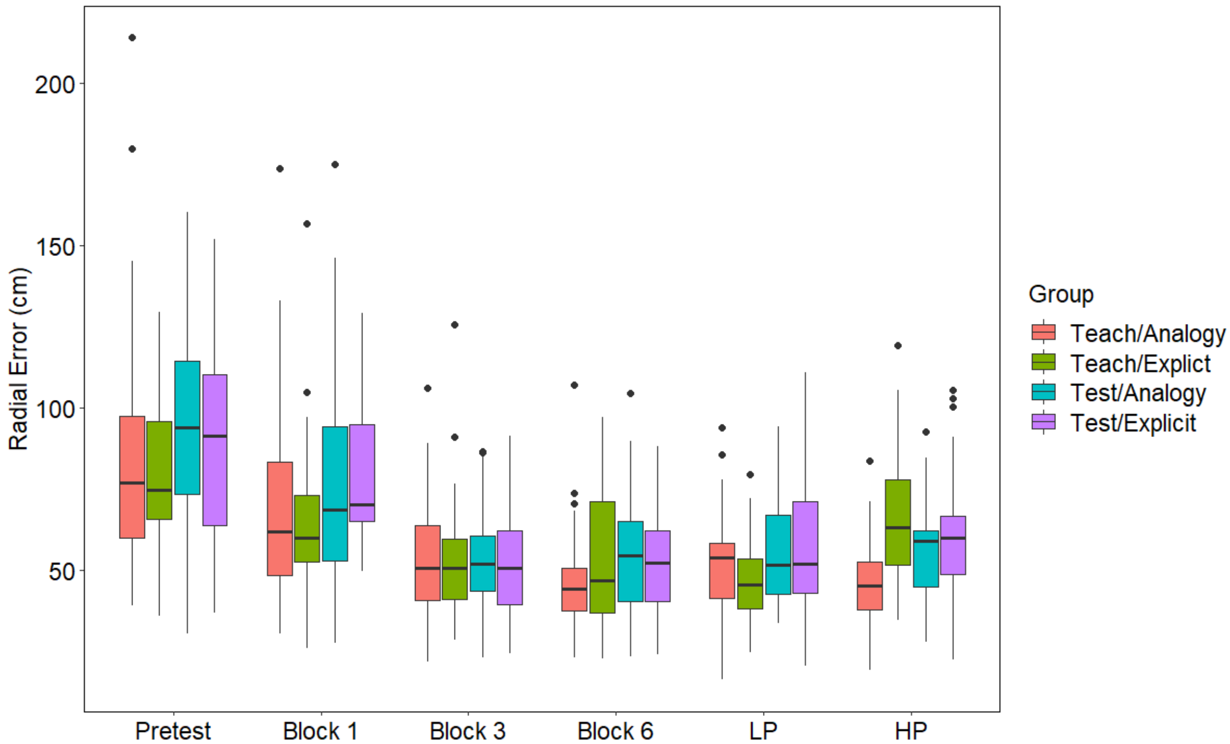
The primary analysis, a 2 (Expectation) x 2 (Instructions) x 2 (Post-test) mixed-factor ANCOVA with repeated-measures on the last factor, post-test radial error serving as the dependent variable, and pretest radial error serving as the covariate showed main effects of expectation ($F(1, 147) = 4.05$, $p = .046$, $\eta^2_p = .03$) and instructions ($F(1, 147) = 6.87$, $p = .01$, $\eta^2_p = .05$) as well as an Instructions x Post-test interaction ($F(1, 147) = 18.25$, $p < .001$, $\eta^2_p = .11$). As predicted, these effects were superseded by an Expectation x Instructions x Post-test interaction ($F(1, 147) = 8.62$, $p = .004$, $\eta^2_p = .06$). There were no other significant effects for this ANCOVA (post-test: $p = .493$, $\eta^2_p < .01$; Expectation x Post-test: $p = .170$, $\eta^2_p = .01$; Expectation x Instructions: $p = .377$, $\eta^2_p < .01$). The subsequent Expectation x Instructions ANCOVA for the low-pressure post-test revealed the predicted main effect of expectation ($F(1, 147) = 5.72$, $p = .018$, $\eta^2_p = .04$), such that participants who practiced with the expectation of teaching exhibited lower radial error than those who practiced with the expectation of being tested (Figure 1). There were no other significant effects for this ANCOVA (instructions: $p = .499$, $\eta^2_p < .01$; Expectation x Instructions: $p = .247$, $\eta^2_p < .01$). The Expectation x Instructions ANCOVA for the high-pressure post-test showed a main effect of instructions ($F(1, 147) = 23.19$, $p < .001$, $\eta^2_p = .14$), which was superseded by the predicted Expectation x Instructions interaction ($F(1, 147) = 6.71$, $p = .011$, $\eta^2_p = .04$). The main effect of expectation was not significant ($p = .493$, $\eta^2_p < .01$). To continue using pretest radial error as a covariate, we employed univariate ANCOVAs (expectation), rather than the planned one-sided *t*-tests (expectation), for the analogy and explicit instructions groups. As predicted, results revealed that participants who practiced with the expectation of teaching and analogy instructions exhibited lower radial error than those who

interact with any of the other independent variables or their interaction terms in any of the models ($p$s $\geq .066$, $\eta^2_p$s $< .03$). The analyses described above can be found in the R Markdown file (https://osf.io/vpr92/, R Markdown Files for Statistical Analyses and Results).

practiced with the expectation of being tested and analogy instructions ($F(1, 73) = 6.98$, $p = .01$, $\eta^2_p = .09$), whereas participants who practiced with the expectation of teaching and explicit instructions did not have significantly lower radial error than those who practiced with the expectation of being tested and explicit instructions ($p = .249$, $\eta^2_p = .02$) (Figure 1).

**Figure 1.**

*Putting Accuracy as a Function of Group and Phase*




*Note.* Radial error (y-axis) represents putting accuracy, with lower values indicating greater accuracy. Study phase (x-axis) shows pretest, practice blocks 1, 3, and 6, as well as low-pressure (LP) and high-pressure (HP) post-test. Each rectangular box represents the interquartile range, and the horizontal line in the middle of each box identifies the median. The lines extending from each rectangular box are 1.5 x the interquartile range. Each dot represents an individual data point.

### *Practice*

The 2 (Expectation) x 2 (Instructions) x 3 (Block) mixed-factor ANCOVA with repeated-measures on the last factor, practice radial variable error serving as the dependent variable, and pretest radial variable error serving as the covariate did not reveal significant effects (block: $p$ = .102, $\eta^2_p = .02$; expectation: $p$ = .195, $\eta^2_p = .01$; instructions: $p$ = .764, $\eta^2_p < .01$; Expectation x Block: $p$ = .191, $\eta^2_p = .01$; Instructions x Block: $p$ = .714, $\eta^2_p < .01$; Expectation x Instructions: $p$ = .666, $\eta^2_p < .01$; Expectation x Instructions x Block: $p$ = .107, $\eta^2_p = .02$).[6]

**Self-Reported Anxiety**

Cronbach's α for Competitive State Anxiety Inventory-2 cognitive and somatic anxiety subscales was .85 and .89 for low-pressure post-test, and .85 and .89 for the high-pressure post-test, respectively. Thus, we averaged items within the subscales. The correlation coefficient between the cognitive and somatic anxiety subscales for low-pressure post-test was $r$ = .63 and for high-pressure post-test it was $r$ = .64, so we averaged the subscales together for each post-test. As predicted, the 2 (Expectation) x 2 (Instructions) x 2 (Post-test) mixed-factor ANOVA with repeated measures on the last factor and averaged self-reported anxiety serving as the dependent variable showed a main effect of post-test ($F$(1, 148) = 136.03, $p < .001$, $\eta^2_p = .48$), such that participants reported more anxiety during the high-pressure post-test relative to the low-pressure post-test (Table 2). No other effects were significant (expectation: $p$ = .589, $\eta^2_p < .01$; instructions: $p$ = .667, $\eta^2_p < .01$; Expectation x Instructions: $p$ = .983, $\eta^2_p < .01$; Expectation x Post-test: $p$ = .790, $\eta^2_p < .01$; Instructions x Post-test: $p$ = .650, $\eta^2_p < .01$; Expectation x Instructions x Post-test: $p$ = .372, $\eta^2_p < .01$).

[6] We predicted a main effect of block on radial error during practice, but our ability to detect this effect was likely constrained by including a covariate of pretest radial error, which explained a substantial amount of variance in radial error during practice ($p < .001$, $\eta^2_p = .21$), reducing the amount of variance that could be explained by block.

**Table 2.**
***Self-Reported Anxiety by Post-Test and Group***

| | Teach/Analogy | | Teach/Explicit | | Test/Analogy | | Test/Explicit | |
|---|---|---|---|---|---|---|---|---|
| | *M* | 95% *CI* | *M* | 95% *CI* | *M* | 95% *CI* | *M* | 95% *CI* |
| Low-Pressure Anxiety | 20.07 | 15.16 – 24.99 | 19.31 | 14.51 – 24.11 | 19.64 | 14.26 – 25.02 | 17.47 | 13.28 – 21.65 |
| High-Pressure Anxiety | 30.52 | 23.85 – 37.19 | 29.01 | 24.67 – 33.34 | 28.12 | 21.89 – 34.35 | 28.23 | 22.19 – 34.27 |

**Free Recall**

Data from one participant was illegible, so we removed it from the dataset. The multivariate Shapiro-Wilk test revealed that the multivariate normality assumption was violated ($W(151) = .75$, $p < .001$), and the Box test showed that covariance matrices were not homogenous ($M(9) = 29.1$, $p < .001$). Therefore, we proceeded with two separate 2 (Expectation) x 2 (Instructions) ANOVAs with all concepts and hypothesis testing serving as dependent variables.

For all concepts, there was a main effect of instructions ($F(1, 147) = 49.89$, $p < .001$, $\eta^2_p = .25$), such that participants who practiced with explicit instructions reported using more concepts during the high-pressure post-test than those who practiced with analogy instructions (Table 3). Contrary to predictions, no other effects were significant (expectation: $p = .404$, $\eta^2_p < .01$; Expectation x Instructions: $p = .794$, $\eta^2_p < .01$). For hypothesis testing, contrary to hypotheses, no effects were significant (expectation: $p = .330$, $\eta^2_p < .01$; instructions: $p = .200$, $\eta^2_p = .01$; Expectation x Instructions: $p = .080$, $\eta^2_p = .02$).

**Table 3.**
*Declarative Knowledge Use in High-Pressure Post-Test by Group*

| | Teach/Analogy | | Teach/Explicit | | Test/Analogy | | Test/Explicit | |
|---|---|---|---|---|---|---|---|---|
| | *M* | 95%*CI* | *M* | 95%*CI* | *M* | 95%*CI* | *M* | 95%*CI* |
| Free Recall All Concepts | 2.05 | 1.58 – 2.53 | 3.85 | 3.24 – 4.45 | 1.76 | 1.32 – 2.21 | 3.69 | 3.10 – 4.29 |
| Free Recall Hypothesis Testing | 0.53 | 0.28 – 0.78 | 0.59 | 0.30 – 0.88 | 0.63 | 0.29 – 0.98 | 0.22 | 0.06 – 0.39 |

**Other Variables Collected**

Secondary analyses for post-test and practice putting with bivariate variable error serving as the dependent variable can be found at https://osf.io/vpr92/, Supplemental Results. Movement

Specific Reinvestment Scale data, Competitive State Anxiety Inventory-2 self-confidence subscale data for the low- and high-pressure post-test, and Intrinsic Motivation Inventory data for practice were not analyzed (other than computing Cronbach's α for the Movement Specific Reinvestment Scale, which was .78), but they can be found at https://osf.io/vpr92/, Data (Item-Level Questionnaire Data). Finally, we did not extract physiological data (heart rate and heart rate variability) because Competitive State Anxiety Inventory-2 cognitive and somatic subscales exhibited good reliability. Since we did not extract physiological data, we deviated from our plan to consider replacing participants who did not follow instructions to abstain from alcohol within 24 hr of the experiment, caffeine within 3 hr of the experiment, and/or did not get similar amounts of sleep the night before each day of the experiment. This follows because these instructions were intended only to facilitate physiological data collection. After consultation with an editor, we decided to conduct a sensitivity analysis of the primary statistical model excluding the participants who failed to follow instructions ($n = 19$), but we decided not to recruit additional participants, even if the statistical significance of the results of the primary analysis changed when excluding these participants. (Results of the primary analysis were similar when excluding these participants, and the predicted Expectation x Instructions x Post-test interaction remained significant [$p = .026$, $\eta^2_p = .04$]). For complete results of this sensitivity analysis, see https://osf.io/vpr92/, Supplemental Results.)

## Discussion

Having learners practice a motor skill with the expectation of teaching it has been revealed to enhance skill learning and promote large gains in declarative knowledge (Daou, Buchanan et al., 2016; Daou, Hutchison et al., 2019; Daou, Lohse et al., 2016; Daou et al., 2018; Daou, Rhoads et al., 2019). However, the learning benefit vanishes when the skill is performed under pressure, likely due to the use of declarative knowledge during skill execution, although this explanation lacked strong evidence (Daou, Hutchison et al., 2019). Therefore, the present study aimed to investigate whether this choking effect is, in fact, caused by an accrual of declarative knowledge during skill practice and whether choking is prevented if a technique to minimize the accrual of declarative knowledge during practice (analogy instruction) is applied. Our results showed that, on a low-pressure post-test, participants who practiced with the expectation of teaching exhibited greater putting accuracy than those who practiced with the expectation of being tested, but, on a high-pressure post-test, this advantage was only present for participants who also practiced with analogy instructions. Therefore, we provide evidence that the choking observed for participants who practiced with the expectation of teaching is indeed caused by declarative knowledge use and that analogy instruction is an effective learning strategy to maintain the benefit of expecting to teach when participants are performing under pressure.

Our results are consistent with the bulk of the evidence showing that expecting to teach a motor skill while practicing it enhances skill learning (Batista et al., 2022; Daou, Buchanan et al., 2016; Daou, Hutchison et al., 2019; Daou, Lohse et al., 2016; Daou et al., 2018; Daou, Rhoads et al., 2019; see McKay, Hussien, Yantha, Carter, & Ste-Marie (2021) for a high-powered study that did not observe this effect on learning). Specifically, we found a main effect of expectation in both the Expectation x Instructions x Post-test ANCOVA and the Expectation x Instructions ANCOVA for the low-pressure post-test, which is the pressure condition under which previous studies have observed an expecting to teach advantage (Batista et al., 2022; Daou, Buchanan et al., 2016; Daou, Hutchison et al., 2019; Daou, Lohse et al., 2016; Daou et al., 2018; Daou, Rhoads et al., 2019). Although expecting to teach has frequently been shown to benefit motor learning, the mechanisms behind this effect remain unclear, despite several candidates being examined, such as motivation and pressure (Daou, Buchanan et al., 2016; Daou, Lohse et al., 2016; Daou et al., 2018; Daou, Rhoads et al., 2019), length of motor preparation (Daou, Lohse et al., 2016; Daou, Rhoads et al., 2019), and cortical dynamics during motor preparation (Daou et al., 2018). Thus, more research is warranted to elucidate the underpinnings of the expecting to teach benefit to motor learning.

