Educational Psychology Review (2024) 36:57
https://doi.org/10.1007/s10648-024-09892-z

META-ANALYSIS

The Effects of Contextual Interference Learning on the Acquisition and Relatively Permanent Gains in Skilled Performance: A Critical Systematic Review with Multilevel Meta-Analysis

Achraf Ammar[1,2,3,4] · Khaled Trabelsi[4,5] · Mohamed Ali Boujelbane[1,4] · Atef Salem[1,4] · Omar Boukhris[6,7] · Jordan M. Glenn[8] · Piotr Zmijewski[9] · Haitham A. Jahrami[10,11] · Hamdi Chtourou[4,12] · Wolfgang I. Schöllhorn[1]

Accepted: 6 May 2024 / Published online: 4 June 2024

Abstract

The paradoxical effects of contextual interference (CI) assume that high CI practices hinder performances during the acquisition phase of learning, while providing more permanent enhancement during the retention phase. This meta-analysis evaluates the possible generalizability of the CI phenomenon in physical education (PE) and sports contexts, with regard to the acute and relatively permanent gains in performance outcomes. A total of 933 records from five electronic databases were screened using the PICOS criteria, of which 36 studies were selected. Outcomes evaluating the performance changes (Δ) from pre-post, post-retention, and pre-retention tests were included. Out of 183 overall pooled outcomes, Δ in only 37 performance outcomes (20%) agreed with the paradoxical CI effects on the acquisition or the relatively permanent gains. No statistically significant overall difference was detected for "Δ pre-post" between low (blocked) (28.9 ± 59.5%) and high (random/serial) (27.9 ± 52.8%) CI (effect size (ES) = 0.1, $p = 0.35$). An overall significant difference ($p = 0.001$) in favor of high CI practice was detected in "Δ post-retention." However, this difference was not large enough (ES = −0.35) to produce an overall greater long-term gain following high (24.56 ± 4.4%), compared to low (21.9 ± 9.8%) CI (ES = −0.13, $p = 0.18$). Out of 10 tested variables, only the age significantly moderated both CI effects ($p < 0.0001$ for both Δ pre-post and Δ pre-retention) and the female proportion significantly moderated only the first CI effect ($p = 0.009$ for Δ pre-post). These findings found very limited evidence supporting the recommendation to employ high CI practices to gain a longer-term performance advantage, calling into question the generalization of the CI model

Achraf Ammar and Khaled Trabelsi contributed equally as first authors.

Extended author information available on the last page of the article

to PE and sports practices. High-quality follow-up research evaluating alternative motor-learning models are therefore needed.

Keywords Physical education · Motor learning · Retention · Practice schedule · Gross motor skills · Variability of practice · Differential learning

Introduction

In the domain of motor learning, the contextual interference (CI) approach and its efficacy in maximizing skill acquisition and retention have generated discussion among coaches, physical education (PE) teachers, practitioners, and sports scientists (Wang et al., 2022). Originally, the CI learning effects were described as a combination of two phenomena occurring one after the other. When practice trials are interrupted by the execution of other tasks, detrimental effects can be observed immediately after the acquisition phase but are followed by greater benefits in the subsequent retention phase compared to uninterrupted sequence (Magill & Hall, 1990). Due to the analogous structure of the experimental design in teaching institutions, where testing is often conducted immediately after the teaching phase, it is tempting to directly transfer the findings to an educational context. However, the problem resides in the alluring attempt to generalize experimental findings from the original field of verbal (Battig, 1966, 1972) and fine motor learning (Shea & Morgan, 1979) to the parallel learning of several gross motor movements through epistemologically critical induction. This issue of induction (first proposed in 1740), identified by Hume (2007) as the inference from a part (specific examples) to the whole (general conclusions), is reflected in four essential steps in the research timeline of the CI effect. The first induction's problematic step occurred when the learning of letter sequences (Battig, 1979) was extended to the learning of individual fine motor movements (Shea & Morgan, 1979). The second arose in the extension to the parallel learning of several fine motor movement sequences (Wright et al., 1992). The third involved generalizing to the learning of gross motor movements (Wrisberg & Liu, 1991), and the fourth involved generalizing to the parallel learning of several large motor movements (Apidogo et al., 2024). Initially, only the single and undifferentiated criterion of learning served as the basis for a possible generalization (Battig, 1979), but it has since become apparent that the CI effect, similar to the working memory model developed in parallel (Baddeley & Hitch, 1974), could only be systematically demonstrated for sequences of fine motor movements with a high visual-spatial component (Schöllhorn et al., 2022). The experimental study by Mechsner et al. (2001) suggests a similar potential regulatory role of visual perception on the motor system. Nonetheless, according to Battig (1979), the interference effect is in general considered as a consequence of adaptation processes that occurs when a learner reacts to a varied input during an acquisition phase.

The original and most common forms of structured practice in studies on the CI effect are blocked (often referred to as "low CI practice") and random (often referred to as "high CI practice") practice. Blocked practice requires the learner to complete

all trials of a task before moving on to a different task in predictable patterns, thereby allowing them to repeat multiple tasks with less variability (Schmidt et al., 2011). Whereas random practice interleaves tasks in an unpredictable order with tasks rarely repeated in consecutive trials, producing high variability (Schmidt et al., 2011). High CI practice has been assumed to involve learners in (i) deeper and more elaborate information processing, known as the "elaborative hypothesis" suggested by Shea and Zimny (1983), or (ii) increased forgetting, which requires extra retrieval practice and reconstruction of tasks' plans-of-action, known as the "forgetting/reconstruction hypothesis" (Lee & Magil 1983, 1985). According to both hypotheses, more detailed and permanent representations of the task should be memorized during the random practice, which may enhance the retention performance. Based on the cognitive load theory (Gentile, 1998), overloading the working memory during the random practice is suggested to hinder performances during the acquisition phase.

Following a review of the literature, Brady (1998) speculated that the effects of CI are generalizable to motor abilities. However, Shewokis (1997), Barreiros et al. (2007), and Wang et al. (2022), based on numerous contradictory data, noted the difficulty of generalizing the CI phenomenon to all motor abilities and particularly to applied sports skills as the effects appeared to be mediated by various variables such as the learners' characteristics, the nature of the skills and intervention durations. The CI literature includes studies with participants of different ages, genders, and experience levels, with tasks of different types and complexity, along with learning durations ranging from 1 day to several weeks. Although the majority of the CI studies employed a standardized experimental design including an acquisition phase followed by a retention phase with pre-, post-, and retention tests, inconsistent findings on the CI effects have been reported (Barreiros et al., 2007; Brady, 2004; Wang et al., 2022). Despite the enormous effort but epistemologically fatal attempt to reach a decision on a possible generalization by democratically weighing the respective number of positive studies, the issue remains debatable. While many studies have reported superior acquisition performance for the blocked group and superior retention performance for the random group (Hajihosseini, 2016; Kalkhoran & Shariati, 2015; Memmert, 2006; Pasand et al., 2016), others have found no significant differences (Naimo et al., 2013; Porter et al., 2020; Sharp et al., 2020; Yanci et al., 2013; Zetou et al., 2007) or even reversed values at post-acquisition and/or retention tests (Farrow & Maschette, 1997; Hall et al., 1994; Zipp & Gentile., 2010).

Although meta-analyses follow a similar epistemological logic of weighted majorities, they also provide a quantitative summary overview of the results of previous studies, on the basis of which research directions could be rethought. Since the CI proponents' claim is to improve motor learning in general with a random acquisition schedule, it is important to remember that motor learning was defined by Schmidt and Lee (2019) as "a set of processes associated with practice or experience leading to relatively permanent gains in the capability for skilled performance." According to this definition, a proper evaluation of the effect of a motor learning approach is primarily related to the amount of performance gain obtained during acquisition (acute gain) and/or retention (sustainable gain) phase. However, the

majority of CI studies (original research or reviews) have attempted to validate selected CI effects by comparing the performance of blocked and random groups separately for post and retention tests (Ammar et al., 2023; Wang et al., 2022). In this context, a recent systematic review and meta-analysis with meta-regression showed little evidence to support the CI hypothesis and its generalization to applied PE and sport settings (Ammar et al., 2023). This meta-analysis pooled only the performance outcomes of blocked and random groups at each test session (post- and retention tests), while little attention was paid to the possible magnitude of gain (e.g., Δ pre-post, Δ post-retention, and Δ pre-retention). The lack of gain analyses calls into question the majority of previous conclusions derived from original and review studies evaluating CI effects.