An important finding from Daou, Hutchison et al. (2019) was that the expecting to teach advantage vanished when participants were performing under pressure. Here, we replicated this result by finding an Expectation x Instructions x Post-test interaction due to a main effect of expectation during the low-pressure post-test, but not high-pressure post-test, where we found an Expectation x Instructions interaction. Daou, Hutchison et al. argued that participants who practiced with the expectation of teaching choked likely due to their accrual of declarative knowledge while practicing, although the authors failed to provide evidence to support this conclusion. Specifically, Daou, Hutchison et al. reported that declarative knowledge use during post-test did not mediate the relationship between expecting to teach and choking, possibly because declarative knowledge use was measured via free recall, which may be an imprecise measure of thoughts while putting. For example, Beilock, Wierenga, and Carr (2002) reported that trained novice golfers (like those in the present study) had only modest episodic memories of putts they executed. Here, we circumvented the problem of measuring declarative knowledge use during putting by successfully minimizing its availability in half the participants who practiced with the expectation of teaching by giving them analogy instructions (see main effect of instructions for free recall of all concepts). We reasoned that if they ceased to choke under pressure, then the choking effect could be attributed to declarative knowledge use. Results did indeed show that participants who practiced with the expectation of teaching and analogy instructions maintained their advantage under high-pressure, whereas those who practiced with explicit instructions did not. This was evidenced by an Expectation x Instructions interaction for the high-pressure post-test, caused by a main effect of expectation for participants who practiced with analogy instructions, but not for participants who practiced with explicit instructions. In sum, results strongly suggest that the reason learners who practice with the expectation of teaching choke under pressure is due to declarative knowledge use, and that limiting their accrual of declarative knowledge during practice with analogy instructions can maintain the expecting to teach benefit under pressure.

In the Expectation x Instructions x Post-test ANCOVA, we found a main effect of instructions that was superseded by an Instructions x Post-test interaction. This interaction was due to a main effect of instructions for the high-pressure post-test, such that the analogy groups presented greater accuracy, relative to the explicit groups, whereas no effect of instructions was observed for the low-pressure post-test. As analogy instructions promote implicit learning, these results are consistent with reinvestment theory (Masters & Maxwell, 2008) and a recent meta-analysis showing that implicit learning is beneficial to performance under pressure (Cabral, Wilson, & Miller, 2022). This feature of implicit learning can be attributed to the lack of declarative knowledge to reinvest in conscious control of movement (Masters & Maxwell, 2008). Although research seems to point to a benefit of implicit learning on performance under pressure, there is evidence that the benefit is overestimated due to reporting bias and underpowered studies (Cabral, Wilson et al., 2022). Besides these problems in implicit learning studies, there have been critiques about how explicit learning manipulations are implemented in the studies. For example, explicit learning groups are often given a greater number of instructions and ones that are of poor quality and have different meanings than those given to implicit learning groups (Bobrownicki, Collins, Sproule, & MacPherson, 2018). Therefore, pre-registered studies with greater statistical power and careful consideration of instructions given to explicit learning groups are recommended (Bobrownicki, Carson, MacPherson, & Collins, 2021; Cabral, Wilson, et al., 2022).

The present study's results have a few limitations, including their generalizability. Specifically, the sample consisted of healthy, young adults, like other studies examining the effect of expecting to teach on motor learning (Batista et al., 2022; Daou, Buchanan et al., 2016; Daou, Hutchison et al., 2019; Daou, Lohse et al., 2016; Daou et al., 2018; Daou, Rhoads et al., 2019; McKay et al., 2021), so it is unclear whether expecting to teach effects apply to other populations. Notably, analogy instructions have been shown to promote stable motor performance in other populations, such as older adults (Tse, Wong, & Masters, 2017). The present study's sample was also relatively novice at the motor skill, like most motor learning studies (Williams, Fawver, & Hodges, 2017), and thus whether the effects of expecting to teach and analogy instruction extend to experts is unclear. Finally, the generalizability of the expecting to teach and analogy instructions effects to motor skills other than closed ones is uncertain because most studies examining these effects have done so with skills such as golf putting, dart throwing, and basketball shooting (e.g., Batista et al., 2022; Daou, Hutchison et al., 2019; Lam et al., 2009a). Another limitation of the present study's results is the ecological validity of the analogy instructions benefit, since the control (explicit instructions) group practiced with instructions that may be unrepresentative of those given in the real world (Bobrownicki, Carson, & Collins, 2022). Finally, participants who practiced with the expectation of teaching and analogy instructions did not choke under pressure, strongly suggesting that the choking effect for those who practiced with the expectation of teaching and explicit instructions was caused by declarative knowledge use; yet, participants who practiced with the expectation of teaching did not recall using more declarative knowledge during the high-pressure post-test. However, it is important to reiterate that free recall is likely an imprecise measure of declarative knowledge use, such that it is only sensitive enough to detect very large differences in declarative knowledge use (e.g., differences associated with analogy vs. explicit instructions).

The present study is methodologically strong, and its results can inform future research and practical applications. The methodological strength is rooted in the study being a registered report, which allowed the study design to undergo a peer review that led to high statistical power and well-planned experimental manipulations, analyses, and outcome measures. These attributes along with the open science practices (e.g., access to experimental materials, statistical processing code, and data) required for the registered report increase the positive predictive value of the study results (Ioannidis, 2005), add precision to the effect size estimates, and facilitate future research, including replications and meta-analyses. Other avenues for future research include addressing the limitations to the generalizability of the expecting to teach and analogy instructions benefits by conducting studies with different populations (including experts) and tasks. Further, researchers are encouraged to investigate other methods aiming to maintain the benefits of expecting to teach when performing under pressure. These methods could involve other implicit learning strategies (e.g., errorless learning; Zhu et al., 2011), but also techniques commonly associated with theories of choking under pressure other than reinvestment theory. For example, researchers could examine whether having participants practice a motor skill with the expectation of teaching and psychological pressure (anxiety training) prevents choking due to distraction (distraction theories; DeCaro et al., 2011) in subsequent high-pressure post-test (Low et al., 2022; Oudejans & Pijpers, 2010). Researchers could also investigate whether having participants practice with the expectation of teaching and providing an intervention at the time of a high-pressure post-test prevents choking under pressure. Such interventions could include pre-performance routines, mental imagery, self-talk, and meditation (Chen et al., 2017; Fekih et al., 2021; Mesagno, Geukes, & Larkin, 2015; Walter, Nikoleizig, & Alfermann, 2019). In practice, the present results indicate that learners should train under the expectation of teaching and analogy instructions to enhance motor learning and preserve skill performance under pressure. Instructors may also consider having learners practice with the expectation of teaching and, pending future study results, implementing strategies other than analogy instructions to prevent choking under pressure.

## Conclusion

In conclusion, the present study indicates that learners who practice with the expectation of teaching exhibit superior learning, and when they choke under pressure it is likely due to their accrual of declarative knowledge during practice, since the choking effect is prevented by having them practice with analogy instructions. Accordingly, instructors are recommended to have learners practice with the expectation of teaching and techniques that minimize the accrual of declarative knowledge, such as analogy instructions.

# Chapter 3: The effects of anxiety training on motor performance under pressure: Investigating the psychological mechanisms

## Introduction

The ability to perform a motor skill under psychological pressure is critical and a common feature of most sports and activities involving motor performance. However, it is common to experience an acute and substantial decrease in skill execution due to increased anxiety and pressure, a phenomenon known as choking under pressure (Mesagno & Hill, 2013). The decisive free throw in a basketball game, a penalty during a soccer match, and the match point in a tennis game are examples of pressure situations where people could choke. Some possible explanations for the choking phenomenon are based on distraction and self-focus theories, such as the attentional control theory (ACT) (Eysenck et al., 2007) and the reinvestment theory (Masters & Maxwell, 2008), respectively. In addition to these traditional accounts of choking, the biopsychosocial model (BPSM) of challenge and threat has been suggested as a possible explanation for reduced performance under pressure (Blascovich, 2008; Seery, 2013).

### *Theories of Choking*

The ACT proposes that anxiety impacts performance due to disruption of top-down (goal-directed) attentional control, while bottom-up (stimulus-driven) processing is increased (Eysenck et al., 2007). In other words, individuals performing under pressure are more likely to direct their attentional resources toward worries and detection of possible threat stimuli, while less attention is directed toward task-relevant information (Eysenck et al., 2007). Additionally, individuals performing a task under enhanced anxiety levels might try to compensate for the harmful effects of anxiety by increasing mental effort invested in the task in an attempt to maintain performance (i.e., decreased processing efficiency) (Eysenck et al., 2007). Supporting this prediction, Wilson and Smith (2007) showed that golf players avoided choking under pressure at the expense of increased mental effort and decreased target final fixation, meaning that participants spent less time looking at the target before putting (considered a less efficient visual search strategy), a combination that reflects reduced processing efficiency.

The self-focus theories of choking propose that anxiety leads to disruption of automatic components of the movement by making individuals consciously monitor and/or control the step-by-step execution of the skill. A well-known self-focus theory is the reinvestment theory (Masters & Maxwell, 2008), which suggests that anxiety leads to the disruption of the skill's performance due to the individual reinvesting attentional resources for skill execution by using declarative knowledge acquired earlier during practice. This focus of attention on the movement reduces its automaticity and hinders its performance (Wulf & Su, 2007), and is likely due to inefficient muscle activation as well as invariable and uncorrelated effector movement (Lohse et al., 2014).

Alternatively, the BPSM of challenge and threat contends that individuals evaluate task demands and determine if they have the necessary resources to cope with those demands (Blascovich, 2008; Seery, 2013). Individuals who perceive they have available resources to deal with imposed task demands perceive the task as a challenge. Conversely, individuals who perceive they do not have the available resources to cope with the demands perceive the task as a threat (Seery, 2013). Feelings of challenge and threat lead to distinct physiological outcomes, mostly evidenced by changes in the cardiovascular system. For instance, a challenge state would elicit increasing heart rate (HR) and ventricular contractility, leading to increased blood supply to the brain and muscular vessels (Fawler et al., 2020). A threat state would elicit a similar pattern as described above, but would also elicit the release of cortisol, which would lead the vessels to contract. This causes less efficient energy mobilization, which disfavors physical activity (Blascovich, 2008; Seery, 2013).

***Anxiety Training***

Anxiety training (AT) might be a learning strategy to lead individuals to focus on task-relevant stimuli, prevent attempts to consciously monitor and control movement, and perceive a task as challenging rather than threatening. AT requires individuals to practice a skill under enhanced levels of pressure. This learning strategy is believed to reduce the harmful effects of anxiety when participants are performing the skill under pressure, thus avoiding choking (Beseler et al., 2016; Oudejans & Pijpers, 2009, 2010). AT may train individuals to increase their mental effort to maintain attention on the task goal rather than worries about the task outcome, sensations related to anxiety (Oudejans & Pijpers, 2010), or a need to monitor and control movement, although the latter has yet to be examined. Further, AT may cause individuals to believe they have resources to cope with high-pressure situations, consequently increasing their self-confidence (Vine et al., 2016), which is linked to attending to task goals (rather than attending to movement; Wulf & Lewthwaite, 2016), and perceiving the task as challenging rather than threatening. Thus, AT may elicit positive performance and adaptive physiological outcomes (e.g., higher cardiac output, blood flow and vessel dilation) (Seery, 2013; Vine et al., 2016). A recent meta-analysis provided evidence of AT benefits to performance under pressure ($g$ = 0.67, 95% *CI* [0.43, 1.12]), regardless of the task or participants' level of experience (Low et al., 2021). For example, Oudejans and Pijpers (2010), which was included in the meta-analysis, randomly assigned novice dart players to either AT or control group. While the AT group practiced the skill under mild levels of anxiety, the control group practiced the skill without any pressure manipulation. During the post-tests, the AT group maintained performance from low-to mild- to high-pressure conditions, whereas the control group choked from mild- to high-pressure. Interestingly, both groups showed systematic increases in mental effort, as measured by a self-reported continuous scale ranging from 0 ("no effort at all") to 10 ("most effort ever"), across post-tests, thus supporting the ACT prediction that on-task effort is increased to counteract the negative effects of anxiety. The authors argued that the mental effort shown by the AT group was adaptative, whereas the increase seen in the control group was not. This discrepancy was likely due to the AT group's development of more effective coping strategies to deal with anxiety and pressure situations during practice (Oudejans & Pijpers, 2010). In another study (Oudejans et al., 2011), police officers who practiced a shooting exercise under a high-anxiety condition were able to maintain performance during high-pressure post-test, while the control group exhibited a performance decrease from the low- to the high-pressure post-test condition. Again, both groups showed increased mental effort from low- to high-pressure post-test. These studies corroborate the ACT prediction that enhanced anxiety levels are accompanied by decreases in processing efficiency in order to maintain performance (Eysenck et al., 2007). However, the literature lacks studies that account for other possible explanations for the benefits of AT.

***The Current Experiment***

In the present study, we will attempt to conceptually replicate previous findings showing AT benefits performance under pressure. More specifically, we are grounding our design on Oudejans & Pijpers (2010), wherein the authors tested whether training under mild levels of anxiety would prevent choking under higher levels of anxiety. This is an important feature of Oudejans and Pijpers' study because it reflects the reality that competition elicits levels of anxiety that are difficult, if not impossible, to reproduce during training. Thus, we will add to the AT literature by attempting to conceptually replicate this study with a much larger sample size, which should lead to a highly informative result irrespective of the outcome (Simonsohn, 2015). Additionally, we aim to expand on the AT literature by measuring variables related to other theories of choking besides ACT (Alder et al., 2016; Oudejans et al., 2011; Oudejans & Pijpers, 2009, 2010). In addition to mental effort (accounting for ACT), we will measure movement reinvestment and challenge-threat appraisal to account for self-focus and BSPM of challenge and threat, respectively, which will give us a broader perspective of mechanisms behind the expected benefits of AT on performance under pressure.

Therefore, the present study examined whether AT prevents motor performance decrements from a low- to mild- to a high-pressure post-test. We expected the AT group to maintain performance across all post-tests as previously shown (Oudejans & Pijpers, 2010), and the control group to show reduced performance from low- to mild-, and from mild- to high-pressure post-tests. (Although Oudejans & Pijpers (2010)'s control group did not show reduced performance from low- to mild-pressure post-tests, the manipulation used for our mild-pressure post-test has been shown to elicit choking (Daou et al., 2019).) As shown by previous studies (Oudejans et al., 2011; Oudejans & Pijpers, 2009, 2010), we expect anxiety levels and HR to systematically increase for both groups across post-tests. Consequently, we expect mental effort to show a similar pattern for both groups to counteract the effects of anxiety on performance. Additionally, we speculate that participants in the control group will reinvest in their movement and perceive increased threat from low to mild to high pressure post-tests, whereas the anxiety training group will not.

## Methods

The study preregistration, data, questionnaires, statistical assumptions, and supplementary analyses are publicly available at https://osf.io/xpv8e/.

### Participants

Men and women aged 18 to 30 years participated in the study and were offered 6 course credits for participation (Cabral, Daou et al., 2022; Daou et al., 2019). Participants were recruited if they had putted, anything from playing mini golf to playing 18 holes on a standard golf course, between 1 and 30 times in their lifetime and not more than 20 times in the past year (Cabral, Daou et al., 2022; Daou et al., 2019). Additionally, participants must not have been categorized as an elite athlete of any sport, defined by individuals who play(ed) in Division I of the National Collegiate Athletic Association or at a higher level (e.g., professional), because research has shown that athletes generally have reduced stress levels before a stressful event and increased mental toughness, compared to non-athletes (Jalili, 2011; Verner et al., 2010). (Although Low et al. (2021)'s meta-analysis suggests AT is effective irrespective of participant skill level, limiting the heterogeneity of the present study's sample may increase statistical power by reducing variability in outcome measures.) Participants were free from physical illness, injury, or disability that could make putting difficult (for more detail on the initial questionnaires and demographics questions, see supplementary material). Participants were asked to avoid alcohol/drug consumption within 24 hr of all days of the study, as well as caffeine consumption within 3 hrs of each testing period. Lastly, participants were instructed to try to get a similar amount of sleep each night prior to a day of the study.