To address this problem and provide a more differentiated and holistic evaluation of the CI effect, the current systematic review and meta-analysis focused on the difference between high and low CI schedules in terms of acute (Δ pre-post) and delayed (Δ post-retention and Δ pre-retention) gains in skilled performance. Likewise, by performing a thorough subgroup analysis and/or meta-regression of possible moderators, the present systematic review tends to offer more reliable explanations of previously reported discrepancies, while also highlighting useful perspectives that may aid future investigations to optimize gains in skilled performances.

Method

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 statement (Page et al., 2021a, b) and the recent PRISMA guidelines adapted to sports science (Rico-González et al., 2022) were followed in the present systematic review. Since no participants were required for this study, obtaining ethical approval was not necessary.

Eligibility Criteria

Inclusion criteria were applied following the PICO (participants, intervention, comparators, and outcomes) model. Articles were included if they were conducted on healthy participants and examining the effect of CI on motor learning with an intervention group of random/serial practice order (high CI (Lage et al., 2022)) and at least a control group with blocked practice order (low CI). No specific criteria were specified for the age of the studied population. Any outcome evaluating the acquisition, retention, and/or transfer of a learned sports/gross motor skill was included. Gross motor skills are defined as the ability to engage the use of large muscle groups (e.g., in the torso, arms, legs), coordinating major body movements to perform activities such as maintaining balance, walking, running, jumping, and throwing objects (Esposito & Vivanti, 2013). These eligibility criteria resulted in the exclusion of studies that were conducted in a human population with physical or mental disability (e.g., mentally retarded, Down syndrome, mental handicap) and/

or conducted with lack of practice for the control (blocked) group. Studies focusing only on cognitive performance and/or fine motor skills and did not describe the effects of CI on at least one measure/outcome of performance in a movement task (sport/gross motor skill) were also excluded. Furthermore, narrative and systematic reviews, books, citations, conference proceedings, commentaries, and studies, of which only the abstracts were available, were excluded to be able to make a reliable risk-of-bias assessment. Importantly, studies failing to guarantee the intervention and control groups are equivalent at baseline were excluded from the quantitative synthesis (meta-analysis).

Search Strategy

Up to April 16th, 2022, a comprehensive systematic search of studies was performed electronically in five electronic databases—PubMed, Web of Science, Scopus, Taylor and Francis, and SciELO. To ensure a thorough retrieval of relevant studies, our search strategy included no filters, such as language, species, or publication year. Furthermore, for PubMed and Web of Science, we used the "All Fields" search option, while for Taylor and Francis, we selected "Anywhere" (equivalent to "All Fields"). In Scopus, primarily an abstract and indexing database that does not include full texts, we utilized "TITLE-ABS-KEY" which is the default search field. Search terms (including mapping to appropriate MeSH terms where appropriate) described low and high contextual interference classes (contextual interference OR random practice OR alternating practice OR non-repeated practice OR variable practice OR high interference OR blocked practice OR repeated practice OR low interference), in combination with keywords relating to sports and motor skills (sports performance OR sports skill OR motor skill OR motor task OR motor learning OR motor skill learning OR motor skill acquisition OR retention performance OR transfer performance OR skill transfer OR skill retention). The search was applied to titles and abstracts and was conducted with additional filters excluding non-human studies (NOT mice OR rats OR animals) and studies in injury/disease (NOT mice OR rats OR animals OR injury OR diseases OR impairment OR rehabilitation OR surgery OR stroke OR Parkinson OR patient OR treatment OR disorder OR review OR meta-analysis OR systematic review OR commentary). Full search strategies for each database are presented in Appendix S1. To minimize the risk of missing relevant publications, an additional search was performed by screening the reference lists of the included studies and related review papers using Google Scholar. Personal files were also searched and subject experts in the field were contacted for information on potential upcoming studies. Authors of included studies were contacted for missing information.

Selection Process

The search strategies were combined, and duplicates were removed by "Endnote 20" (Camelot UK Bidco Limited-Clarivate, UK, London). Following the removal of duplicates, titles and abstracts of all unique hits were screened by two authors (AA

and MAB) for eligibility, and disagreements were resolved by consensus. Once all relevant articles were located, the full texts were screened for their appropriateness for inclusion in accordance with the predetermined PICO criteria. The reason for excluding an article during each step (abstract and full-text screenings) was recorded.

Data Collection Process

To avoid any selection bias and data extraction flaws, two authors independently collected data using a pilot-tested extraction form and resolved disagreements by consensus. Extracted data included (i) author(s) and year of publication, (ii) study design, (iii) sample size (intervention and control groups), (iv) participant characteristics (i.e., age, sex, and experience level), (v) sport and specific skill names, (vi) and intervention protocol including the intervention duration, number of training sessions, total number of acquisition trials, duration between test sessions (pre-post; post-retention, and/or post-transfer), and testing protocol (e.g., block and/or random), (v) and the difference between groups at pre-, post-, retention-, and/or transfer-test.

The mean and standard deviation (SD) of the main performance outcomes (sports/gross motor skills) at the different testing sessions (pre, post, retention, and/or transfer) were extracted and used to perform the meta-analysis and meta-regression. When data were displayed graphically and no further data were provided upon request, mean and SD were retrieved using WebPlot Digitizer (Pacifica, CA, USA, Version: 4.5) (https://automeris.io/WebPlotDigitizer, accessed on 8 June 2022). Studies not including or mentioning the results of pre-test measurements, as well as the studies showing significant differences between the low (control) and high (intervention) CI groups at pre-test (i.e., baseline), were excluded from the meta-analysis and meta-regression. The same applies to studies that did not mention SD (or other measures which could be converted to SD) with the mean values.

In studies that focused solely on serial and blocked (low CI) groups without including random practice (high CI), data from serial practice were categorized under the broader high CI group for comparison against low CI.

During data extraction, assessments conducted before the intervention or training begins, or referred to as "baseline," were considered "pre-test"; assessments; those occurring immediately after the intervention or training program were considered "post-test"; measures conducted after the completion of the post-test, which can range from the same day (Boyce & Del Rey, 1990) to a year later (Memmert, 2006), were considered "retention." For studies that included multiple retention tests, such as at 24 h and 1 week (e.g., Boyce & Del Rey, 1990), data were extracted separately for each specified retention period. The same approach was applied to studies with multiple acquisition tests.

Quality Assessment

Given the absence of a specific scale for evaluating methodological quality in motor learning research, the Physiotherapy Evidence Database (PEDro) scale (Verhagen et al., 1998) has been employed in the present study. Originally designed for evaluating physiotherapeutic interventions on clinical populations, the PEDro scale is universally recognized for its efficacy in appraising study quality within the broader spectrum of clinical research, including motor control studies (Ammar et al., 2020a, b, c; Maher et al., 2003). The PEDro scale's focus on key research design elements includes randomization, blinding, and statistical analysis, serving as universal benchmarks of methodological rigor in clinical research.

The PEDro scale is a reliable and objective tool based on the Delphi list developed by Verhagen et al. (1998). To identify which of the trials are likely to be externally (criteria 1) and internally (criteria 2–9) valid with interpretable results, each paper was independently assessed twice by two independent authors using the 11-item checklist to yield a maximum score of 10 (the sum of awarded points for criteria 2–11). Discrepant results were resolved through consensus meetings. Points were only awarded when a criterion was clearly satisfied. According to a recent article evaluating the measurement properties and summarizing the clinimetric properties of the PEDro scale, total PEDro scores of 0–3 are considered "poor," 4–5 "fair," 6–8 "good," and 9–10 "excellent" (Cashin & McAuley, 2020). Additionally, for trials evaluating complex interventions (e.g., exercise), a total PEDro score of 8/10 is considered optimal.

Meta-Analysis and Risk of Bias Assessment

The meta-analysis was conducted following the recent recommendations of Hagger (2022) and Gucciardi et al. (2022) utilizing the R programming language and software environment (R Version 4.3.1; R Core Team, 2021).

Calculation of Effect Size and Variance

Standardized mean difference (SMD) was used as the measure of effect size (ES) and was calculated between the low and high CI for the acute (Δ pre-post) and delayed (Δ post-retention and Δ pre-retention) gains in skilled performance. For each performance outcome, SMD calculation was based on the initially extracted data of mean and SD at each test point, along with the sample size, and the trial correlation. Because no correlation values were published in any of the included investigations, the recommendation of Higgins and Green (2011), assuming a conservative estimation of $r = 0.50$, was followed. To calculate the SMD for each performance outcome regarding acquisition gain (Δ pre-post), the initial step involves computing the mean difference (MD) and SD change between pre- and post-test for each group (low and high CI) using the following formulas (Deeks et al., 2019):

$$\mathrm{MD} = \mathrm{Post} - \mathrm{Prescore}$$

$$\text{SDChange} = \text{Sqrt}(\text{SDPre}^2 + \text{SDPost}^2 - 2 \times 0.5 \times \text{SDPre} \times \text{SDPost})$$

Then, the raw difference (rawDiff) and the pooled SD (SDp) between group were calculated as follows:

$$\text{Rawdiff} = \text{MDLowCI} - \text{MDHighCI}$$

$$\text{SDp} = \text{Sqrt}(((\text{n(LowCI)} - 1) \times \text{SDChange(LowCI)}^2 + (\text{n(HighCI)} - 1) \times \text{SDChange(HighCI)}^2)/(\text{n(LowCI)} + \text{n(HighCI)} - 2))$$

The SMD was then computed as the rawdiff divided by the SDp. The same formulas were applied to determine the SMD for the delayed (Δ post-retention and Δ pre-retention) gains, with the performance in the subsequent test (i.e., retention test) considered as post-test score.

To account for the direction of change, a corrected SMD (cSMD) was calculated as an SMD absolute value bearing a "+" or "-" sign based on the direction of mean change. A positive SMD value indicates that performance changes were in favor of low CI, whereas a negative SMD indicates that performance changes were in favor of high CI. This correction step is crucial in the motor control field, and was performed with caution as the interpretation of effect direction for performance outcome depends on the specific measurements of the performance. Performance measured in terms of "number of errors" or "completion time" are interpreted inversely to performance measured in terms of "successful throws" or "covered distance." For instance, for outcomes measured in terms of errors, a higher Δ change was deemed non-favorable for skill acquisition; thus, a higher Δ change in high CI practice compared to low CI practice was interpreted as positive, indicating a preference for low CI practice. This approach was applied for both the acquisition (pre-post) and the relatively permanent gains (pre-retention). For the Δ post-retention, a smaller decrease in performance (smaller Δ change) could be an indicator of better retention. Consequently, a smaller post-retention decrease in performance was deemed favorable for low CI practice when outcomes were measured in terms of errors, thus interpreting the effect direction as negative, indicating a preference for high CI practice.

The SMD magnitude was interpreted as follows: small (0.20–0.49), moderate (0.50–0.79), and large (>0.80) (Hedges, 1981). Forest plots illustrated point estimates of the effect size and 95% confidence intervals.

Variance in each performance outcome was computed as square of the standard error (SE), which has been determined using the following formula:

$$\text{SE} = \text{Sqrt}(1/\text{n(LowCI)} + 1/\text{n(HighCI)} + \text{SMD}^2/(2 \times (\text{n(LowCI)} + \text{n(HighCI)})))$$

Statistical Synthesis of Effect Sizes

Most included studies in the meta-analysis (94.6%) reported data for multiple performance outcomes when examining the effects of low vs. high CI. It is suggested that effect sizes (Ess) derived from the same study are likely more similar than ESs from different studies (Assink & Wibbelink, 2016). Therefore, the incorporation of multiple ESs from a single study may violate the assumption of independence in ESs in meta-analyses (e.g., Cheung, 2014, Van den Noortgate et al., 2013). To account the dependency issues when including multiple data points from the same studies in the present meta-analysis, a three-level meta-analysis (i.e., a multilevel model) was used (Cheung, 2019). This approach separates the variance components of the pooled effect into sampling variance of the observed effect sizes (level 1), and variance within (level 2) and between studies (level 3). Previous reports have shown that such a multilevel meta-analysis effectively addresses the hierarchical nature of the data (e.g., effect sizes nested within studies), thus preserving information and improving statistical power (Assink & Wibbelink, 2016). A multilevel meta-analysis was conducted separately for each Δ change (Δ pre-post, Δ post-retention, and Δ pre-retention). Forest plots illustrated point estimates of the weighted ESs and 95% CIs.

These analyses were conducted in R language and environment for statistical computing (Version 4.3.1; R Core Team, 2021) using the R-package metafor (Version 4.2.0; Viechtbauer & Cheung, 2010).

Statistical Heterogeneity and Moderators Analysis

Two statistics, Q (Morris, 2008) and I^2 (Higgins et al., 2003), were assessed statistical heterogeneity. Substantial heterogeneity was considered for I^2 values > 50%, indicating a that random-effects model was preferred to a fixed-effects model (Higgins et al., 2003). To identify potential sources of variance and heterogeneity, while controlling for several dependency issues, moderator analysis was performed using multivariate subgroup analysis for categorical variables including the age groups (5–9, 10–14, 15–19, 20–24, 25–32, and > 65 years old), the experience level of the participants (i.e., inexperienced, and experienced), the investigated sports and skills, and the testing protocol (i.e., only blocked, only random, or blocked for the control group/random for the intervention group). Additionally, multivariate meta-regression for integer or decimal variables, including the sex of the participants (i.e., expressed in % of females' subjects in any given study), intervention duration (days), number of training sessions, total number of acquisition trials, and duration between test sessions (pre-post, post-retention, and pre-retention for the correspondent Δ changes) was also performed. The codes used to run these analyses are available in the supplementary file.

Outliers Detection

In our meta-analytic models, we conducted diagnostics for leverage, outliers, and influential cases using hat values, Cook's distance, and studentized residuals, following methodologies from Stevens (1984), Viechtbauer and Cheung (2010), and Aguinis et al. (2013). Cases were flagged if their hat values and Cook's distance exceeded three times their mean, with studentized residuals greater than 3 in absolute terms.

Publication Bias

The risk of small study bias was visualized through funnel plots. Moreover, Begg's rank correlation test and the multilevel model of the Egger's test were used to detect publication bias. The variance of effects was converted into standard error and the moderation model was performed (Fernández-Castilla et al., 2021). The robust estimated results of the Egger's test were produced for all analyses.

A significance level of $p < 0.05$ was adopted for all analyses. z-curve has also been conducted to estimate the power of the included studies. According to Bartoš and Schimmack, (2022) and McKay et al. (2023), a z-curve analysis estimates the statistical power (i.e., average power) of the included studies, on the basis of the significant results that are present. The expected discovery rate estimated by the z-curve analysis and reflecting the expected rate of significant results is equivalent to the power estimate, while the observed discovery rate is equivalent to the actual rate of significant results. A discrepancy between the observed discovery rate and the

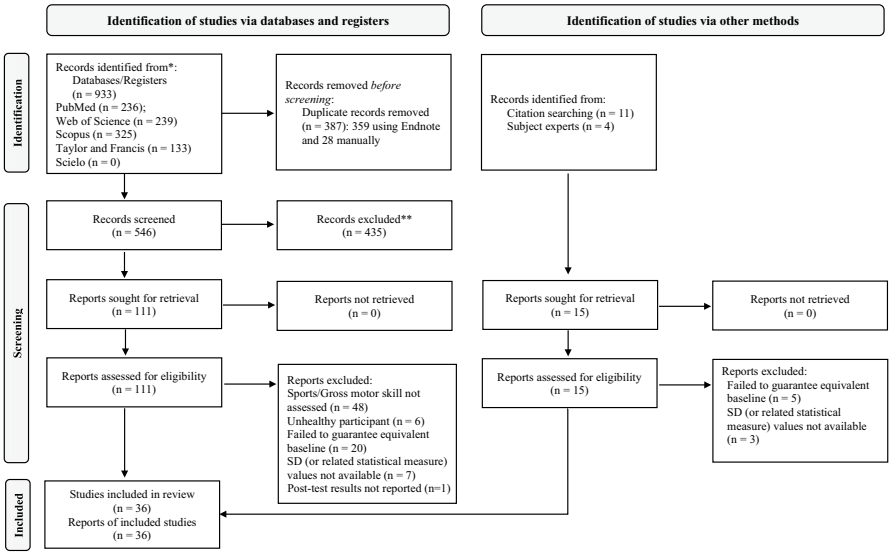

Fig. 1 PRISMA 2020 flow diagram

95% confidence interval of the expected discovery rate estimate provides evidence of reporting bias.