**Sample Size Calculation**

We first used the General Linear Mixed Model Power and Sample Size (GLIMMPSE) 3.0 tool to estimate the required sample size for our study (Kreidler et al., 2013). To confirm this result, we then used the ANOVA_power R Shiny app (Lakens & Caldwell, 2021) to perform 2000 Monte Carlo simulations to estimate the study's power using the sample size returned by GLIMMPSE, along with the same parameters we inputted in GLIMMPSE. For the GLIMMPSE parameters, we selected a power of 80%, an alpha of .05 and the Geisser-Greenhouse correction, which we planned to use if we violated sphericity. We set radial error as our outcome measure, specified post-test as a categorical repeated measure with three levels (low pressure, mild pressure, and high pressure) with no clustering, entered a nominal fixed predictor of group with two levels (AT and control), hypothesized a Group x Post-Test interaction with a null hypothesis that group differences would be zero at each post-test, left the contrast comparison constant set at zero, and set equal group sizes. Next, we set the radial error means and standard deviations for each group for each post-test. We chose our values based on the standardized effect size from Low et al. (2021)'s meta-analysis (which did not find evidence that the estimate was biased [i.e., no funnel plot asymmetry]), so we inputted standardized values, which we did not scale. Specifically, we used a standard deviation of one for each post-test. For the means, we expected the AT group to maintain similar performance across all post-tests, so we inputted one for each post-test. For the control group, we expected a medium performance decrement from the low- to mild-pressure post-test and a large performance decrement from the low- to high-pressure post-test, so we chose means of 1, 1.5 and 1.8 for the low-, mild-, and high-pressure post-tests, respectively. We based our expectations for performance decrement on Low et al.'s finding that AT resulted in a medium ($g$ = 0.67) benefit under pressure conditions like our mild-pressure post-test. For our post-test correlation matrix, we based our input on two studies (Daou et al., 2019; Thompson, 2016) conducted in our lab ($N$s = 82 and 83, respectively) wherein participants performed motor skills in low- and high-pressure conditions, with the high-pressure condition being like our mild-pressure condition and eliciting significantly higher anxiety and performance decrement than the low-pressure condition. Daou et al. (2019) reported a correlation coefficient of .36 and Thompson reported a coefficient of .38, so we assumed a correlation coefficient of .35 for adjacent post-tests (low-pressure to mild-pressure, and mild-pressure to high-pressure) and a slightly lower coefficient (.30) for low- to high-pressure. GLIMMPSE calculated a required sample size of 84 (42/group). ANOVA_power simulations indicated 96% power for the Group x Post-Test interaction with $N$ = 84, so we were confident that GLIMMPSE's estimated sample size would be sufficient. Therefore, we aimed to collect data from 90 participants to account for data loss and/or participant dropout and ensure complete data from 84 participants. However, I realized that I would not be able to collect the full sample size by my graduation deadline, so I re-calculated the sample size necessary to achieve 67% power, while holding all other parameters in the calculation constant. The required sample was 64. We then collected 70 participants to account for dropouts and any data collection problems.

## Task

Participants used a standard golf putter (88.90 cm) to putt a standard golf ball from a starting location indicated by a 5 cm line painted in white washable paint on a foam mat designed to simulate a medium-speed (10 – 11 stimpmeter), natural putting green (BirdieBall Inc., Evergreen, CO, USA) to a target cross (+) comprised of two 10.80 cm lines painted in white washable paint and located 300.00 cm away from the starting position. Participants' objective was to make the ball stop as close to the center of the target as possible.

## Procedure

### *Day 1 – Pre-Test and Practice Session 1*

Participants completed the experiment individually. After signing the informed consent form, participants completed a demographic questionnaire asking their sex, age, putting experience, any illness, injury, or disability that could make putting difficult, if they consumed alcohol or any drugs in the last 24 hrs, if they consumed caffeine within the last 3 hrs, and how many hours of sleep they had the previous night. After confirmation that participants met the inclusion criteria, they put a heart rate (HR) monitor (Polar ProTrainer, model H10), wetted to ensure a better connection, around their chest. Participants wore the HR monitor throughout the experiment to collect data for potential exploratory analyses. Next, participants started the pre-test, which consisted of 1 block of 10 putts, and completed the Revised Competitive State

Anxiety Inventory-2 (CSAI-2R), as well as mental effort, reinvestment, and challenge-threat questionnaires (i.e., baseline levels).

After pre-test, we used the website random.org to quasi-randomly assigned (based on sex) participants to either the AT or control group. This was because sex can moderate performance under pressure, where the relationship between anxiety and reduced performance is more pronounced in females, relative to males (Englert & Seiler, 2020). Participants from both groups were asked to follow a list of explicit instructions related to golf putting execution, so they could accrue declarative knowledge, which has been suggested to increase the risk of choking under pressure (Masters & Maxwell, 2008). With this measure, we expected to increase the likelihood of generating a choking effect that could be moderated by AT. Further, explicit instructions should reduce variability in participants' putting motion (i.e., deter irregular putting techniques), some of which may reduce the propensity for choking under pressure (Iso-Ahola et al., 2016). The instructions were adapted from previous golf putting studies (Daou et al., 2019; Vine et al., 2013; Vine & Wilson, 2010). Next, participants performed 2 blocks of 75 trials with a 2-min break between blocks. Then, participants received their group-specific instructions. Participants in the AT group were told "During the next two days of practice, you will perform 2 blocks of 75 putts per practice day. Additionally, you will be recorded and critically analyzed by a golf expert who will give you a grade." Next, the experimenter affixed an iPad to the edge of a table in front of and to the left of right-handed participants or to the right of left-handed participants. The iPad's screen was faced toward participants so that they could see themselves. After the iPad was set-up, the experimenter told participants, "The combination of the golf expert grade and your performance during the next two days of practice will allow you to compete against the rest of the participants for the 1st prize of $50, 2nd prize of $40, 3rd prize of $30, 4th prize of $20, and 5th prize of $10. In summary, you will be putting for money. Make sure you do your best." As the experimenter explained the award, he took an envelope from a cabinet, pulled money from it, and showed the monetary awards to participants, after which he placed the money on a countertop, visible to participants. Conversely, the control group was told "During the next two days of practice, you will perform 2 blocks of 75 putts per practice day. Make sure you do your best in each day." The pressure manipulation included two categories of pressure that yielded choking in previous studies: performance-contingent outcomes and monitoring by others. Additionally, we have successfully implemented these manipulations in our lab in the past (Cabral, Daou et al., 2022; Daou et al., 2019). Next, participants performed 2 blocks of 75 trials with a 2-min break between them. After the first 10 putts of practice, participants from both groups completed the CSAI-2R, as well as the reinvestment, challenge-threat, and mental effort questionnaires.

***Days 2 –Practice Session 2***

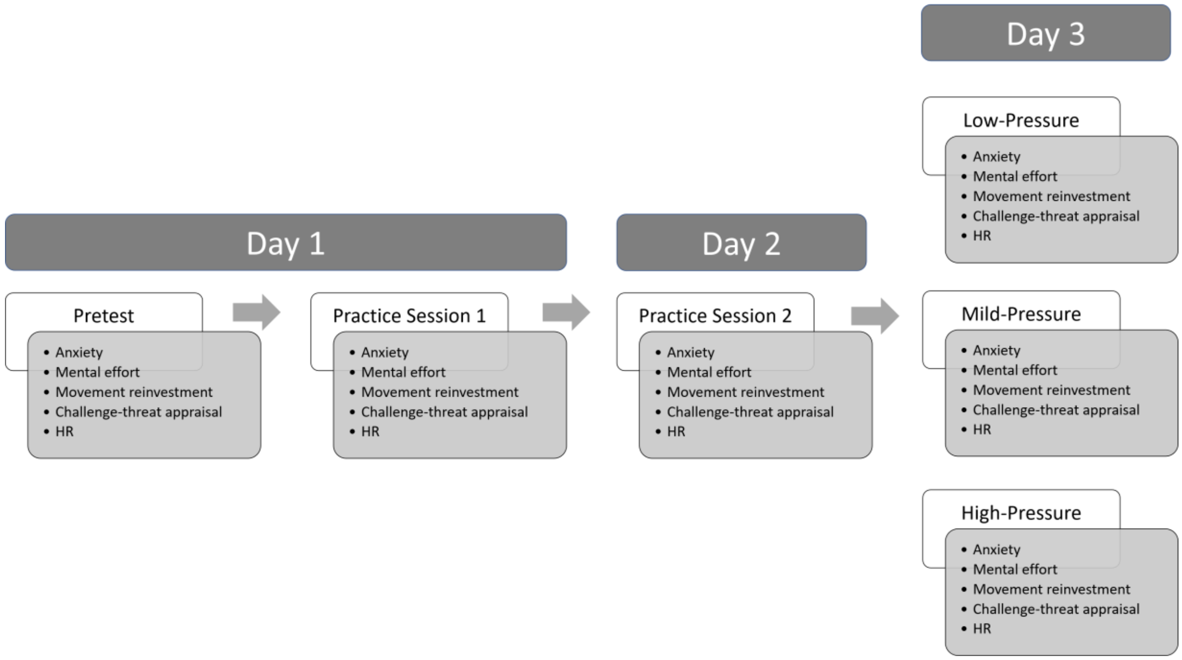
Day 2 was the second practice session and occurred approximately 2 days after Day 1. Before the practice session, participants provided verbal consent to continue the study and were asked if they consumed alcohol or any drugs in the last 24 hrs, if they consumed caffeine within the last 3 hrs, and how many hours of sleep they had the previous night. Next, participants put on a HR monitor. Then, participants performed the practice session like they did on Day 1, with minor changes in what they are told (e.g., "During this *practice session*…"). Again, participants completed the self-reported questionnaires after the first 10 putts of practice. At the end of Day 2, participants had practiced 300 putts, as in similar studies with golf putting and AT (Beilock & Carr, 2001; Lawrence et al., 2014).

***Day 3 - Post-Tests***

Approximately one week after the first day of practice, participants completed three post-tests. Before doing so, participants provided verbal consent to continue the study and were asked if they consumed alcohol or any drugs in the last 24 hrs, if they consumed caffeine within the last

3 hrs, and how many hours of sleep they had the previous night. Next, participants put on a HR monitor. Both groups performed low-, mild- and high-pressure post-tests in a counterbalanced order. In all post-tests, participants were instructed to perform 1 block of 10 putts. For the low-pressure test, participants were instructed to perform the putts as best as they could, and were told that this test was just to complete the dataset (Oudejans & Pijpers, 2010). For the mild-pressure condition, all participants putted while going through the same pressure manipulation as the AT group during the practice phase. For the high-pressure condition, one extra pressure manipulation was added to the manipulations applied during the mild-pressure condition. Specifically, participants were told that they had been randomly partnered with another participant and if both showed an individual improvement of 50% relative to the last 10 practice trials on Day 2, they would each earn $100 in addition to any other monetary award they earned. The money was displayed to participants and placed on the counter as well. However, if one of them failed, then neither would get the $100 bonus. The experimenter told participants that their partner had already succeeded by improving their performance 50% from the last 10 practice trials on Day 2, so the $100 bonuses entirely depended on the participant's performance (Beilock et al., 2004; Beilock & Carr, 2005; Beilock & Decaro, 2007). Additionally, they were told that their partner would be watching them live while they performed. Accordingly, we had a confederate ostensibly watching their performance live through a laptop. Finally, after each condition, participants completed the CSAI-2R, the movement-specific reinvestment, challenge-threat, and mental effort questionnaires.

**Figure 1.**
*Timeline of the Study for Both Groups*




*Note:* Both groups performed pre-test, two practice sessions and three post-tests in a counterbalanced order. After each condition, participants completed the CSAI-2R (Revised Competitive State Anxiety Inventory-2), the movement-specific reinvestment, challenge-threat, and mental effort questionnaires, and had their HR (heart rate) measured from the first to the last putt of each condition. Putting performance was measured during pre-test, putts 1 - 10, 65 - 75, and 140 - 150 of each practice session, and during the post-tests.

## Measures

### *Self-Reported Anxiety*

The CSAI-2R (Cox et al., 2003) has been used to evaluate anxiety in many motor skill studies (e.g., Allsop & Gray, 2014; Elliott et al., 2014; Kuan et al., 2018) and has good psychometric properties (Cox et al., 2003). The cognitive anxiety, somatic anxiety, and self-confidence subscales were relevant since the pressure manipulation was intended to influence anxiety and decrease self-confidence (Jackson et al., 2006). The cognitive and somatic anxiety subscale items required participants to state how much they perceived several indicators of anxiety during that moment and were used as the primary indicator of anxiety levels. The self-confidence subscale required them to indicate how confident they were when performing the task, and we used this subscale to conduct exploratory analyses. Participants were asked to report a number: 1 ("not at all"), 2 ("somewhat"), 3 ("moderately so"), or 4 ("very much so") in response to items asking how anxious or confident they felt.

### *Challenge and Threat Appraisal*

We measured challenge and threat appraisals using an adapted version of the cognitive appraisal ratio (Tomaka et al., 1997). Specifically, the first item assessed perceived demands: "How demanding was the golf task for you?", and the second item assessed perceived resources: "How able were you to cope with the demands of the golf task?" The two items were scored on a 6-point scale going from 1 ("not at all") to 6 ("extremely"). The two appraisal variables were combined to yield a single index of cognitive appraisal (Feinberg & Aiello, 2010; Laborde et al., 2015; Moore et al., 2012). This index was calculated as a ratio of the first item to the second item and reflected whether participants perceived they had available resources to match task demands (Blascovich & Tomaka, 1996). Ratios equal to or greater than 1 indicate perceived threat and values lower than 1 indicate a challenge state. The challenge-threat ratio has been previously used to investigate performance in golf putting tasks (Moore et al., 2012).

### *Movement Reinvestment*

To assess movement reinvestment, participants completed the Conscious Motor Processing Subscale of the Movement Specific Reinvestment Scale (Orrell et al., 2009) adapted for golf putting (Cooke et al., 2011; Vine et al., 2013). Specifically, participants were asked to indicate what they thought about while they were putting during the previous block of putts. For example, ''*I thought about my stroke*'' and ''*I thought about bad putts*.'' Each item was scored on a 5-point Likert scale anchored by 1 ("never") and 5 ("always").

### *Mental Effort*

To measure mental effort, we used the Rating Scale of Mental Effort (Zijlstra, 1993). This scale rates the mental effort required to perform a task, and the score ranges from 0 (“absolutely no effort”) to 150 (beyond “extreme effort”). The higher the score, the greater the mental effort. Participants were required to indicate a point on the scale that represented how much effort they had invested in the task. This scale has been previously used to quantify mental effort in motor performance under pressure studies (Alder et al., 2016; Bobrownicki et al., 2015; Wilson et al., 2007).

### *Heart Rate*

The Polar H10 heart rate monitor has been previously validated for laboratory studies and has shown good reliability (Georgiou et al., 2018; Gilgen-Ammann et al., 2019). The device provides raw RR interval time series with a sampling rate of 1 kHz. The output file allowed us to quantify both HR and HR variability (HRV) for the periods of interest, which were from the first to the last putt for pre-test, first and second days of practice, and post-tests.

## Data Processing

### *Putting Performance*

Putts were recorded with an iPad that was placed on the ceiling above the target cross. We measured the ball’s location relative to the target using a custom-developed program written in the National Instruments LabVIEW graphical programming language (Neumann & Thomas, 2008). Putting accuracy was indexed by recording radial error (Hancock et al., 1995): $Radial\ Error = (x^2 + y^2)^{1/2}$, where $x$ and $y$ represent the magnitude of error along the respective axes. Radial error was averaged over pre-test, over trials 1 – 10 (“early”), over trials 65 – 75 (“middle”), and over trials 140 – 150 (“late”) of each practice day (to provide a glimpse into training performance during early, middle, and late practice trials on each day), and radial error was averaged over the low-pressure post-test, over the mild-pressure post-test, and over the high-pressure post-test.

### *Self-Reported Anxiety and Self-Confidence*

Items in cognitive, somatic, and self-confidence subscales were averaged within each subscale separately for each time point, and the Cronbach's α was calculated for each subscale for each time point. A Cronbach's α of ≥ .7 was considered good reliability for these subscales and all other scales. Since cognitive and somatic anxiety subscale scores were strongly correlated (*r*s > .84), they were averaged to make one anxiety score for each time point.

### *Movement Specific Reinvestment Scale*

Items in the Movement Specific Reinvestment Scale were averaged together for each time point and Cronbach's α was calculated for each point time.

### *Heart Rate*

A text file was extracted from the HR monitor device and imported to the Kubios software (Tarvainen et al., 2014). Artifact correction due to ectopic beats (i.e., irregular heart rhythm due to a premature heartbeat), missed beat detections, and technical artifacts (i.e., any possible error during data collection) was implemented by using Kubios filtering algorithms. We then calculated HR for pre-test, early, middle, and late phases from each practice day, and post-tests. The parameters used to calculate HRV can be found in the supplementary material at https://osf.io/xpv8e/.

## Statistical Analysis

### *Pressure Manipulation Check*

To verify if the post-test pressure manipulation was successful, we conducted separate 2 (Group: AT vs. control) x 3 (Post-Test: low-, mild-, high-pressure) ANOVAs with repeated measures on the last factor, and the average of the cognitive and somatic subscales of the CSAI-

2R and HR serving as dependent variables. For both anxiety and HR, we predicted a main effect of post-test, which we planned to follow up with one-tailed paired *t*-tests. We expected participants to show systematic increases in both anxiety and HR from low-, to mild-, to high-pressure post-tests. If at least one measure (anxiety or HR) from the mild-pressure post-test was greater than from the low-pressure post-test, and if at least one measure from the high-pressure post-test was greater than from the mild-pressure post-test, the post-test manipulation would be considered successful. For the self-confidence subscale, we did not have specific predictions and planned to follow up the 2 (Group) x 3 (Post-Test) ANOVA dependent upon its result.

To verify if the pressure manipulation was successful during practice for the AT group, we planned to conduct separate one-way (Group) ANCOVAs with the scores from the average of the cognitive and somatic anxiety subscales of the CSAI-2R and HR serving as dependent variables, and pre-test measures of these variables serving as covariates in their respective ANCOVAs. Notably, we only compared pre-test and the early portion (putts 1 – 10) of the first practice day because participants in the AT group might have reduced anxiety levels and/or HR during the second practice session as they acclimated to pressure. If at least one measure (anxiety or HR) was greater for the AT group than the control group, the group manipulation would be considered successful.[7]

***Putting Accuracy During Post-Tests***

[7] For the ANCOVA assessing whether the AT group reported higher anxiety early in practice on Day 1, the assumption that the covariate was independent from the treatment effect was not met, as revealed by a Wilcoxon signed-rank test with group serving as the independent variable and pre-test anxiety serving as the dependent variable ($p < .001$) (Field et al., 2012). Thus, we performed this manipulation check by conducting a one-sided Wilcoxon rank-sum test with group serving as the independent variable and the scores from the average of the cognitive and somatic anxiety subscales of the CSAI-2R after the early portion of practice on Day 1 serving as the dependent variable. We used a Wilcoxon rank-sum test because anxiety scores were non-normally distributed and *t*-tests would not have been robust to normality violations (Field et al., 2012).

Our primary analysis of interest was a 2 (Group) x 3 (Post-Test) ANOVA with repeated measures on the last factor, and post-test radial error as the dependent variable. We predicted a Group x Post-Test interaction, which we planned to follow up with separate one-way repeated-measures ANOVAs for the AT and control groups. For the former, we predicted a non-significant effect of post-test, whereas for the latter we predicted a significant effect of post-test. We planned to follow up the significant post-test effect for the control group with one-tailed paired *t*-tests between each post-test condition for this group, and one-tailed independent *t*-tests to explore between-group differences for each post-test. We expected the control group to show a systematic increase in radial error from low-, to mild-, to high-pressure post-tests. In terms of between-group predictions, we expected no group differences for the low-pressure post-test. For the mild- and high-pressure post-tests, we expected the AT group to show lower radial error, relative to the control group.

***Mechanistic Variables During Post-Tests***

For mental effort, we ran a 2 (Group) x 3 (Post-Test) ANOVA with repeated measures on the last factor, and post-test mental effort as the dependent variable. We predicted a main effect of post-test, which we planned to follow up with one-tailed paired *t*-tests. We expect participants from both groups to show systematic increases in mental effort from low-, to mild-, to high-pressure post-test.

For challenge-threat and movement reinvestment, we ran separate 2 (Group) x 3 (Post-Test) ANOVAs with repeated measures on the last factor, and post-test challenge-threat and movement reinvestment serving as dependent variables in their respective ANOVAs. We predicted a significant Group x Post-Test interaction, which we planned to follow up with separate one-way repeated-measures ANOVAs for the AT and control groups. For the former, we predicted a non-significant effect of post-test, whereas for the latter we predicted a significant effect of post-test. We planned to follow up this significant effect for the control group with one-tailed paired *t*-tests between each post-test condition for this group, and one-tailed independent *t*-tests to explore between-group differences for each post-test. We expected the control group to show a systematic increase in perceive threat and movement reinvestment from low-, to mild-, to high-pressure post-tests. In terms of between-group predictions, we expected no group differences for the low-pressure post-test. For the mild- and high-pressure post-tests, we expected the control group to show higher perceived threat and movement reinvestment, relative to the control group.

***Putting Accuracy During Practice***

For practice performance, we conducted a 2 (Group:) x 2 (Day: 1, 2) x 3 (Phase: early, middle, late) ANOVA with repeated measures on the latter two factors, and radial error as the dependent variable. We did not have specific predictions, so we planned to follow up the ANOVA dependent upon its results.

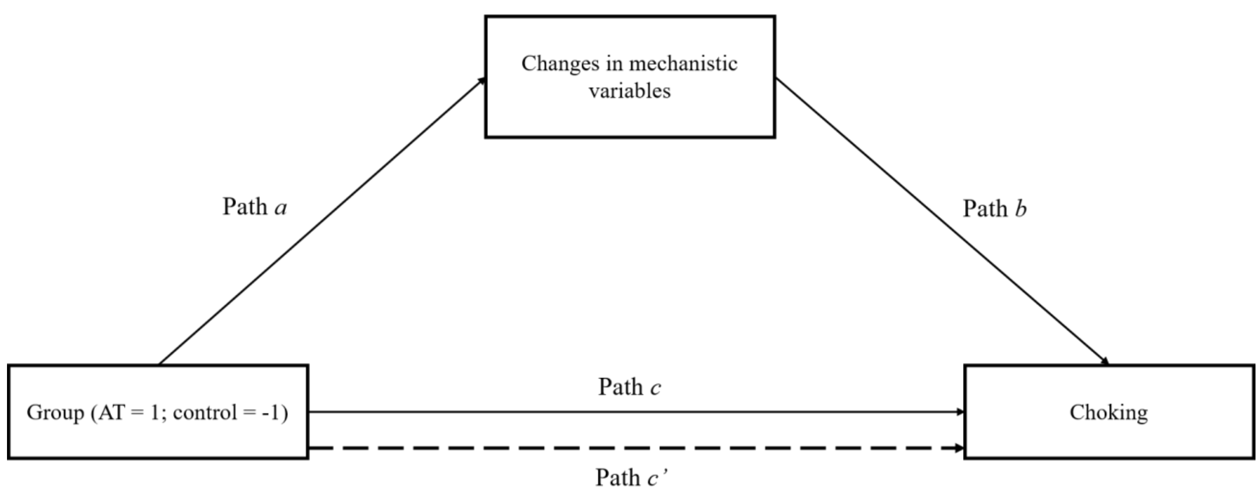
***Mechanistic Variables During Practice***

For mental effort, challenge-threat, movement reinvestment, we conducted separate 2 (Group) x 2 (Day) ANOVAs with repeated measures on the last factor. We did not have specific predictions for these variables, so we followed up the ANOVAs dependent upon their results.

***Mediation Analyses***

If group significantly affected choking under pressure and at least one mechanistic variable, we planned to conduct separate mediation analyses for each mechanistic variable significantly affected by group. Group would serve as the independent variable and be dummy coded as AT = 1 and control = -1. Choking would serve as the dependent variable and be calculated as the difference between the post-tests exhibiting the largest significant effect size difference between group. The mediating variable would be the difference in the mechanistic variable between the same post-tests used to define choking. We planned to conduct separate mediation models for each mechanistic variable because the mechanistic variables overlap conceptually and may be strongly correlated (Kenny, 2021). For each model, we planned to first test the effect of group on choking (Path *c* in Figure 2), followed by testing the effect of group on changes of the mechanistic variable (Path *a* in Figure 2). Then, we planned to test if the change in the mechanistic variable influenced choking, when controlling for group (Path *b* in Figure 2). We were interested in whether the group effect on choking could be attributed to changes in the mechanistic variables. Thus, we planned to test the significance of the indirect effect of group on choking, which is the product of Path *a* and *b* (Path *c'* in Figure 2). Accordingly, we planned to use bootstrapping (5000 iterations) to extract the 95% confidence interval for the indirect effect (Preacher & Hayes, 2008).

**Figure 2.**
*Mediation Model With Group as the Independent Variable, Changes in the Mechanistic Variables as the Mediator, and Choking as the Dependent Variable*




## Sensitivity Analysis

We conducted sensitivity analyses of the primary effect of interest (Group x Post-Test effect) excluding participants who consumed alcohol/drugs within 24 hr of the third day (post-tests), caffeine within 3 hr of the third day, and/or reported a difference of >2 hr of sleep relative to the average duration of sleep preceding the previous 2 days of data collection. In terms of outliers, we only excluded participants if one of their post-tests radial error values had a *z*-score ≥ 3.00, in which case we reported the primary statistical results with and without the inclusion of the participant (this result can be found in the supplementary material at https://osf.io/xpv8e/).

## Results

### Descriptive Data

We collected data from 70 participants. However, 6 were excluded due to dropping out ($n = 4$), playing golf more than 30 times in their lifetime ($n = 1$), and not being able to complete data collection ($n = 1$), resulting in 32 participants per group and a total sample size of 64 participants. After computing *z*-scores for low-, mild-, and high-pressure post-tests, three participants were identified as outliers and excluded. Thus, our results are based on analyses of 61 participants. Table 1 shows age, sex, and putting experience for each group.

**Table 1.**
*Age, Sex, and Putting Experience by Group*

| | Control Group ($n$ = 30; 12 females) | | Anxiety Training ($n$ = 31; 12 females) | |
|---|---|---|---|---|
| | *M* | *SD* | *M* | *SD* |
| Age (Years) | 23.67 | 3.73 | 22.61 | 3.94 |
| Lifetime Putting Experience[a] | 1.40 | 0.62 | 1.48 | 0.68 |
| Past-Year Putting Experience[a] | 0.43 | 0.63 | 0.58 | 0.56 |

*Note*: [a]0 = Never putted; 1 = Putted 1 – 10 times; 2 = Putted 11 – 20 times; 3 = Putted 21 – 30 times.

### Reliability of the Self-Reported Questionnaires

**Table 2.**
*Cronbach's Alpha for Applicable Self-Reported Measures as a Function of Study Phase*

| | Pre-Test | Practice 1 | Practice 2 | Low-Pressure | Mild-Pressure | High-Pressure |
|---|---|---|---|---|---|---|
| Cognitive anxiety subscale | .75 | .80 | .84 | .85 | .89 | .87 |
| Somatic anxiety subscale | .68 | .76 | .82 | .76 | .87 | .84 |
| Self-confidence subscale | .83 | .87 | .85 | .87 | .89 | .89 |
| Movement reinvestment | .71 | .72 | .75 | .88 | .82 | .82 |

## Pressure Manipulation Check

### *Post-Test Pressure*

The 2 (Group) x 3 (Post-Test) mixed-factor ANOVA with repeated-measures on the last factor and anxiety serving as the dependent variable showed a main effect of post-test ($F$(1.72, 101.25) = 42.37, $p$ < .001, $\eta^2_p$ = .42, $\varepsilon$ = .83). There was no main effect of group ($p$ = .829, $\eta^2_p$ < .01) and no Group x Post-Test interaction ($p$ = .337, $\eta^2_p$ = .02). Because anxiety scores were non-normally distributed and *t*-tests would not have been robust to normality violations (Field et al., 2012), we followed up with one-sided Wilcoxon signed-rank tests to investigate within-subject differences between each post-test condition. Results revealed that anxiety scores were higher in the high- (*M* = 16.18, *SD* = 5.58), compared to mild- (*M* =13.18, *SD* = 4.50; *Z* = -5.30, *p* < .001) and low-pressure (*M* = 11.50, *SD* = 2.83; *Z* = -5.91, *p* < .001) post-tests. Additionally, anxiety was higher in the mild-, compared to the low-pressure (*Z* = -3.55, *p* < .001) post-test.

The 2 (Group) x 3 (Post-Test) mixed-factor ANOVA with repeated-measures on the last factor and HR[8] serving as the dependent variable showed a main effect of post-test ($F$(2, 116) = 20.69, $p$ < .001, $\eta^2_p$ = .26). There was no main effect of group ($p$ = .075, $\eta^2_p$ = .05) and no Group x Post-Test interaction ($p$ = .114, $\eta^2_p$ = .04). We followed up with one-tailed paired *t*-tests to investigate within-subject differences between each post-test condition. Results revealed that HR was higher in the high- (*M* = 88.78, *SD* = 13.66), compared to mild- (*M* = 86.27, *SD* 13.44; *t*(59)= -4.47, *p* < .001) and low-pressure (*M* = 84.88, *SD* = 12.78; *t*(59) = -5.74, *p* < .001) post-tests. Additionally, HR was higher in the mild-, compared to the low-pressure (*t*(59) = -2.26, *p* = .014) post-test. Since participants showed systematic increases in anxiety and HR across the post-tests, the post-test pressure manipulation was considered successful.

[8] HR data from one participant is missing (ID 36) due to connection issues with the HR monitor device.