Results

Study Selection

Figure 1 provides the flow diagram of the search process. The predefined search strategies yielded a preliminary pool of 933 possible papers. Among these, 387 were excluded as duplicates (359 using Endnote and 28 manually). The remaining 546 published papers were screened by titles and abstracts for eligibility with 435 being removed and 111 reports remained for full-text assessment. By applying the PICO eligibility criteria, 54 full texts were excluded based on the population (i.e., 6 studies focused on mentally disabled participants) and outcome (i.e., 48 only focus on cognitive performance and/or fine motor skills with no focus on at least one outcome of performance in a movement task (sport/gross motor skill)). In total, 57 published studies met the PICO inclusion criteria. After conducting an additional search, 15 articles were identified as potentially relevant studies, resulting in a total of 72 relevant studies. After a careful review of the 72 full texts, 36 studies were excluded from the quantitative synthesis (meta-analysis) for the following reasons: 25 studies failed to guarantee the intervention and control groups were equivalent at baseline (15 studies without pre-test measurement, 9 studies did not mention the pre-test results, and one study failed to recruit groups with equivalent pre-test performances); 10 studies did not report SD values or other measures that could be converted to SD such as standard error (SE), confidence intervals (CI), variance, coefficient of variation (CV) in some or all measurements; and 1 study did not report post-test results. Detailed references of the excluded full texts are presented in Appendix S2. In total, 36 articles with a wide variety of motor tasks and enough comparable data were included in the systematic review (Table 1). However, only 35 of these were incorporated into the quantitative synthesis due to the study by Pasand et al. (2016) being identified as an outlier in all tested Δ changes.

Study Characteristics

Characteristics of the 36 studies as well as the difference between the low and high CI group in the Δ pre-post, Δ post-retention, and Δ pre-retention of each performance outcome are presented in Table 1. Most of the studies (26 out of 36) applied a randomized between-subject design. Six studies followed a non-randomized between-subject design with group assignment based on pre-test results, and the four remaining studies applied either a quasi-random or matched between-subject design, a randomized or a counterbalancing within-subject design. Regarding the examined sports-related skills, 6 studies focused on volleyball, 4 studies on golf, 3 studies on tennis, 2 studies on either badminton, baseball, basketball, field hockey, table tennis, or shooting (e.g., pistol or kneeling), 1 study on either football, Australian rules football, gymnastics, dancing or bench press, and the remaining 6 studies

Table 1 Study characteristics

N	Study	Discipline	Skills/techniques	Measurements	mean age Y	*N* participants		Female %	Level of experience	Duration between test session	Study design	Intervention duration	Number of acquisition sessions	Total acquisition trials number	Testing protocol	Block vs. Random/serial		
						Block Grp	Rndm/serial Grp									Diff in Δ pre-post	Diff in Δ post-retention	Diff in Δ pre-retention
1	Boyce and Del Rey (1990)	Kneeling shooting	Kneeling shooting task	Acquisition 1	≈ 16	30	30	41,67	Inexperienced	4 days	Randomised between subject design	14 days	4 sessions	20 trials	random	BlK > Rdm		
				Acquisition 2						7 days						BlK > Rdm		
				Acquisition 3						10 days						BlK ≈ Rdm		
				Acquisition 4						14 days						BlK ≈ Rdm		
				Retention 1						same day					block/random		Rdm > BlK	BlK ≈ Rdm
				Retention 2						7 days							Rdm > BlK	BlK ≈ Rdm
2	French et al. (1990)	Volleyball skills	Pass	Post-test	≈ 16	12	12	54,67	Inexperienced	9 days	Randomised between subject design	9 days	9 sessions	90 trials	block	Rdm > BlK		
			Set													BlK ≈ Rdm		
			Overhead serve													BlK ≈ Rdm		
			Pass	Retention						2 days							BlK ≈ Rdm	BlK ≈ Rdm

Table 1 (continued)

N	Study	Discipline	Skills/techniques		Measurements	mean age Y	*N* participants		Female %	Level of experience	Duration between test session	Study design	Intervention duration	Number of acquisition sessions	Total acquisition trials number	Testing protocol	Block vs. Random/serial		
							Block Grp	Rndm/serial Grp									Diff in Δ pre-post	Diff in Δ post-retention	Diff in Δ pre-retention
			Set															BlK ≈ Rdm	BlK ≈ Rdm
			Overhead serve															BlK ≈ Rdm	BlK ≈ Rdm
3	Wrisberg and Liu (1991)	Badminton	Long serves	Female	Acquisition	≈ 20	26	26	38,5	Inexperienced	14 days	Counterbalancing within subjects design	14 days	5 sessions	45 trials	random	BlK ≈ Rdm		
				Male													Rdm > BlK		
			Short serves	Female													BlK > Rdm		
				Male													BlK > Rdm		
			Long serves	Female	Retention						same day							Rdm > BlK	BlK ≈ Rdm
				Male														BlK ≈ Rdm	BlK ≈ Rdm
			Short serves	Female														Rdm > BlK	BlK ≈ Rdm
				Male														Rdm > BlK	Rdm > BlK

Table 1 (continued)

N	Study	Discipline	Skills/techniques		Measurements	mean age Y	*N* participants		Female %	Level of experience	Duration between test session	Study design	Intervention duration	Number of acquisition sessions	Total acquisition trials number	Testing protocol	Block vs. Random/serial		
							Block Grp	Rndm/serial Grp									Diff in Δ pre-post	Diff in Δ post-retention	Diff in Δ pre-retention
4	Hall et al. (1994)	Baseball skills	Combined (Fastballs, Curveballs, Changeup)	Random testing	Acquisition	≈ 19	10	10	0	Experienced	6 weeks	Randomised between subject design	6 weeks	12 sessions	NM	random	Rdm > BlK		
				Block testing												block	Rdm > BlK		
6	Goodwin and Meeuwsen (1996)	Golf	Putting 2.43m		Acquisition	≈ 26	10	10	100	Inexperienced	2 days	Randomised between subject design	2 days	2 sessions	66 trials	block	BlK ≈ Rdm		
			Putting 5.47m														BlK ≈ Rdm		
			Putting 3.95m														BlK ≈ Rdm		
			Putting 2.43m		Retention						1 day							BlK ≈ Rdm	BlK ≈ Rdm
			Putting 5.47m															BlK ≈ Rdm	BlK ≈ Rdm
			Putting 3.95m															Rdm > BlK	BlK ≈ Rdm

Table 1 (continued)

N	Study	Discipline	Skills/techniques		Measurements	mean age Y	*N* participants		Female %	Level of experience	Duration between test session	Study design	Intervention duration	Number of acquisition sessions	Total acquisition trials number	Testing protocol	Block vs. Random/serial		
							Block Grp	Rndm/serial Grp									Diff in Δ pre-post	Diff in Δ post-retention	Diff in Δ pre-retention
7	Hebert et al. (1996)	Tennis table	Forehand and backhand strockes	low skills	Acquisition (block testing)	≈ 22	46 (22 low skills + 24 high skills)	37 (19 low skills + 18 high skills)	45,78	low skills (Inexperienced) and high skills (experienced)	9 days	Randomised between subject design	9 days	9sessions	135 trials	block	BlK > Rdm		
				high skills													BlK ≈ Rdm		
				low skills	Acquisition (randm testing)											random	BlK > Rdm		
				high skills													BlK ≈ Rdm		
8	Pollatou et al. (1997)	Motor skills	Throwing task		Acquisition	≈ 20	20	21	52,38	Inexperienced	2 weeks	Randomised between subject design	2weeks	8sessions	80 trials	block/random	BlK ≈ Rdm		
			Kicking task														BlK ≈ Rdm		
			Throwing task		Retention						1week							BlK ≈ Rdm	BlK ≈ Rdm
			Kicking task															BlK ≈ Rdm	BlK ≈ Rdm