### *Group Pressure During Practice*

The Wilcoxon rank-sum test with group serving as the independent variable and anxiety during Day 1 serving as the dependent variable revealed that the AT group presented higher anxiety levels, compared to the control group ($Z$ = -3.89, $p < .01$). The one-way (Group: AT vs. control) ANCOVA with HR during early practice on Day 1, and pre-test HR[9] serving as the covariate did not show a main effect of group ($p$ = .806, $\eta^2_p < .01$). Since anxiety was greater for the AT group than the control group during early practice on Day 1, the group pressure manipulation was considered successful.

## Putting Accuracy

### *Post-Tests*

The primary analysis, a 2 (Group) x 3 (Post-Test) mixed-factor ANOVA with repeated-measures on the last factor and post-test radial error serving as the dependent variable showed a Group x Post-Test interaction ($F(1.76, 103.84) = 8.38$, $p < .001$, $\eta^2_p = .12$, $\varepsilon = .88$). There was no main effect of group ($p$ = .124, $\eta^2_p = .04$) or post-test ($p$ = .067, $\eta^2_p = .05$). The subsequent one-way ANOVA for the AT group did not reveal a main effect of post-test ($p$ = .550, $\eta^2_p = .02$). The one-way ANOVA for the control group revealed a main effect of post-test ($F(2, 58) = 13.38$, $p < .001$, $\eta^2_p = .32$), which we followed up with one-tailed paired $t$-tests between each post-test condition. Results showed that the radial error was lower in the low- ($M$ = 40.64, $SD$ = 8.21), compared to the mild- ($M$ = 44.90, $SD$ = 13.28; t(29) = -2.10, $p$ = .022) and high-pressure ($M$ = 53.16, $SD$ = 16.01; $t(29) = -4.60$, $p < .001$) post-tests. Additionally, the radial error was lower in the mild- compared to the high-pressure post-test ($t(29) = -3.20$, $p$ = .002) (Figure 3). We also conducted one-tailed independent $t$-tests to assess potential between-group differences in radial

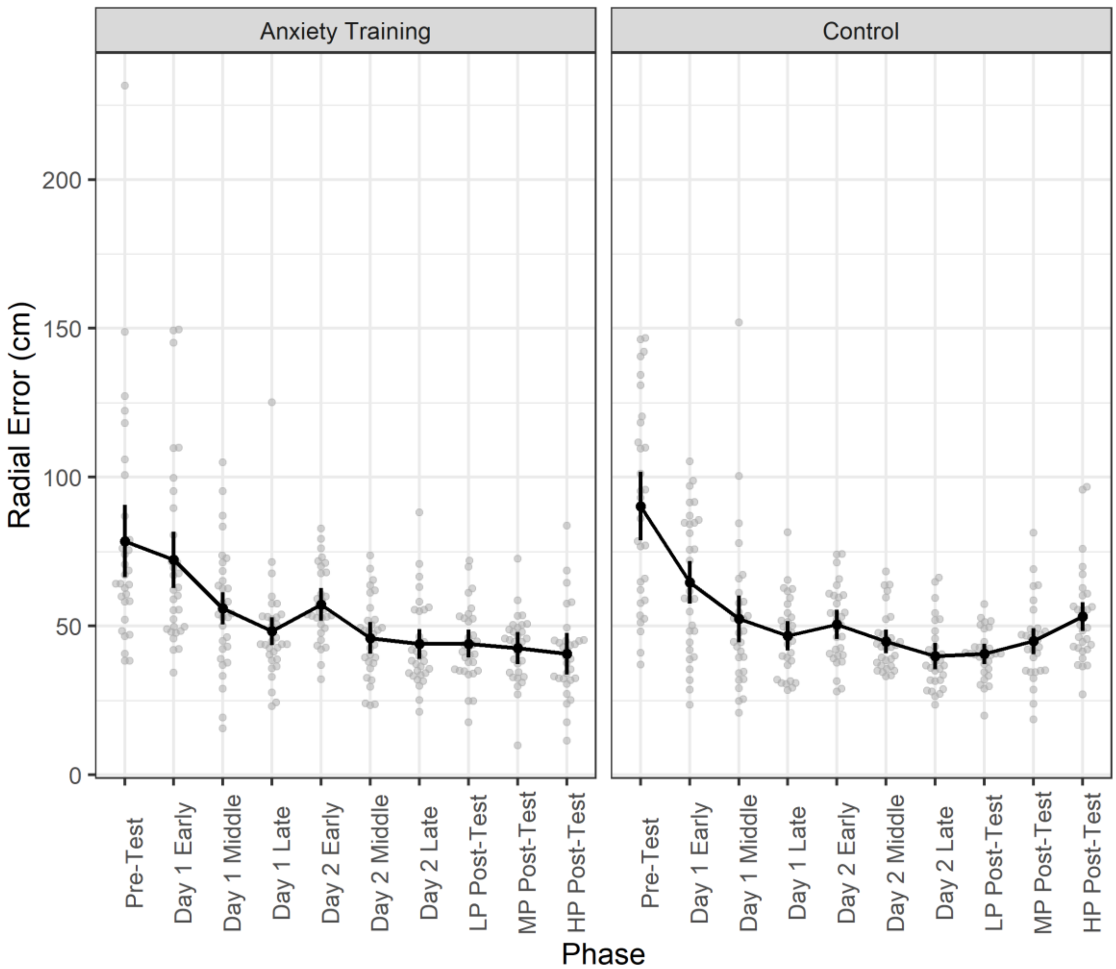
[9] For pre-test, HR data is missing from one participant (ID 70) due to connection issues with the HR monitor device.

error for each post-test. Results revealed that participants in the AT group presented lower radial error in the high-pressure post-test, compared to the control group (AT: $M = 40.72$, $SD = 14.82$; control: $M = 53.16$, $SD = 16.01$; $t(59) = -3.15$, $p = 001$). There were no between-group differences for the low- and mild-pressure post-tests ($p \geq .108$).

***Practice***

The 2 (Group) x 2 (Day) x 3 (Block) mixed-factor ANOVA with repeated-measures on the last factor and practice radial error serving as the dependent variable showed a main effect of day ($F(1, 59) = 25.12$, $p < .001$, $\eta^2_p = .30$), such that participants presented lower radial error on day 2, relative to day 1. There was also a main effect of block ($F(2, 118) = 44.34$, $p < .001$, $\eta^2_p = .43$). There was no main effect of Group ($p = .179$, $\eta^2_p = .03$) and no interactions ($p$s $\geq .084$). We followed up the main effect of block with two-tailed paired *t*-tests with a Bonferroni correction (i.e., significance set as $p = .017$). Results showed that radial error during the early blocks ($M = 61.25$, $SD = 16.59$) was higher, compared to the middle ($M = 49.79$, $SD = 14.62$; $t(60) = 5.76$, $p < .001$) and late blocks ($M = 44.74$, $SD = 11.33$; $t(60) = 8.98$, $p < .001$). Additionally, radial error was lower in the late blocks compared to the middle blocks ($t(60) = 3.32$, $p = .005$) (Figure 3).

**Figure 3.**
*Putting Accuracy for Each Group as a Function of the Study Phase*




*Note.* Radial error (y-axis) represents putting accuracy, with lower values indicating greater accuracy. Study phase is presented on the x-axis and shows pre-test, days 1 and 2 with their respective blocks (early, middle, and late), as well as low-pressure (LP), mild-pressure (MP), and high-pressure (HP) post-tests. Each large black data point represents the average radial error for each study phase. Smaller grey data points represent individual values of radial error. Error bars represent within-subject standard errors.

## Mechanistic Variables

### *Post-Tests*

The 2 (Group) x 3 (Post-Test) mixed-factor ANOVA with repeated-measures on the last factor and post-test mental effort serving as the dependent variable showed a main effect of post-test ($F$(1.47, 86.60 = 67.60, $p < .001$, $\eta^2_p = .53$, $\varepsilon = .73$), which was superseded by a Group x

Post-Test interaction ($F$(1.47, 86.60) = 5.25, $p$ = .013, $\eta^2_p$ = .08, $\varepsilon$ = .73). There was no main effect of group ($p$ = .598, $\eta^2_p$ < .01). We followed up with one-way ANOVAs for each group. The ANOVA for the AT group revealed a main effect of post-test ($F$(1.54, 46.24) = 27.49, $p$ < .001, $\eta^2_p$ = .48, $\varepsilon$ = .77). We then used one-tailed $t$-tests to investigate within-subject differences between each post-test condition. Results revealed that mental effort was higher in the high- ($M$ = 91.84, $SD$ = 26.44), compared to both mild- ($M$ = 80.87, $SD$ = 25.13; ($t$(30) = -6.60, $p$ < .001) and low-pressure ($M$ = 76.77, $SD$ = 25.12; $t$(30) = -5.80, $p$ < .001) post-tests. Additionally, mental effort was higher in the mild-, compared to low-pressure post-test ($t$(30) = -2.12, $p$ = .021). The one-way ANOVA for the control group also revealed a main effect of post-test ($F$(1.39, 40.33) = 41.50, $p$ < .001, $\eta^2_p$ = .59, $\varepsilon$ = .70), which we followed up with one-tailed $t$-tests. Similar to the AT group, results for the control group revealed that mental effort was higher in the high- ($M$ = 91.10, $SD$ = 26.43), compared to both mild- ($M$ = 80.67, $SD$ = 29.85; $t$(29) = 6.29, $p$ < .001) and low-pressure post-tests ($M$ = 66.80, $SD$ = 33.44; $t$(29) = 7.37, $p$ < .001). Additionally, mental effort was higher in the mild- compared to the low-pressure post-test ($t$(29) = 4.95, $p$ < .001). We also conducted $t$-tests to identify potential between-group differences for each post-test, but results did not reveal any statistically significant result ($p$s ≥ .194) (Figure 4A).

The 2 (Group) x 3 (Post-Test) mixed-factor ANOVA with repeated-measures on the last factor and post-test challenge-threat appraisal serving as the dependent variable showed a main effect of post-test ($F$(1.77, 104.25) = 23.38, $p$ < .001, $\eta^2_p$ = .28, $\varepsilon$ = .88). There was no main effect of group ($p$ = .734, $\eta^2_p$ < .01) and no Group x Post-Test interaction ($p$ = .227, $\eta^2_p$ = .03). Because challenge-threat appraisal data were non-normally distributed, we used one-sided Wilcoxon signed-rank tests to investigate within-subject differences between each post-test condition. Results revealed that challenge-threat appraisal was higher (i.e., higher scores indicate higher perceived threat and lower perceived challenge, whereas lower scores indicate the opposite) in the high- (*M* = 2.21, *SD* = 1.71*)*, compared to both mild- (*M* = 2.02, *SD* = 1.72; *Z* = -4.27, *p* < .001) and low-pressure post-tests (*M* = 1.89, *SD* = 1.71; *Z* = -5.21, *p* < .001). Additionally, challenge-threat appraisal was higher in the mild- compared to the low-pressure post-test (*Z* = -2.95, *p* = .002) (Figure 4B).

The 2 (Group) x 3 (Post-Test) mixed-factor ANOVA with repeated-measures on the last factor and post-test movement reinvestment serving as the dependent variable showed a main effect of post-test ($F(1.72, 101.42) = 8.02$, $p = .001$, $\eta^2_p = .12$, $\varepsilon = .86$), which was superseded by a Group x Post-Test interaction ($F(1.72, 101.42) = 3.31$, $p = .048$, $\eta^2_p = .05$, $\varepsilon = .86$). There was no main effect of group ($p = .900$, $\eta^2_p < .01$). We followed up with one-way ANOVAs for each group. The ANOVA for the AT group did not reveal a main effect of post-test ($F(1.67, 50.06) = 2.94$, $p = .071$, $\eta^2_p = .09$, $\varepsilon = .83$). Conversely, the ANOVA for the control group did show a main effect of post-test ($F(1.46, 42.42) = 6.20$, $p = .009$, $\eta^2_p = .18$, $\varepsilon = .73$), which we followed up with one-sided Wilcoxon signed-rank tests to investigate within-subject differences between each post-test condition, because reinvestment data were non-normally distributed. Results revealed that movement reinvestment was higher in the high- (*M* = 2.88, *SD* = 1.60*)*, compared to mild- (*M* = 2.77, *SD* = 1.70*; Z* = -1.95, *p* = .025) and low-pressure post-test (*M* = 2.51, *SD* = 1.74; *Z* = -3.02, *p* = .001). Additionally, movement reinvestment was higher in the mild-, compared to the low-pressure post-test (*Z* = -2.58, *p* = .005). We also conducted one-sided Wilcoxon rank-sum tests to identify potential between-group differences for each post-test. Results did not reveal any statistically significant result (*p*s ≥ .314) (Figure 4C).

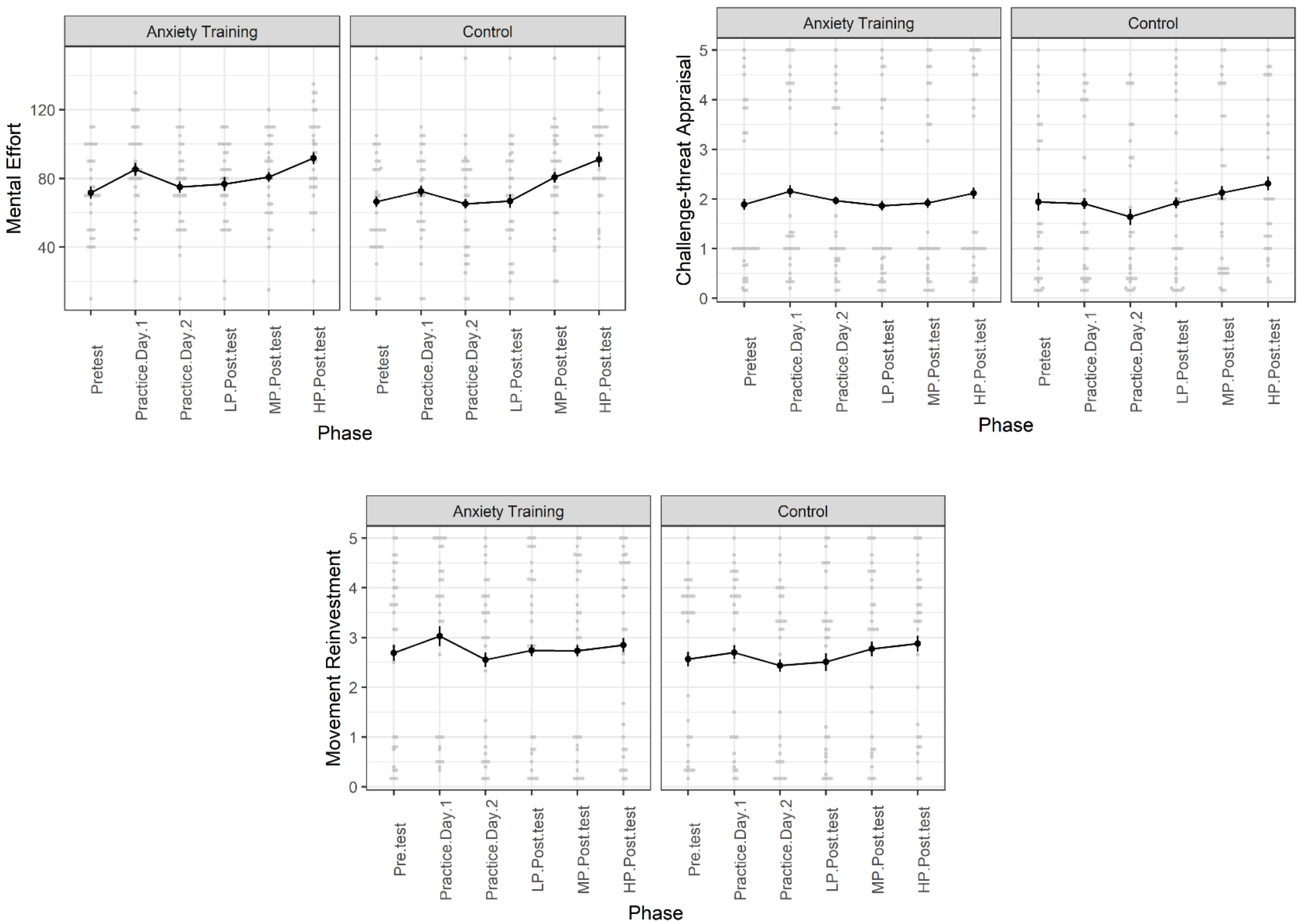
***Practice***

The 2 (Group) x 2 (Day: first, second day of practice) ANOVA with repeated measures on the last factor and mental effort serving as dependent variable revealed a main effect of day ($F(1, 59) = 29.51$, $p < .001$, $\eta^2_p = .33$), such that participants reported higher mental effort in the first day, compared to the second day. There was no main effect of group ($p$ =. 113, $\eta^2_p = .04$) and no Group x Day interaction ($p = .343$, $\eta^2_p = .02$).