Table 1 (continued)

N	Study	Discipline	Skills/techniques		Measurements	mean age Y	*N* participants		Female %	Level of experience	Duration between test session	Study design	Intervention duration	Number of acquisition sessions	Total acquisition trials number	Testing protocol	Block vs. Random/serial		
							Block Grp	Rndm/serial Grp									Diff in Δ pre-post	Diff in Δ post-retention	Diff in Δ pre-retention
9	Bortoli et al. (2001)	Motor Skill	Underhand throwing	Nonvariable schedules	Acquisition	≈ 9	12	12	50	Experienced in throwing and hurdle running and not experienced in complex jumping	8 weeks	Randomised between subject design	8 weeks	14 sessions	126 trials	block	BlK ≈ Rdm		
			Quintuple jumping														BlK ≈ Rdm		
			Hurdle running														BlK ≈ Rdm		
			Underhand throwing	Variable schedules													BlK ≈ Rdm		
			Quintuple jumping														BlK ≈ Rdm		
			Hurdle running														BlK ≈ Rdm		

Table 1 (continued)

N	Study	Discipline	Skills/techniques	Measurements	mean age Y	*N* participants		Female %	Level of experience	Duration between test session	Study design	Intervention duration	Number of acquisition sessions	Total acquisition trials number	Testing protocol	Block vs. Random/serial		
						Block Grp	Rndm/serial Grp									Diff in Δ pre-post	Diff in Δ post-retention	Diff in Δ pre-retention
10	Meira and Tani (2003)	Volleyball	Underhand serves	Acquisition	≈ 13	18	18	100	Inexperienced	8 days	not randomised (groups assignment based on pretest) between subject design	8 days	8 sessions	288 trials	block/random	BlK ≈ Rdm		
			Overhand serves													BlK ≈ Rdm		
11	Smith et al. (2003)	Gymnastics	Cartwheel	Acquisition	≈ 20	16	16	50	Inexperienced	same day	Randomised between subject design	1 day	1 session	192 trials	block/random	BlK ≈ Rdm		
				Retention						7days					block		BlK ≈ Rdm	BlK ≈ Rdm

Table 1 (continued)

N	Study	Discipline	Skills/techniques		Measurements	mean age Y	*N* participants		Female %	Level of experience	Duration between test session	Study design	Intervention duration	Number of acquisition sessions	Total acquisition trials number	Testing protocol	Block vs. Random/serial		
							Block Grp	Rndm/serial Grp									Diff in Δ pre-post	Diff in Δ post-retention	Diff in Δ pre-retention
12	Vera and Montilla (2003)	Throwing	Vertical Target	Tennis ball 3m	Acquisition	≈ 6	34	37	50	Inexperienced	6 weeks	not randomised (groups assignment based on pretest) between subject design	6 weeks	18 sessions	1044 trials	block	BlK ≈ Rdm		
				Feather-fly ball 3m													BlK ≈ Rdm		
				Tennis ball 5m													BlK ≈ Rdm		
				Feather flyball 5m													Rdm > BlK		
			Horizontal Target	Tennis ball 3m													BlK ≈ Rdm		
				Feather fl ball 3m													Rdm > BlK		

Table 1 (continued)

N	Study	Discipline	Skills/techniques	Measurements	mean age Y	*N* participants		Female %	Level of experience	Duration between test session	Study design	Intervention duration	Number of acquisition sessions	Total acquisition trials number	Testing protocol	Block vs. Random/serial		
						Block Grp	Rndm/serial Grp									Diff in Δ pre-post	Diff in Δ post-retention	Diff in Δ pre-retention
			Tennis ball 5m													BlK ≈ Rdm		
			Feather flyball 5m													BlK ≈ Rdm		
			Retention							2 weeks							BlK ≈ Rdm	Rdm > BlK
13	Keller et al. (2006)	Pistol-shooting	Basic pistol-shooting skills	Acquisition	≈ 32	6	6	50	Inexperienced	same day	Randomised between subject design	1 day	1 session	30 trials	block	BlK ≈ Rdm		
				Retention													BlK ≈ Rdm	BlK ≈ Rdm
14	Memmert (2006)	Basketball	Shooting performance	Acquisition	≈ 25	16	16	50	Inexperienced	same day	quasi-random between subject design	1 day	1 session	160 trials	block	BlK > Rdm		
				Retention						1 year							Rdm > BlK	BlK ≈ Rdm
15	Jones and French (2007)	3 Volley Ball skills	Underhand Serve	Acquisition	≈ 16	17	17	90	Inexperienced	9 days	Randomised between subject design	9 days	9 sessions	90 trials	block/random	BlK ≈ Rdm		

Table 1 (continued)

N	Study	Discipline	Skills/techniques	Measurements	mean age Y	*N* participants		Female %	Level of experience	Duration between test session	Study design	Intervention duration	Number of acquisition sessions	Total acquisition trials number	Testing protocol	Block vs. Random/serial		
						Block Grp	Rndm/serial Grp									Diff in Δ pre-post	Diff in Δ post-retention	Diff in Δ pre-retention
			Overhead Set													BlK ≈ Rdm		
			Underhand Serve	Retention						2 days							BlK ≈ Rdm	BlK ≈ Rdm
			Overhead Set														BlK ≈ Rdm	BlK ≈ Rdm
16	Zetou et al. (2007)	Volleyball	Set	Acquisition	≈ 12	13	13	100	Inexperienced	10 weeks	Randomised between subject design	10 weeks	20 sessions	1800 trials	block	BlK ≈ Rdm		
			Pass													BlK ≈ Rdm		
			Service													BlK ≈ Rdm		
			Set	Retention						2 weeks							BlK ≈ Rdm	BlK ≈ Rdm
			Pass														BlK ≈ Rdm	BlK ≈ Rdm
			Service														BlK ≈ Rdm	BlK ≈ Rdm
17	Vera et al. (2008)	Football	Dribling	Acquisition	≈ 10	22	23	49	Inexperienced	5 weeks	Randomised between subject design	5 weeks	15 sessions	NM	block	BlK ≈ Rdm		

Table 1 (continued)

N	Study	Discipline	Skills/techniques	Measurements	mean age Y	*N* participants		Female %	Level of experience	Duration between test session	Study design	Intervention duration	Number of acquisition sessions	Total acquisition trials number	Testing protocol	Block vs. Random/serial		
						Block Grp	Rndm/serial Grp									Diff in Δ pre-post	Diff in Δ post-retention	Diff in Δ pre-retention
			Kicking													BlK ≈ Rdm		
			Dribling	Retention						2 weeks							BlK ≈ Rdm	BlK ≈ Rdm
			Kicking														BlK ≈ Rdm	BlK ≈ Rdm
18	Bertollo et al. (2010)	Dancing	Rhythmic footstep sequence	Acquisition	≈ 16	20	20	100	Inexperienced	3 weeks	Randomised between subject	3 weeks	6 sessions	NM	block	BlK > Rdm		
				Retention													BlK ≈ Rdm	BlK ≈ Rdm
19	Menayo et al. (2010)	Tennis	Flat forehand shot	Acquisition	≈ 23	8	8	0	Experienced	3 weeks	Randomised between subject design	3 weeks	12 sessions	576 trials	random	BlK > Rdm		
			Topspin backhand shot													BlK ≈ Rdm		
			Slice backhand volley shot													BlK ≈ Rdm		
			Slice serve shot													BlK ≈ Rdm		

Table 1 (continued)

N	Study	Discipline	Skills/techniques	Measurements	mean age Y	*N* participants		Female %	Level of experience	Duration between test session	Study design	Intervention duration	Number of acquisition sessions	Total acquisition trials number	Testing protocol	Block vs. Random/serial		
						Block Grp	Rndm/serial Grp									Diff in Δ pre-post	Diff in Δ post-retention	Diff in Δ pre-retention
			Flat forehand shot	Retention 1						2 days							BlK ≈ Rdm	BlK ≈ Rdm
				Retention 2						14 days							BlK ≈ Rdm	BlK ≈ Rdm
				Retention 3						28 days							Rdm > BlK	BlK ≈ Rdm
				Retention 4						42 days							Rdm > BlK	Rdm > BlK
			Topspin backhand shot	Retention 1						2 days							Rdm > BlK	Rdm > BlK
				Retention 2						14 days							BlK > Rdm	BlK > Rdm
				Retention 3						28 days							BlK > Rdm	BlK > Rdm
				Retention 4						42 days							Rdm > BlK	Rdm > BlK
			Slice backhand volley shot	Retention 1						2 days							Rdm > BlK	Rdm > BlK
				Retention 2						14 days							Rdm > BlK	Rdm > BlK

Table 1 (continued)

N	Study	Discipline	Skills/techniques	Measurements	mean age Y	*N* participants		Female %	Level of experience	Duration between test session	Study design	Intervention duration	Number of acquisition sessions	Total acquisition trials number	Testing protocol	Block vs. Random/serial		
						Block Grp	Rndm/serial Grp									Diff in Δ pre-post	Diff in Δ post-retention	Diff in Δ pre-retention
				Retention 3						28 days							BlK > Rdm	BlK ≈ Rdm
				Retention 4						42 days							BlK > Rdm	BlK > Rdm
			Slice serve shot	Retention 1						2 days							BlK ≈ Rdm	BlK ≈ Rdm
				Retention 2						14 days							Rdm > BlK	Rdm > BlK
				Retention 3						28 days							BlK ≈ Rdm	BlK ≈ Rdm
				Retention 4						42 days							BlK > Rdm	BlK > Rdm
20	Travlos (2010)	Volleyball	Underhand volleyball serve	Acquisition	≈ 14	12	12	50	Inexperienced	3 days	Randomised between subject	3 days	3 sessions	45 trials	block	BlK ≈ Rdm		
21	Rendell et al. (2011)	Australian Rules Football	Pass (Hand-Ball)	Acquisition	≈ 29	9	10	42	Inexperienced	4 weeks	Matched between subject design	4 weeks	7 sessions	320 trials	block	BlK ≈ Rdm		
			Kicking													Rdm > BlK		
			Pass (Hand-Ball)	Retention						5 weeks							Rdm > BlK	BlK ≈ Rdm

Table 1 (continued)

N	Study	Discipline	Skills/techniques	Measurements	mean age Y	*N* participants		Female %	Level of experience	Duration between test session	Study design	Intervention duration	Number of acquisition sessions	Total acquisition trials number	Testing protocol	Block vs. Random/serial		
						Block Grp	Rndm/serial Grp									Diff in Δ pre-post	Diff in Δ post-retention	Diff in Δ pre-retention
			Kicking														Rdm > BlK	Rdm > BlK
22	Cheong et al. (2012)	Field hockey	Indian dribble	Control	Acquisition 1	≈ 18	11	10	46	Inexperienced	14 days	Randomised between subject design	10 days	3 sessions	45 trials	block	BlK ≈ Rdm	
				Mvt form													Rdm > BlK	
			Push pass	Accuracy													BlK ≈ Rdm	
				Speed													BlK > Rdm	
				Movt form													BlK ≈ Rdm	
			Hitting	Accuracy													BlK ≈ Rdm	
				Speed													BlK ≈ Rdm	
				Movt form													BlK ≈ Rdm	
			Indian dribble	Control	Acquisition 2					28 days		21 days	6 sessions	90 trials		BlK ≈ Rdm		
				Mvt form													Rdm > BlK	

Table 1 (continued)

N	Study	Discipline	Skills/techniques		Measurements	mean age Y	*N* participants		Female %	Level of experience	Duration between test session	Study design	Intervention duration	Number of acquisition sessions	Total acquisition trials number	Testing protocol	Block vs. Random/serial		
							Block Grp	Rndm/serial Grp									Diff in Δ pre-post	Diff in Δ post-retention	Diff in Δ pre-retention
			Push pass	Accuracy													BlK ≈ Rdm		
				Speed													BlK > Rdm		
				Movt form															
			Hitting	Accuracy													BlK ≈ Rdm		
				Speed													BlK ≈ Rdm		
				Movt form													BlK ≈ Rdm		
			Indian dribble	Control	Retention						7 days							BlK ≈ Rdm	BlK ≈ Rdm
				Movement form														BlK ≈ Rdm	Rdm > BlK
			Push pass	Accuracy														BlK ≈ Rdm	BlK ≈ Rdm
				Speed														BlK ≈ Rdm	BlK > Rdm
				Movt form														BlK ≈ Rdm	BlK ≈ Rdm

Table 1 (continued)

N	Study	Discipline	Skills/techniques		Measurements	mean age Y	*N* participants		Female %	Level of experience	Duration between test session	Study design	Intervention duration	Number of acquisition sessions	Total acquisition trials number	Testing protocol	Block vs. Random/serial		
							Block Grp	Rndm/serial Grp									Diff in Δ pre-post	Diff in Δ post-retention	Diff in Δ pre-retention
			Hitting	Accuracy														BlK ≈ Rdm	BlK ≈ Rdm
				Speed														BlK ≈ Rdm	BlK ≈ Rdm
				Movt form														BlK ≈ Rdm	BlK ≈ Rdm
23	Naimo et al. (2013)	Bench press	1RM		Acquisition	≈ 21	8	8	38	Inexperienced	4 weeks	not randomised between subject design	4 weeks	12 sessions	528 trials	block	BlK ≈ Rdm		
			1 RM gain (%)														BlK > Rdm		
			component checklist scores														Rdm > BlK		
			1RM		Retention						9 days							BlK ≈ Rdm	BlK ≈ Rdm
			1 RM gain (%)															Rdm > BlK	BlK ≈ Rdm
			component checklist scores															BlK > Rdm	BlK ≈ Rdm

Table 1 (continued)

N	Study	Discipline	Skills/techniques	Measurements	mean age Y	*N* participants		Female %	Level of experience	Duration between test session	Study design	Intervention duration	Number of acquisition sessions	Total acquisition trials number	Testing protocol	Block vs. Random/serial		
						Block Grp	Rndm/serial Grp									Diff in Δ pre-post	Diff in Δ post-retention	Diff in Δ pre-retention
24	Sadri et al. (2013)	Badminton	Serves/ Stroke	Acquisition	≈ 11	15	15	100	Inexperienced	NM	Randomised between subject design	NM	10 sessions	150 trials	block	Blk > Rdm		
				Retention 1						same day							Rdm > Blk	Blk > Rdm
				Retention 2						2 days							Rdm > Blk	Blk ≈ Rdm
25	Yanci et al. (2013)	Motor Skill	Change of direction	Acquisition	≈ 6	15	16	54	Inexperienced	3 weeks	Randomised between subject design	3 weeks	6 sessions	NM	block	Blk ≈ Rdm		
			Sprinting at 5 m (s)													Blk ≈ Rdm		
			Sprinting at 15 m (s)													Blk ≈ Rdm		
26	Krause et al. (2014)	Golf	Putting task (Novice/ Experts)	Distance	Acquisition	≈ 26	20	20	27	Inexperienced	2 days	Randomised between subject design	2 days	2 sessions	120 trials	block	Rdm > Blk		
				Angle												Blk ≈ Rdm		
				Distance					Experienced							Blk ≈ Rdm		

Table 1 (continued)

N	Study	Discipline	Skills/techniques	Measurements	mean age Y	*N* participants		Female %	Level of experience	Duration between test session	Study design	Intervention duration	Number of acquisition sessions	Total acquisition trials number	Testing protocol	Block vs. Random/serial		
						Block Grp	Rndm/serial Grp									Diff in Δ pre-post	Diff in Δ post-retention	Diff in Δ pre-retention
				Angle												BlK ≈ Rdm		
				Distance	Retention				Inexperienced	1 day							BlK ≈ Rdm	BlK ≈ Rdm
				Angle													BlK ≈ Rdm	BlK ≈ Rdm
				Distance					Experienced								BlK ≈ Rdm	BlK ≈ Rdm
				Angle													BlK ≈ Rdm	Rdm > BlK
27	Broadbent et al. (2015)	Tennis	Anticipatoty tennis shots (laboratory, accuracy)	Acquisition	≈ 13	9	9	NM	Experienced	3 weeks	not randomised between subject	3 weeks	3 sessions	216 trials	block/random	BlK ≈ Rdm		
				Retention						7 days							Rdm > BlK	Rdm > BlK
28	Yanci et al. (2015)	Motor Skill	Agility (MAT)	Acquisition	≈ 6	19	19	42	Inexperienced	4 weeks	Randomised between subject	4 weeks	8 sessions	24 trials	block	BlK ≈ Rdm		