The 2 (Group) x 2 (Day) ANOVA with repeated measures on the last factor and challenge-threat appraisal serving as dependent variable revealed a main effect of day ($F(1, 59) = 17.44$, $p < .001$, $\eta^2_p = .23$), such that participants reported higher perceived threat/lower perceived challenge in the first day, compared to the second day. There was no main effect of group ($p$ =. 497, $\eta^2_p < .01$) and no Group x Day interaction ($p$ =.494, $\eta^2_p < .01$).

The 2 (Group) x 2 (Day) ANOVA with repeated measures on the last factor and movement reinvestment serving as dependent variable revealed a main effect of Day ($F(1, 59) = 27.87$, $p < .001$, $\eta^2_p = .32$), such that participants reported higher movement reinvestment in the first day, compared to the second day. There was no main effect of group ($p$ =. 605, $\eta^2_p < .01$) and no Group x Day interaction ($p$ =.133, $\eta^2_p = .04$).

**Figure 4.**
***Mechanistic Variables for Each Group as a Function of the Study Phase***




*Note*. Mental effort (Panel A), challenge-threat appraisal (Panel B), and movement reinvestment (Panel C) are depicted in the y-axis. Study phase (x-axis) shows pre-test, the first and second practice sessions, as well as low-pressure (LP), mild-pressure (MP), and high-pressure (HP) post-tests. Each large black data point represents the average of each mechanistic variable for each study phase. Smaller grey data points represent individual values for each mechanistic variable. Error bars represent within-subject standard errors.

## Mediation Analysis

Choking, our dependent variable, was considered the difference in radial error between low- and high-pressure post-tests, since this difference exhibited the largest effect size between groups (Cohen's $d$ = 0.94). The independent variable, group, was coded as AT = 1 and control = -

1. We did not proceed with the mediation analysis for challenge-threat appraisal because it was not affected by group (i.e., no main effect of group and no Group x Post-Test interaction). Since mental effort and movement reinvestment showed a significant Group x Post-Test interaction, we conducted two separate mediation models with changes in mental effort and reinvestment from low- to high-pressure as the mediator variable.

For the model with change in mental effort as the mediator, group predicted choking (Path *c*; $\beta_{unstandardized} = 8.39$ cm, $p < .001$), and change in mental effort (Path *a*; $\beta_{unstandardized} = 4.62$, $p = .031$). Change in mental effort did not predict choking when controlling for group (Path *b*; $\beta_{unstandardized} = -0.10$ cm, $p = .461$). Group did not have an indirect effect on choking via change in mental effort ($\beta_{unstandardized} = -0.46$ cm, Lower *CI* = -2.14 cm, Upper *CI* = 0.97 cm).

For the model with change in movement reinvestment as the mediator, group predicted choking (Path *c*; $\beta_{unstandardized} = 8.39$ cm, $p < .001$), and change in movement reinvestment (Path *a*; $\beta_{unstandardized} = 0.13$, $p = .040$). Change in movement reinvestment did not predict choking when controlling for group (Path *b*; $\beta_{unstandardized} = 6.39$ cm, $p = .162$). Group did not have an indirect effect on choking via change in mental effort ($\beta_{unstandardized} = -0.83$ cm, Lower *CI* = -7.54 cm, Upper *CI* = 3.16 cm).

**Secondary Analyses**

We reran the primary analysis including the outliers and the primary analysis excluding participants who failed to abstain from caffeine within 3 hr or alcohol within 24 hr prior to the third day (post-tests), and/or reported a difference of >2 hr of sleep relative to the average duration of sleep preceding the previous 2 days of data collection. Moreover, we investigated changes in self-confidence and HRV throughout practice and post-tests, as well as changes in HR from the first to the second day of practice. All secondary analysis results are available in the supplementary material at https://osf.io/xpv8e/.

## Discussion

Having learners practice a motor skill under psychological pressure has been shown to prevent choking under pressure in a subsequent post-test with the same pressure level as the practice phase (Beseler et al., 2016; Oudejans et al., 2011; Oudejans & Pijpers, 2009, 2010). However, only one study to date investigated whether this advantage can be transferred to a higher-stakes environment (Oudejans & Pijpers, 2010). Moreover, the AT literature is mostly based on the ACT (Alder et al., 2016; Oudejans & Pijpers, 2009, 2010), which has limited investigations into the mechanisms underlying the AT advantage, leaving them unclear (Kegelaers & Oudejans, 2022). Therefore, the present study examined whether practicing a motor skill under mild levels of pressure prevents motor performance decrement under higher levels of pressure, as well as potential mechanistic variables, based on different theories of choking. Our results showed that participants in the AT group maintained performance across low-, mild-, and high-pressure post-tests, whereas participants in the control group reduced performance (i.e., choked under pressure) from low- to mild-pressure, and from mild- to high-pressure post-tests. We not only replicated previous findings on AT, but we also extended the current literature by providing evidence that the advantage of AT can be transferred to a higher-stakes condition, and we did so with a much larger sample size, compared to typical AT experiments (Alder et al., 2016; Oudejans & Pijpers, 2009, 2010). However, we failed to provide strong evidence in favor of mental effort, movement reinvestment, or perceived challenge/threat as potential mediators of the association between AT and choking, although the former two were affected by AT, rendering them the most likely candidates to explain its benefits. Results will be discussed considering these two mechanistic variables and their associated theories (ACT and reinvestment theory, respectively), but not perceived challenge threat or its associated theory (BPSM of challenge and threat), since it was not affected by group.

Our results revealed a Group x Post-Test interaction for mental effort. Follow-up analyses showed a main effect of post-test for both AT and control groups, such that participants systematically increased mental effort across post-tests, which supports the ACT prediction that increases in anxiety are accompanied by increases in mental effort. Since mental effort was affected by group, we examined whether it mediated the relationship between AT and choking, but results did not support this hypothesis. A possible reason is that increased mental effort itself does not necessarily translate into an effective mechanism to prevent performance decrements. Instead, it might reflect adaptive processes, such as improved visual search behavior (Alder et al., 2016), in some participants and maladaptive processes, such as worries about performance outcomes, in others. This suggests that, although mental effort did not explain choking in our study, it may affect processes that affect performance under pressure, but which were not measured in our experiment. Such processes (e.g., visual search behavior) could be examined in future AT studies. Another possible explanation is that the mental effort scale we used was not precise enough to correlate individual-level changes in mental effort with individual-level changes in performance, precluding an association between mental effort and choking. In fact, research has suggested other methods to measure mental effort, such as changes in pupil size (van der Wel & Steenbergen, 2018) and electroencephalography variables (e.g., theta and alpha frequency; Zhu, Wang & Zhang, 2021). Future AT studies are encouraged to combine self-reported and physiological measurements to capture a broader picture of changes in mental effort and examine its relationship with choking. Nonetheless, our study suggests that AT benefits performance under pressure concomitant to smaller increases in mental effort, making it a candidate mechanism to explain the AT benefit on performance.

Results also revealed a Group x Post-Test interaction for movement reinvestment. Follow-up analyses showed a main effect of post-test for the control group, but not for the AT group. Specifically, participants in the control group systematically increased movement reinvestment across post-tests. Since movement reinvestment was affected by group, we examined whether it mediated the relationship between AT and choking, but it did not. Our results are consistent with a recent systematic review of 29 studies that measured performance under pressure and movement reinvestment (Sullivan et al., 2022). The review showed that pressure generally increased movement reinvestment, but only two studies assessed whether it mediated performance under pressure, with neither study finding that it did. Importantly, other researchers have claimed that movement reinvestment may not necessarily lead to choking (Roberts, Jackson & Grundy, 2019). For instance, Toner and Moran (2011) argued that golfers can reinvest in technical changes to their putting stroke without losing proficiency, and Hill et al. (2010) reported that some golfers reinvest declarative knowledge in their putting stroke to mitigate poor scoring. An important limitation of reinvestment studies is the lack of detail about the movement reinvestments that could be harmful or helpful to performance (Roberts, Jackson & Grundy, 2019). Therefore, the increased movement reinvestment for control group participants in our study may not necessarily lead to choking. Importantly, this is the first experiment to show that AT simultaneously affects choking and movement reinvestment, and suggests that movement reinvestment, like mental effort, is a plausible candidate to explain the benefits of AT to performance under pressure.

The present study adds important insights to the AT literature as it is methodologically strong with a large sample size, and it has pre-registered hypotheses and analyses. Our study also provides ecological validity as we used a motor skill that is common around the world and pressure manipulations that can be easily implemented. Moreover, our design reflects the reality in sports environments such that competition usually, if not always, elicits levels of anxiety that are difficult to reproduce during training. By adding the high-pressure post-test, we were able to investigate whether the benefits of training under mild levels of anxiety are in fact transferrable to higher-stakes environments as expected in most competitions. With that said, a few limitations are worth mentioning. First, our sample consisted of young, healthy adults; thus, we cannot generalize the results to other populations, such as individuals who are used to performing in high-stakes environments (e.g., professional athletes), but who were intentionally excluded from participation in this study. Second, our intervention was considered as having a “medium” duration (i.e., two or more sessions; Low et al., 2021), so our results cannot be generalized beyond that AT length. Finally, we only investigated three potential mediators of the relationship between AT and choking, and the way these mediators were measured may be improved, for example by measuring processes reflected by increases in mental effort.

Despite AT being a promising intervention, the literature is still in its early stages with some gaps that should be addressed. First, the mechanisms through which AT prevents choking are not clear. Here, we provide limited evidence for two potential candidates (i.e., mental effort and movement reinvestment), but there are other mechanisms that could explain the benefits of AT. For example, AT has been posited to increase mental toughness and coping skills, such as self-talk, which could explain performance benefits (Kegelaers & Oudejans, 2022). Therefore, future studies examining AT should not only look at its efficacy in performance but also continue investigating potential mechanistic variables. Second, AT has possible downsides. For example, AT could be misused and create a destructive sport culture (Owusu-Sekyere & Gervis, 2016). Future work should investigate how AT can impact well-being and mental health to establish concrete pitfalls of practicing under pressure. Third, AT studies have mostly relied on brief interventions (Low et al., 2021), which have been shown to promote beneficial results on performance. However, longer interventions seem to have smaller effect sizes (Low et al., 2021) possibly due to the impact of a specific pressure manipulation losing its efficacy when used for a prolonged duration (Kagalaers & Oudejans, 2022). Thus, future experiments should investigate optimal AT duration.

Based on the current state of the literature, Fletcher and Arnold (2021) have recently proposed guidelines for implementing AT. They propose that coaches should work together with sports psychologists to design and discuss specific aims and how AT would be implemented for a specific skill or sport. Fletcher and Arnold contend that the pressure level during training should be gradually increased and modified as necessary and that practitioners should evaluate the efficacy of AT on performance and have athletes engage in reflective activities (e.g., awareness of their emotions, behavior under pressure, and personal stress triggers). Additionally, in real-world settings, pressure levels could be increased by manipulating training consequences, such as missing practice time, place in the team, and having spectators, as long these manipulations do not prompt an abusive environment (Kagalaers & Oudejans, 2022; for a complete list of possible ways to manipulate pressure during AT, see Kagelaers, Wylleman & Oudejans, 2020). Nonetheless, although recommendations are beginning to appear, it is important to stress that the AT literature is in its early stages and more research is warranted to better understand how to apply AT.

## Conclusion

In conclusion**,** the present study showed that AT is beneficial to performance under pressure, even to higher levels of pressure than experienced during training. Although we did not find evidence that mental effort, movement reinvestment, or perceived challenge/threat mediated the relationship between AT and choking, the former two were affected by AT, making them candidate mechanisms for explaining the AT performance benefit. We suggest that practitioners have learners practice under pressure in consultation with sport psychologists and that researchers further investigate the mechanisms underlying AT performance benefits.

## References


Alder, D., Ford, P. R., Causer, J., & Williams, A. M. (2016). The Effects of High- and Low-Anxiety Training on the Anticipation Judgments of Elite Performers. *Journal of Sport & Exercise Psychology*, 38(1), 93–104. doi:10.1123/jsep.2015-0145.

Allsop, J., & Gray, R. (2014). Flying under pressure: Effects of anxiety on attention and gaze behavior in aviation. *Journal of Applied Research in Memory and Cognition*, 3(2), 63–71. doi: 10.1016/j.jarmac.2014.04.010.

Bargh, J. A., & Schul, Y. (1980). On the cognitive benefits of teaching. *Journal of Educational Psychology*, 72, 593 – 604. doi: 10.1037/0022-0663.72.5.593.

Batista, M. T. S., Figueiredo, L. S., Martins, C. A., Nogueira, N. G. D. H. M., Ferreira, B. D. P., & Benda, R. N. (2022). Expectation of teaching and self-controlled KR in motor skills acquisition: Are there additive effects? *Research Quarterly for Exercise and Sport*, 1-10. doi: 10.1080/02701367.2021.1999891.

Baumeister, R. F. (1984). Choking under pressure: Self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of Personality and Social Psychology*, 46(3), 610–620. doi: 10.1037//0022-3514.46.3.610.

Baumeister, R. F., & Showers, C. J. (1986). A review of paradoxical performance effects: Choking under pressure in sports and mental tests. *European Journal of Social Psychology*, 16(4), 361–383. doi: 10.1002/ejsp.2420160405.

Beilock, S. L., & Carr, T. H. (2001). On the fragility of skilled performance: What governs choking under pressure? *Journal of Experimental Psychology*. General, 130(4), 701–725. doi: 10.10.1037//0096-3445.130.4.70 1 1.

Beilock, S. L., & Carr, T. H. (2005). When High-Powered People Fail: Working Memory and "Choking Under Pressure" in Math. *Psychological Science*, 16(2), 101–105. doi:10.1111/j.0956-7976.2005.00789.x.

Beilock, S. L., & Decaro, M. S. (2007). From poor performance to success under stress: Working memory, strategy selection, and mathematical problem solving under pressure. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 33(6), 983–998. doi:10.1037/0278-7393.33.6.983.

Beilock, S. L., Kulp, C. A., Holt, L. E., & Carr, T. H. (2004). More on the fragility of performance: Choking under pressure in mathematical problem solving. *Journal of Experimental Psychology*. General, 133(4), 584–600. doi:10.1037/0096-3445.133.4.584.