Table 1 (continued)

N	Study	Discipline	Skills/techniques	Measurements	mean age Y	*N* participants		Female %	Level of experience	Duration between test session	Study design	Intervention duration	Number of acquisition sessions	Total acquisition trials number	Testing protocol	Block vs. Random/serial		
						Block Grp	Rndm/serial Grp									Diff in Δ pre-post	Diff in Δ post-retention	Diff in Δ pre-retention
29	Cheong et al. (2016)	Field hockey	Close dribble	Acquisition (close skills)	≈ 22	24	20	31	Inexperienced	3 weeks	Not randomised between subject design	3 weeks	6 sessions	NM	block/random	BlK ≈ Rdm		
			Push pass acc													BlK ≈ Rdm		
			Push pass sp													BlK ≈ Rdm		
			Close dribble	Retention 1						1 week							BlK ≈ Rdm	BlK ≈ Rdm
				Retention 2						3 weeks							BlK ≈ Rdm	BlK ≈ Rdm
			Push pass accuracy	Retention 1						1 week							Rdm > BlK	BlK ≈ Rdm
				Retention 2						3 weeks							Rdm > BlK	BlK ≈ Rdm
			Push pass speed	Retention 1						1 week							BlK ≈ Rdm	BlK ≈ Rdm
				Retention 2						3 weeks							BlK ≈ Rdm	BlK > Rdm
30	Pasand et al. (2016)	Volleyball	Combined (Forearm pass, set, service)	Acquisition	≈ 23	15	15	0	Inexperienced	3 weeks	Randomised between subject design	3 weeks	9 sessions	1215 trials	block/random	BlK > Rdm		

Table 1 (continued)

N	Study	Discipline	Skills/techniques	Measurements	mean age Y	*N* participants		Female %	Level of experience	Duration between test session	Study design	Intervention duration	Number of acquisition sessions	Total acquisition trials number	Testing protocol	Block vs. Random/serial		
						Block Grp	Rndm/serial Grp									Diff in Δ pre-post	Diff in Δ post-retention	Diff in Δ pre-retention
				Retention						2 days							Rdm > BlK	Rdm > BlK
31	Buszard et al. (2017)	Tennis	Serves-in (%)	Acquisition	≈ 12	8	8	50	Experienced	9 weeks	Randomised between subject design	7 weeks	10 sessions	400 trials	block	BlK > Rdm		
			Service displacement													BlK > Rdm		
			Service velocity													BlK ≈ Rdm		
32	Fazeli et al. (2017)	Golf	Golf putting	Acquisition	≈ 27	10	10	0	Inexperienced	6 days	Randomised between subject design	6 days	6 sessions	1080 trials	block/random	BlK ≈ Rdm		
				Retention						7 days							Rdm > BlK	Rdm > BlK
33	Aiken and Genter (2018)	Golf	Chipping tasks	Acquisition	≈ 20	12	12	58	Inexperienced	same day	Randomised between subject design	1 day	1 session	54 trials	block/random	BlK > Rdm		
				Retention (block)											block		BlK ≈ Rdm	BlK > Rdm
				Retention (random)											random		Rdm > BlK	Rdm > BlK

Table 1 (continued)

N	Study	Discipline	Skills/techniques		Measurements	mean age Y	*N* participants		Female %	Level of experience	Duration between test session	Study design	Intervention duration	Number of acquisition sessions	Total acquisition trials number	Testing protocol	Block vs. Random/serial		
							Block Grp	Rndm/serial Grp									Diff in Δ pre-post	Diff in Δ post-retention	Diff in Δ pre-retention
34	North et al. (2019)	Table tennis	Backhand shot		Acquisition	≈ 28	10	10	40	Inexperienced	same day	Not randomised (assignment based on pretest)	1 day	1 session	150 trails	block/random	Rdm > BlK		
					Retention						1 day							BlK ≈ Rdm	BlK ≈ Rdm
35	Porter et al. (2020)	Basket ball	Shot		Acquisition	≈ 25	12	12	50	Inexperienced	11 days	Randomised between subject design	11 days	6 sessions	360 trials	block/random	BlK ≈ Rdm		
					Retention 1						same day					random		BlK ≈ Rdm	BlK ≈ Rdm
					Retention 2						7 days					random		BlK ≈ Rdm	BlK ≈ Rdm
36	Sharp et al. (2020)	Baseball	Solid hits/swing (%)	Dominant	Acquisition	≈ 16	5	6	0	Experienced	4 weeks	Randomised between subject design	4 weeks	8 sessions	90 trials	block	BlK ≈ Rdm		
				N-dominant													BlK ≈ Rdm		

Table 1 (continued)

N	Study	Discipline	Skills/techniques		Measurements	mean age Y	*N* participants		Female %	Level of experience	Duration between test session	Study design	Intervention duration	Number of acquisition sessions	Total acquisition trials number	Testing protocol	Block vs. Random/serial		
							Block Grp	Rndm/serial Grp									Diff in Δ pre-post	Diff in Δ post-retention	Diff in Δ pre-retention
			Solid hits/swing (ball accuracy)	Dominant													BlK ≈ Rdm		
				N-dominant													BlK ≈ Rdm		
37	Jeon et al. (2021)	Wii Fit Dynamic Balance	Timed Up and go test (TUG)		Acquisition	≈ 65	22	19	NM	Inexperienced	9 days	randomized, pretest-posttest design	10 days	10 sessions	60 trials	block	BlK > Rdm		
			Mobility-test (POMA)														BlK > Rdm		
			Functional-reach test (FRT)														BlK > Rdm		
			TUG	Ret 1	Retention						3 days							BlK ≈ Rdm	BlK > Rdm
				Ret 2							9 days							Rdm > BlK	BlK > Rdm
			POMA	Ret 1							3 days							Rdm > BlK	BlK ≈ Rdm
				Ret 2							9 days							BlK > Rdm	BlK > Rdm

Table 1 (continued)

N	Study	Discipline	Skills/techniques		Measurements	mean age Y	*N* participants		Female %	Level of experience	Duration between test session	Study design	Intervention duration	Number of acquisition sessions	Total acquisition trials number	Testing protocol	Block vs. Random/serial		
							Block Grp	Rndm/serial Grp									Diff in Δ pre-post	Diff in Δ post-retention	Diff in Δ pre-retention
			FRT	Ret 1							3 days							BlK ≈ Rdm	BlK > Rdm
				Ret 2							9 days							Rdm > BlK	BlK ≈ Rdm

Rdm > *BlK*, superior performance following random compared to blocked practice; *BlK* > *Rdm*, superior performance following blocked compared to random practice; *BlK* ≈ *Rdm*, no significant difference between performances measured following random vs. blocked practice; *block/random*, block testing for block-practice group and random testing for random-practice group

focused on some sport-related gross motoric skills (e.g., balance, agility, throwing, jumping), without sports specification.

Of the 36 studies included, 22 investigated more than one performance outcome (i.e., different sports/gross motor skills and/or different performance variables). Similarly, most of the studies (27 out of 36) examined both the acquisition and retention phases, while nine studies focused on the acquisition phase with no retention test. Additionally, some studies focused on more than one population (e.g., male vs. female or experienced vs. inexperienced), or testing protocol (e.g., blocked vs. random testing). Therefore, a specific ES was calculated for each outcome, population, and/or test session, and the 36 selected studies yielded a total of 264 ES. These ES are distributed between Δ pre-post (103 ES), Δ post-retention (81 ES), and Δ pre-retention (81 ES).