Beilock, S. L., Wierenga, S. A., & Carr, T. H. (2002). Expertise, attention, and memory in sensiormotor skill execution: Impact of novel task constraints on dual-task performance and episodic memory. *The Quarterly Journal of Experimental Psychology*, 55A, 12 – 1240. doi: 10.1080/02724980244000170.

Bellomo, E., Cooke, A., & Hardy, J. (2018). Chunking, Conscious Processing, and EEG During Sequence Acquisition and Performance Pressure: A Comprehensive Test of Reinvestment Theory. *Journal of Sport & Exercise Psychology*, 40(3), 135–145. doi: 10.1123/jsep.2017-0308.

Benware, C. A., & Deci, E. L. (1984). Quality of learning with an active versus passive motivational set. *American Educational Research Journal*, 21, 755 – 765. doi: 10.3102/00028312021004755.

Beseler, B., Mesagno, C., Young, W., & Harvey, J. (2016). Igniting the pressure acclimatization training debate: Contradictory pilot-study evidence from Australian football. *Journal of Sport Behavior*, 39(1), 22–38.

Blascovich, J. (2008). Challenge, threat, and health. *In Handbook of motivation science* (pp. 481–493). The Guilford Press.

Blascovich, J., & Tomaka, J. (1996). The Biopsychosocial Model of Arousal Regulation. In M. P. Zanna (Ed.), Advances in Experimental Social Psychology (Vol. 28, pp. 1–51). *Academic Press*. doi:10.1016/S0065-2601(08)60235-X.

Bobrownicki, R., Carson, H. J., & Collins, D. (2022). Conducting systematic reviews of applied intervations: A comment on Cabral et al. (2022). *Sport, Exercise, and Performance Psychology*, 11, 264 – 274. doi: 10.1037/spy0000299.

Bobrownicki, R., Carson, H. J., MacPherson, A. C., & Collins, D. (2021). Unloading the dice: Selection and design of comparison and control groups in controlled trials to enhance translational impact within motor learning and control research. *International Journal of Sport and Exercise Psychology*,1 – 15. doi: 10.1080/1612197X.2021.1956567.

Bobrownicki, R., Collins, D., Sproule, J., & MacPherson, A. C. (2018). Redressing the balance: Commentary on “Examining motor learning in older adults using analogy instruction” by Tse, Wong, and Masters (2017). *Psychology of Sport and Exercise*, 38, 211 – 214. doi: 10.1016/j.psychsport.2018.05.014.

Bobrownicki, R., MacPherson, A. C., Coleman, S. G. S., Collins, D., & Sproule, J. (2015). Re-examining the effects of verbal instructional type on early stage motor learning. *Human Movement Science*, 44, 168–181. doi:10.1016/j.humov.2015.08.023.

Bracken, M. B. (1992). Statistical methods for analysis of effects of treatment in overviews of randomized trials. Effective Care of the Newborn Infant. Oxford University Press, Oxford.

Bright, J. E. H., & Freedman, O. (1998). Differences between implicit and explicit acquisition of a complex motor skill under pressure: An examination of some evidence. *British Journal of Psychology*, 89(2), 249–263. doi:10.1111/j.2044-8295.1998.tb02683.x.

Cabral, D. A. R., Daou, M., Bacelar, M. F. B., Parma, J. O., & Miller, M. W. (2022). Does learning a skill with the expectation of teaching it impair the skill's execution under psychological pressure if the skill is learned with analogy instructions? *Psychology of Sport and Exercise*, 51, 101757. doi:10.1016/j.psychsport.2020.101757.

Cabral, D. A. R., Wilson, A. E., & Miller, M. W. (2022). The effect of implicit learning on motor performance under psychological pressure: A systematic review and meta-analysis. *Sport, Exercise, and Performance Psychology*. doi: 10.1037/spy0000286.

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., Boardley, I., Bouza, B., Cheval, B., Chow, Z. R., Contreras, B., Dieter, B., Halperin, I., Haun, C., Knudson, D., Lahti, J., … 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. doi:10.1007/s40279-019-01227-1.

Chen, K. W., Berger, C. C., Manheimer, E., Forde, D., Magidson, J., Dachman, L., & Lejuez, C. W. (2012). Meditative therapies for reducing anxiety: A systematic review and meta-analysis of randomized controlled trials. *Depression and Anxiety*, 29, 545 – 562. doi: 10.1002/da.21964.

Cochran, W. G. (1954). The Combination of Estimates from Different Experiments. Biometrics, 10(1), 101–129. JSTOR. doi:10.2307/3001666.

Cohen, J. (1962). The statistical power of abnormal-social psychological research: A review. *The Journal of Abnormal and Social Psychology*, 65(3), 145–153. doi:10.1037/h0045186.

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Routledge.

Cooke, A., Kavussanu, M., McIntyre, D., Boardley, I. D., & Ring, C. (2011). Effects of competitive pressure on expert performance: Underlying psychological, physiological, and kinematic mechanisms. *Psychophysiology*, 48(8), 1146–1156. doi:10.1111/j.1469-8986.2011.01175.x.

Cox, R. H., Martens, M. P., & Russell, W. D. (2003). Measuring anxiety in athletics: The Revised Competitive State Anxiety Inventory–2. *Journal of Sport and Exercise Psychology*, 25, 519 – 533. doi: 10.1123/jsep.25.4.519.

Daou, M., Buchanan, T. L., Lindsey, K. R., Lohse, K. R., & Miller, M. W. (2016). Expecting to teach enhances learning: Evidence from a motor learning paradigm. *Journal of Motor Learning and Development*, 4, 197 – 207. doi: 10.1123/jmld.2015-0036.

Daou, M., Hutchison, Z., Bacelar, M., Rhoads, J. A., Lohse, K. R., & Miller, M. W. (2019). Learning a skill with the expectation of teaching it impairs the skill's execution under psychological pressure. *Journal of Experimental Psychology*: Applied, 25(2), 219–229. doi:10.1037/xap0000191.

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. doi: 10.1016/j.ijpsycho.2018.02.013.

Daou, M., Lohse, K. R., & Miller, M. W. (2016). Expecting to teach enhances motor learning and information processing during practice. *Human Movement Science*, 49, 336 – 345. doi: 10.1016/j.humov.2016.08.009.

Daou, M., Rhoads, J. A., Jacobs, T., Lohse, K. R., & Miller, M. W. (2019). Does limiting pre-movement time during practice eliminate the benefit of practicing while expecting to teach? *Human Movement Science*, 64, 153 – 163. doi: 10.1016/j.humov.2018.11.017.

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. doi: 10.1037/a0023466.

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

Elliot, D., Polman, R., & Taylor, J. (2014). The effects of relaxing music for anxiety control on competitive sport anxiety. *European Journal of Sport Science*, 14, 296 – 301. doi: 10.1080/17461391.2012.693952.

Englert, C., & Seiler, R. (2020). Anxiety and volleyball performance in a real testing situation: Gender differences in competitive environments. *Educational Psychology*. doi:10.1080/01443410.2019.1706723.

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: Attentional control theory. *Emotion*, 7(2), 336–353. doi:10.1037/1528-3542.7.2.336.

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. doi: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, 175 – 191. doi: 10.3758/BF03193146.

Fawler, B., Beatty, G., Mann, D., & Janelle, C. (2020). Staying cool under pressure: Developing and maintaining emotional expertise in sport. In Skill acquisition in sports: Research, theory and practice (Vol. 3).

Feinberg, J. M., & Aiello, J. R. (2010). The Effect of Challenge and Threat Appraisals Under Evaluative Presence. *Journal of Applied Social Psychology*, 40(8), 2071–2104. doi:10.1111/j.1559-1816.2010.00651.x.

Fekih, S., Zguira, M. S., Koubaa, A., Bettaieb, A., Hajji, J., Bragazzi, N. L., & Jarraya, M. (2021). Effects of mental training through imagery on the competitive anxiety of young tennis players fasting during Ramadan. *Frontiers in Nutrition*, 8, 1 - 9. doi: 10.3389/fnut.2021.712296.

Field, A., Miles, J., & Field, Z. (2012). Discovering Statistics Using R. London: SAGE Publications.

Fiorella, L., & Mayer, R. E. (2013). The relative benefits of learning by teaching and teaching expectancy. *Contemporary Educational Psychology*, 38, 281 – 288. doi: 10.1016/j.cedpsych.2013.06.001.

Fiorella, L., & Mayer, R. E. (2014). Role of expectations and explanations in learning by teaching. *Contemporary Educational Psychology*, 39, 75 – 85. doi: 10.1016/j.cedpsych.2014.01.001.

Fletcher, D., & Arnold, R. (2021). Stress and pressure training. In *Stress, well-being, and performance in sport* (pp. 261-296). Routledge.

Georgiou, K., Larentzakis, A. V., Khamis, N. N., Alsuhaibani, G. I., Alaska, Y. A., & Giallafos, E. J. (2018). Can Wearable Devices Accurately Measure Heart Rate Variability? *A Systematic Review*. Folia Medica, 60(1), 7–20. doi:10.2478/folmed-2018-0012.

Gilgen-Ammann, R., Schweizer, T., & Wyss, T. (2019). RR interval signal quality of a heart rate monitor and an ECG Holter at rest and during exercise. *European Journal of Applied Physiology*, 119(7), 1525–1532. doi:10.1007/s00421-019-04142-5.

Glen, S. (2016). Cook's Distance / Cook's D: Definition, Interpretation. Elementary Statistics for the Rest of Us! https://www.statisticshowto.com/cooks-distance/.

Grand, K. F., Daou, M., Lohse, K. R., & Miller, M. W. (2017). Investigating the Mechanisms Underlying the Effects of an Incidental Choice on Motor Learning. *Journal of Motor Learning and Development*, 5(2), 207–226. doi:10.1123/jmld.2016-0041.

Gröpel, P., & Mesagno, C. (2019). Choking interventions in sports: A systematic review. *International Review of Sport and Exercise Psychology*, 12(1), 176–201. doi:10.1080/1750984X.2017.1408134.

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. doi:10.1080/00222895.1995.9941714.

Hanin, Y., & Syrjä, P. (1995a). Performance affect in soccer players: An application of the IZOF model. *International Journal of Sports Medicine*, 16(4), 260–265. doi:10.1055/s-2007-973002.

Hanin, Y., & Syrjä, P. (1995b). Performance Affect in Junior Ice Hockey Players: An Application of the Individual Zones of Optimal Functioning Model. *The Sport Psychologist*, 9(2), 169–187. doi:10.1123/tsp.9.2.169.

Hardy, L., Mullen, R., & Jones, G. (1996). Knowledge and conscious control of motor actions under stress. *British Journal of Psychology*, 87, 621 – 636. doi: 10.1111/j.2044-8295.1996.tb02612.x. doi: 10.1123/jcsp.4.3.221.

Hill, D. M., Hanton, S., Matthews, N., & Fleming, S. (2010). A Qualitative Exploration of Choking in Elite Golf, Journal of Clinical Sport Psychology, 4(3), 221-240.

Huffman, J. L., Horslen, B. C., Carpenter, M. G., & Adkin, A. L. (2009). Does increased postural threat lead to more conscious control of posture? *Gait & Posture*, 30, 528 – 532. doi: 10.1016/j.gaitpost.2009.08.001.

Ioannidis, J. P. A. (2005). Why Most Published Research Findings Are False. *Plos Medicine*, 2(8), e124. doi:10.1371/journal.pmed.0020124.

Iso-Ahola, S. E., Dotson, C. O., Jagodinsky, A. E., Clark, L. C., Smallwood, L. L., Wilburn, C., Weimar, W. H., & Miller, M. W. (2016). Improving performance by anchoring movement and "nerves." *Human Movement Science*, 49, 239–247. doi:10.1016/j.humov.2016.07.008.

Jackson, R. C., Ashford, K. J., & Norsworthy, G. (2006). Attentional Focus, Dispositional Reinvestment, and Skilled Motor Performance under Pressure. *Journal of Sport and Exercise Psychology*, 28(1), 49–68. doi:10.1123/jsep.28.1.49.

Jalili, F., Hosseini, S., Jalili, F., & Salehian, M. H. (2011). Comparison of personality dimensions, mental toughness, and social skills of female students athletes (team-individual) and non-athletes. *Annals of Biological Research*, *2*(6), 554-560.

Janacsek, K., & Nemeth, D. (2013). Implicit sequence learning and working memory: Correlated or complicated? *Cortex*, 49(8), 2001–2006. doi:10.1016/j.cortex.2013.02.012.

Kegelaers, J., Hoogkamer, L., & Oudejans, R. R. (2022). Practice and performance management strategies of emerging professional musicians in preparation for orchestra auditions. *Research Studies in Music Education*, *44*(1), 175–191. doi: https://doi.org/10.1177/1321103X211054659.

Kal, E. C., Kamp, J. van der, Houdijk, H., Groet, E., Bennekom, C. A. M. van, & Scherder, E. J. A. (2015). Stay Focused! The Effects of Internal and External Focus of Attention on Movement Automaticity in Patients with Stroke. *Plos One*, 10(8), e0136917. doi:10.1371/journal.pone.0136917.

Kal, E., Prosée, R., Winters, M., & van der Kamp, J. (2018). Does implicit motor learning lead to greater automatization of motor skills compared to explicit motor learning? A systematic review. *Plos One*, 13(9), e0203591. doi:10.1371/journal.pone.0203591.

Kegelaers, J., & Oudejans, R. R. (2022). Pressure makes diamonds? A narrative review on the application of pressure training in high-performance sports. *International Journal of Sport and Exercise Psychology*, 1-19. Doi: 10.1080/1612197X.2022.2134436.

Kenny, D. (2021, April 5). Mediation. https://davidakenny.net/cm/mediate.htm.

Kinrade, N. P., Jackson, R. C., & Ashford, K. J. (2015). Reinvestment, task complexity and decision making under pressure in basketball. *Psychology of Sport & Exercise*, 20, 11 – 19. doi: 10.1016/j.psychsport.2015.03.007.

Klämpfl, M. K., Lobinger, B. H, & Raab, M. (2013).  How to detect the yips in golf. *Human Movement Science*, 32, 1270 – 1287. doi: 10.1016/j.humov.2013.04.004.

Kleynen, M., Braun, S. M., Bleijlevens, M. H., Lexis, M. A., Rasquin, S. M., Halfens, J., Wilson, M. R., Beurskens, A. J., & Masters, R. S. W. (2014). Using a Delphi Technique to Seek Consensus Regarding Definitions, Descriptions and Classification of Terms Related to Implicit and Explicit Forms of Motor Learning. *Plos One*, 9(6), e100227. doi:10.1371/journal.pone.0100227.

Kleynen, M., Braun, S. M., Rasquin, S. M. C., Bleijlevens, M. H. C., Lexis, M. A. S., Halfens, J., Wilson, M. R., Masters, R. S. W., & Beurskens, A. J. (2015). Multidisciplinary Views on Applying Explicit and Implicit Motor Learning in Practice: An International Survey. *Plos One*, 10(8), e0135522. doi:10.1371/journal.pone.0135522.

Koedijker, J. M., Oudejans, R. R. D., & Beek, P. J. (2007). Explicit rules and direction of attention in learning and performing the table tennis forehand. *International Journal of Sport Psychology*, 38(2), 227–244.

Koedijker, J. M., Oudejans, R. R. D., & Beek, P. J. (2008). Table tennis performance following explicit and analogy learning over 10,000 repetitions. *International Journal of Sport Psychology*, 39(3), 237–256.

Kreidler, S. M., Muller, K. E., Grunwald, G. K., Ringham, B. M., Coker-Dukowitz, Z. T., Sakhadeo, U. R., Barón, A. E., & Glueck, D. H. (2013). GLIMMPSE: Online Power Computation for Linear Models with and without a Baseline Covariate. Journal of Statistical Software, 54(10), i10.

Kuan, G., Morris, T., Kueh, Y. C., & Terry, P. C. (2018). Effects of relaxing and arousing music during imagery training on dart-throwing performance, physiological arousal indices, and competitive state anxiety. *Frontiers in Psychology*, 9. doi: 10.3389/fpsyg.2018.00014.

Laborde, S., Lautenbach, F., & Allen, M. S. (2015). The contribution of coping-related variables and heart rate variability to visual search performance under pressure. *Physiology & Behavior*, 139, 532–540. doi:10.1016/j.physbeh.2014.12.003.

Lajeunesse, M. J. (2016). Facilitating systematic reviews, data extraction and meta-analysis with the metagear package for r. *Methods in Ecology and Evolution*, 7(3), 323–330. doi:10.1111/2041-210X.12472.

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

Lakens, D., & Caldwell, A. R. (2021). Simulation-Based Power Analysis for Factorial Analysis of Variance Designs. *Advances in Methods and Practices in Psychological Science*, 4(1), 2515245920951503. doi:10.1177/2515245920951503.

Lam, W. K., Maxwell, J. P., & Masters, R. (2009). Analogy learning and the performance of motor skills under pressure. *Journal of Sport & Exercise Psychology*, 31(3), 337–357. doi:10.1123/jsep.31.3.337.

Lam, W. K., Maxwell, J. P., & Masters, R. S. W. (2009b). Analogy versus explicit learning of a modified basketball shooting task: Performance and kinematic outcomes. *Journal of Sport Sciences*, 27, 179 – 191. doi: 10.1080/02640410802448764.

Lawrence, G. P., Cassell, V. E., Beattie, S., Woodman, T., Khan, M. A., Hardy, L., & Gottwald, V. M. (2014). Practice with anxiety improves performance, but only when anxious: Evidence for the specificity of practice hypothesis. *Psychological Research*, 78(5), 634–650. doi:10.1007/s00426-013-0521-9.

Liao, C. M., & Masters, R. S. (2001). Analogy learning: A means to implicit motor learning. *Journal of Sports Sciences*, 19(5), 307–319. doi:10.1080/02640410152006081.

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. doi:10.1016/j.actpsy.2012.05.009.

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(2), 930–948. doi:10.1037/a0032817.

Lohse, K. R., Sherwood, D. E., & Healy, A. F. (2011). Neuromuscular effects of shifting the focus of attention in a simple force production task. *Journal of Motor Behavior*, 43(2), 173–184. doi:10.1080/00222895.2011.555436.

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. doi:10.1123/jmld.2015-0010.

Low, W. R., Sandercock, G. R. H., Freeman, P., Winter, M. E., Butt, J., & Maynard, I. (2021). Pressure training for performance domains: A meta-analysis. Sport, Exercise, and *Performance Psychology*, 10(1), 149–163. doi:10.1037/spy0000202.

Malhorta, N., Poolton, J. M., Wilson, M. R., Ngo, K., & Masters, R. S. W. (2012). Conscious monitoring and control (reinvestment) in surgical performance under pressure. *Surgical Endoscopy*, 26, 2423 – 2429. doi: 10.1007/s00464-012-2193-8.

Masters, R. S. W. (1992). Knowledge nerves and know-how: The role of explicit versus implicit knowledge in the breakdown of a complex motor skill under pressure. *British Journal of Psychology*, 83, 343 – 358. doi: 10.1111/j.2044-8295.1992.tb02446.x.

Masters, R. S. W., Eves, F. F., & Maxwell, J. P. (2005). Development of a movement specific reinvestment scale. Paper presented at the ISSP 11th *World Congress of Sport Psychology*, Sydney, Australia.

Masters, R. S. W., van Dujin, T., & Uiga, L. (2020). Advances in implicit motor learning. In N. J. Hodges & A. M. Williams (Eds.), *Skill Acquisition in Sport: Research, Theory, and Practice* (3rd ed.). Routledge.

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

Maxwell, J. P., Masters, R. S. W., & Eves, F. F. (2000). From novice to know-how: A longitudinal study of implicit motor learning. *Journal of Sports Sciences*, 18, 111 – 120. doi: 10.1080/026404100365180.

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

McKay, B., & Ste-Marie, D. M. (2020). Autonomy support and reduced feedback frequency have trivial effects on learning and performance of a golf putting task. *Human Movement Science*, 71, 102612. doi: 10.1016/j.humov.2020.102612.

McKay, B., Hussien, J., Carter, M., Yantha, Z., & Ste-Marie, D. (2021). Expecting to teach a novel golf putting task did not enhance retention performance: A replication attempt of Daou and colleagues. *Communications in Kinesiology*, 1. doi: 10.51224/cik.v1i2.39.

Mesagno, C., & Hill, D. M. (2013). Definition of choking in sport: Re-conceptualization and debate. *International Journal of Sport Psychology*, 44, 267–277. doi:10.7352/IJSP.2013.44.267.

Mesagno, C., Geukes, K., & Larkin, P. (2015). Choking under pressure: A review of current debates, literature, and interventions. In S. D. Mellalieu & S. Hanton (Eds.), *Contemporary Advances in Sport Psychology* (pp. 148–174). Routledge/Taylor & Francis Group.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, 6(7), e1000097. doi: 10.1371/journal.pmed.1000097.

Moore, L. J., Vine, S. J., Wilson, M. R., & Freeman, P. (2012). The effect of challenge and threat states on performance: An examination of potential mechanisms. *Psychophysiology*, 49(10), 1417–1425. doi:10.1111/j.1469-8986.2012.01449.x.

Mullen, R., Jones, E. S., Oliver, S., & Hardy, L. (2016). Anxiety and motor performance: More evidence for the effectiveness of holistic process goals as a solution to the process goal paradox. *Psychology of Sport and Exercise*, 27, 142 – 149. doi: 10.1016/j.psychsport.2016.08.009.

Nestojko, J. F., Bui, D. C., Kornell, N., Bjork, E. L. (2014). Expecting to teach enhances learning and organization of knowledge in free recall of text passages. *Memory and Cognition*, 42, 1038 – 1048. doi: 10.3758/s13421-014-0416-z.

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. doi:10.3758/BRM.40.3.892.

Oudejans, R. R. D., & Pijpers, J. R. (2009). Training with anxiety has a positive effect on expert perceptual-motor performance under pressure. *Quarterly Journal of Experimental Psychology*, 62(8), 1631–1647. doi:10.1080/17470210802557702.

Oudejans, R. R. D., & Pijpers, J. R. (2010). Training with mild anxiety may prevent choking under higher levels of anxiety. *Psychology of Sport and Exercise*, 11, 44 – 50. doi: 10.1016/j.psychsport.2009.05.002.

Oudejans, R. R. D., Kuijpers, W., Kooijman, C. C., & Bakker, F. C. (2011). Thoughts and attention of athletes under pressure: Skill-focus or performance worries? *Anxiety, Stress, and Coping*, 24(1), 59–73. doi:10.1080/10615806.2010.481331.

Owusu-Sekyere, F., & Gervis, M. (2016). In the pursuit of mental toughness: Is creating mentally tough players a disguise for emotional abuse? *International Journal of Coaching Science*, *10*(1), 3–23.

Poolton, J. M., Maxwell, J. P., Masters, R. S. W., & Raab, M. (2006). Benefits of an external focus of attention: Common coding or conscious processing? *Journal of Sports Sciences*, 24(1), 89–99. doi:10.1080/02640410500130854.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, 40(3), 879–891. doi:10.3758/BRM.40.3.879.

Renkl, A. (1995). Learning for later teaching: An exploration of mediational links between teaching expectancy and learning results. Learning and Instruction, 5, 21 – 36. doi: 10.1016/0959-4752(94)00015-H.

Richardson, J. T. E. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review*, 6, 135 – 147. doi: 10.1016/j.edurev.2010.12.001.

Roberts, L. J., Jackson, M. S., & Grundy, I. H. (2019). Choking under pressure: Illuminating the role of distraction and self-focus. *International Review of Sport and Exercise Psychology*, *12*(1), 49-69. doi: 10.1080/1750984X.2017.1374432.

Rosenbaum, D. A., Carlson, R. A., & Gilmore, R. O. (2001). Acquisition of intellectual and perceptual motor skills. *Annual Review of Psychology*, 52, 453 – 470. doi: 10.1146/annurev.psych.52.1.453.

Ross, S. M., & Di Vesta, F. J. (1976). Oral summary as a review strategy for enhancing recall of textual material. *Journal of Educational Psychology*, 68, 689 – 695. doi: 10.1037/0022-0663.68.6.689.

Rueden, C. T., Schindelin, J., Hiner, M. C., DeZonia, B. E., Walter, A. E., Arena, E. T., & Eliceiri, K. W. (2017). ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinformatics*, 18(1), 529. doi:10.1186/s12859-017-1934-z.

Ryan, R. M. (1982). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *Journal of Personality and Social Psychology*, 43, 450 – 461. doi: 10.1037/0022-3514.43.3.450.

Schmidt, R., & Lee, T. (2020). Motor Learning and Performance: From Principles to Application (6th ed.). Human Kinetics.

Schücker, L., Ebbing, L., & Hagemann, N. (2010). Learning by Analogies: Implications for Performance and Attentional Processes Under Pressure. *Human Movement*, 2(11), 191–199. doi:10.2478/v10038-010-0025-z.

Schücker, L., Hagemann, N., & Strauss, B. (2013). Analogy vs. Technical Learning in a Golf Putting Task: An Analysis of Performance Outcomes and Attentional Processes Under Pressure. *Human Movement*, 14(2), 175–184. doi:10.2478/humo-2013-0021.

Seery, M. D. (2013). The Biopsychosocial Model of Challenge and Threat: Using the Heart to Measure the Mind. *Social and Personality Psychology Compass*, 7(9), 637–653. doi:10.1111/spc3.12052.

Simonsohn, U. (2015). Small Telescopes: Detectability and the Evaluation of Replication Results. *Psychological Science*, 26(5), 559–569. doi:10.1177/0956797614567341.

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* (Clinical Research Ed.), 366, l4898. doi:10.1136/bmj.l4898.

Sterne, J. A. C., Sutton, A. J., Ioannidis, J. P. A., Terrin, N., Jones, D. R., Lau, J., Carpenter, J., Rücker, 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. doi:10.1136/bmj.d4002.

Sullivan, R., Uiga, L., Masters, R. S., Anson, G., & Nieuwenhuys, A. (2022). Conscious motor processing and the pressure-performance relationship: a systematic review. *International Review of Sport and Exercise Psychology*, 1-26. doi: 10.1080/1750984X.2022.2127327.

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.

Thompson, A. G. (2016). The Influence of Dispositional Mindfulness on State Anxiety and Motor Choking Under Pressure. https://etd.auburn.edu/handle/10415/5293.

Tomaka, J., Blascovich, J., Kibler, J., & Ernst, J. M. (1997). Cognitive and physiological antecedents of threat and challenge appraisal. *Journal of Personality and Social Psychology*, 73(1), 63–72. doi:10.1037//0022-3514.73.1.63.

Toner, J., & Moran, A. (2011). The effects of conscious processing on golf putting proficiency and kinematics. *Journal of Sports Sciences, 29*(7), 673 – 683. https://doi.org/10.1080/02640414.2011.553964.

Tse, A. C. Y., Wong, T. W. L., & Masters, R. S. W. (2017). Examining motor learning in older adults using analogy instruction. *Psychology of Sport and Exercise*, 28, 78 – 84. Doi: 10.1016/j.psychsport.2016.10.005.

van der Wel, P., van Steenbergen, H. Pupil dilation as an index of effort in cognitive control tasks: A review. *Psychon Bull Rev* 25, 2005–2015 (2018). doi: 10.3758/s13423-018-1432-y.

Verner, M., Conzelmann, A., Lehnert, K., Seiler, R., Wassmer, A., & Rammsayer, T. (2010). Subjective stress in female elite athletes and non-athletes: Evidence from cortisol analyses. *Europe's Journal of Psychology*, 6(4), 56–70. doi:10.5964/ejop.v6i4.223.

Viechtbauer, W. (2010). Conducting Meta-Analyses in R with the metafor Package. *Journal of Statistical Software*, 36(1), 1–48. doi:10.18637/jss.v036.i03.

Vine, S. J., & Wilson, M. R. (2010). Quiet Eye Training: Effects on Learning and Performance Under Pressure. *Journal of Applied Sport Psychology*, 22(4), 361–376. doi:10.1080/10413200.2010.495106.

Vine, S. J., Moore, L. J., & Wilson, M. R. (2016). An Integrative Framework of Stress, Attention, and Visuomotor Performance. *Frontiers in Psychology*, 7. doi:10.3389/fpsyg.2016.01671.

Vine, S. J., Moore, L. J., Cooke, A., Ring, C., & Wilson, M. R. (2013). Quiet eye training: A means to implicit motor learning. International Journal of Sport Psychology, 44(4), 367–386. doi: 10.7352/ijsp 2013.00.000.

Walter, N., Nikoleizig, L., & Alfermann, D. (2019). Effects of self-talk training on competitive anxiety, self-efficacy, volitional skills, and performance: An intervention study with junior sub-elite athletes. *Sports*, 7, 1 - 20. doi: 10.3390/sports7060148.

Williams, A. M., Fawver, B., & Hodges, N. J. (2017). Using the" Expert Performance Approach" as a Framework for Improving Understanding of Expert Learning. *Frontline Learning Research*, *5*(3), 64-79. https://journals.sfu.ca/flr/index.php/journal/article/view/267.

Wilson, M., Chattington, M., Marple-Horvat, D. E., & Smith, N. C. (2007). A Comparison of Self-Focus versus Attentional Explanations of Choking. *Journal of Sport and Exercise Psychology*, 29(4), 439–456. doi:10.1123/jsep.29.4.439.

Wulf, G. (2013). Attentional focus and motor learning: a review of 15 years. *International Review of sport and Exercise psychology*, *6*(1), 77-104. doi: 10.1080/1750984X.2012.723728.

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. doi:10.3758/s13423-015-0999-9.

Wulf, G., & Su, J. (2007). An External Focus of Attention Enhances Golf Shot Accuracy in Beginners and Experts. *Research Quarterly for Exercise and Sport*, 78(4), 384–389. doi:10.1080/02701367.2007.10599436.

Zhu, F. F., Poolton, J. M., Wilson, M. R., Maxwell, J. P., & Masters, R. S. W. (2011). Neural coactivation as a yardstick of implicit motor learning and the propensity for conscious control of movement. *Biological Psychology*, 87, 66 – 73. doi: 10.1016/j.biopsycho.2011.02.004.

Zijlstra, F. R. H. (1995). Efficiency in work behaviour: A design approach for modern tools. Delft University Press. https://repository.tudelft.nl/islandora/object/uuid:d97a028b-c3dc-4930-b2ab-a7877993a17f?collection=research.

Zhu, Y., Wang, Q. & Zhang, L. Study of EEG characteristics while solving scientific problems with different mental effort. *Sci Rep* **11**, 23783 (2021). doi: 10.1038/s41598-021-03321-9.