It should be noted that data of forearm pass from Jones and French (2007) were not considered in the meta-analysis, as these studies failed to guarantee equivalence at baselines for these specific skills. Similarly, for the meta-regression assessing the effect of between-test duration, we elected to exclude the retention data of Memmert (2006), which used extreme values for the duration between tests (365 days) compared to the other studies (56 days maximum), and may potentially bias the linear trend.

Furthermore, four studies (Bortoli et al., 2001; Buszard et al., 2017; Keller et al., 2006; Krause et al., 2014) focused on serial practice, but not on random practice. Therefore, data from the corresponding serial practice groups were extracted and analyzed, falling within the broader category of high CI.

Subject Characteristics

The studies in this systematic review included a total of 1135 participants between low ($n = 572$) and high ($n = 563$) CI groups. The number of participants in each trial ranged from 11 to 83 for both groups. These studies targeted healthy populations with a mean age of 19.5 ± 10.2 years. Four studies targeted 5- to 9-year-old participants, seven studies targeted the "10 to 14" years old range, seven studies targeted the "15 to 19" years old range, nine studies targeted the "20 to 24" years old range, and 7 studies targeted the "25 to 29" years old range. One study focused on adults aged 32 years, and one study focused on senior adults (aged above 65 years). In 24 out of 36 studies, participants from both sexes were recruited. Five studies recruited only male participants and five studies recruited only female participants. The gender distributions of the participants were not mentioned in one study. Most of the studies (28 out of 36) examined an inexperienced population (including participants with no previous experience, novice, beginner, unskilled, or with limited experience), while five studies examined experienced athletes, and three studies examined populations with varied experience levels (e.g., experienced and inexperienced groups).

Differences in the Δ of Performance Outcomes Between Low and High CI Practices

As Table 1 indicates, in the acquisition phase, the pre-post Δ was in favor of low CI training in 22 performance outcomes, in favor of high CI training in 12 performance outcomes, with no differences between the groups in 68 performance outcomes. In the retention phase, the post-retention Δ was in favor of low CI training in 7 performance outcomes, in favor of high CI training in 28 performance outcomes, with no differences between the groups in 46 performance outcomes. Combining the acquisition and retention phase, the pre-retention Δ favored low CI training in 12 performance outcomes and high CI in 15 performance outcomes, with no differences between groups in 54 performance outcomes.

Differences Between Low and High CI Practices in the Δ Pre-Post

Data from 36 studies investigating the difference between low and high CI practice in the Δ pre-post of performance outcomes were pooled in our MA.

Outlier analysis identified two performance outcomes from studies by Pasand et al. (2016) and Buszard et al. (2017) as potential outliers due to their extreme effect sizes (|SMD| range between 4.8 and 7.8). Following outlier removal, the analysis proceeded. Given that the study by Pasand et al. (2016) investigated only one performance outcome, identified as outlier, this study was completely excluded from the multilevel meta-analysis.

Accordingly, 35 included studies yielded a total of 101 ES. A small effect size was computed (Fig. 2), based on the three-level meta-analytic model, indicating a non-significant difference between low and high CI practice in the Δ pre-post of performance outcomes (ES = 0.1, standard error (SE) = 0.11, 95% CI = −0.12 to 0.32, t(100) = 0.92, $p = 0.359$). Significant heterogeneity was computed ($Q = 336.84$, df = 100, $p < 0.0001$). The estimated variance components were $\tau^2_{\text{Level 3}} = 0.29$ and $\tau^2_{\text{Level 2}} = 0.11$. This means that $I^2_{\text{Level 3}} = 54.3\%$ of the total variation can be attributed to between-cluster, and $I^2_{\text{Level 2}} = 20.44\%$ to within-cluster heterogeneity, with a total $I^2 = 74.75\%$. To identify potential sources of heterogeneity, subgroup analysis and meta-regression analysis were performed.

Moderator Analyses (of Multilevel Models)

Multilevel *s*ubgroup analysis revealed "age groups" significantly moderated the difference between low and high CI (F(5, 95) = 8.83, $p < 0.00014$). However, "experience level," "sports," "skills," and "testing protocol" did not moderate the difference between low and high CI practice in the Δ pre-post performance (Table 2). No significant differences between low and high CI practices in Δ pre-post performance were detected for the majority of analyzed subgroups in each category (Table 2). Better Δ pre-post performance in favor of low CI practice was apparent for the 20–24 (ES = 0.33 and $p = 0.009$) and > 65 years old age groups (ES = 1.65 and $p = 0.016$) (Table 2), as well as for "Kneeling/pistol shooting" sports (ES = 0.51,

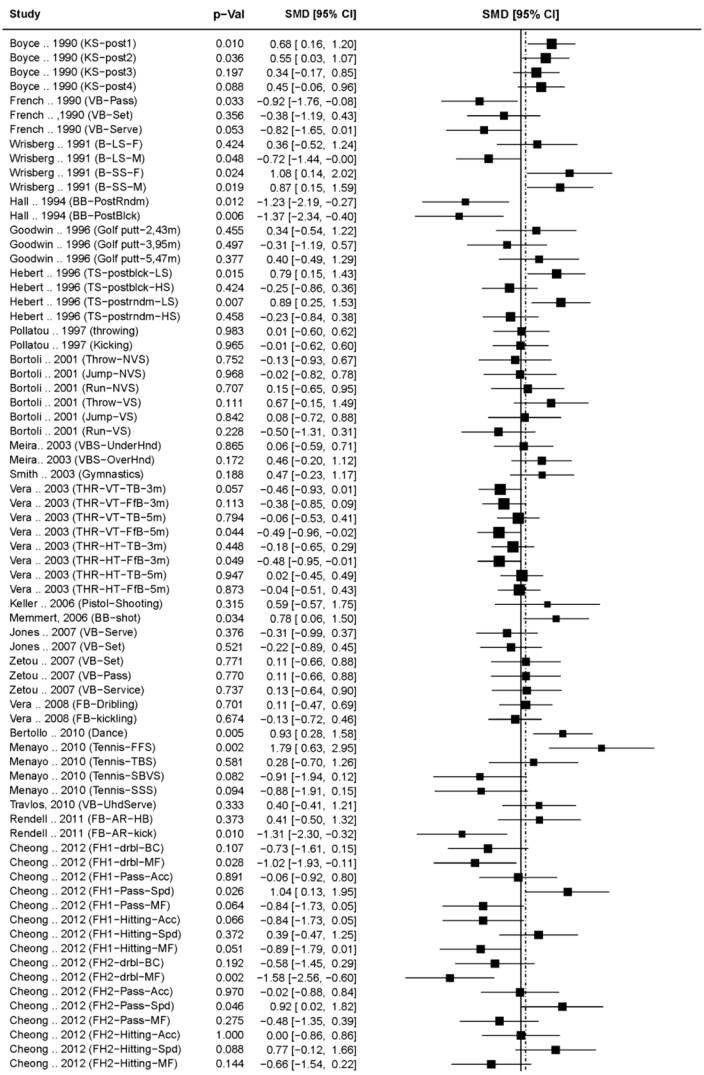

Fig. 2 Forest plot for the weighted difference between blocked and random practices at acquisition gain. *Legends:* KS, kneeling shooting; B, badminton; LS, long serves; SS, short serves; M, male; F, female; TS, tennis strockes; LS, low skills; HS, high skills; VBS, volleyball serve; THR, throwing; VT, vertical target; HT, horizontal target; Tscore, total score; TB, tennis ball; FfB, feather fly ball; VB, volleyball; FB, football; FFS, flat forehand shot; TBS, topspin backhand shot; SBVS, slice backhand volley shot; SSS, slice serve shot; HB, handball; AR, Australian Rules; TUG, timed up and go test; POMA, Tinetti performance-oriented mobility assessment; FRT, functional reach test; FH, field hockey; PassAcc, push pass accuracy; PassSpd, push pass speed; NVS, nonvariable schedules; VS, variable schedules; drbl, dribble; BC, ball control; MF, movement form; FH1, field hockey acquisition 1; FH2, field hockey acquisition 2; Srv, serve; dis, displacement; Bpress, bench press; Uhd, underhand; CODA, change of direction ability; SSP, straight sprinting performance; BSB, baseball; DS, dominant side; NDS, non dominant side; Put, putting; dist, distance; ang, angle; NOV, novice; Exprt, experts

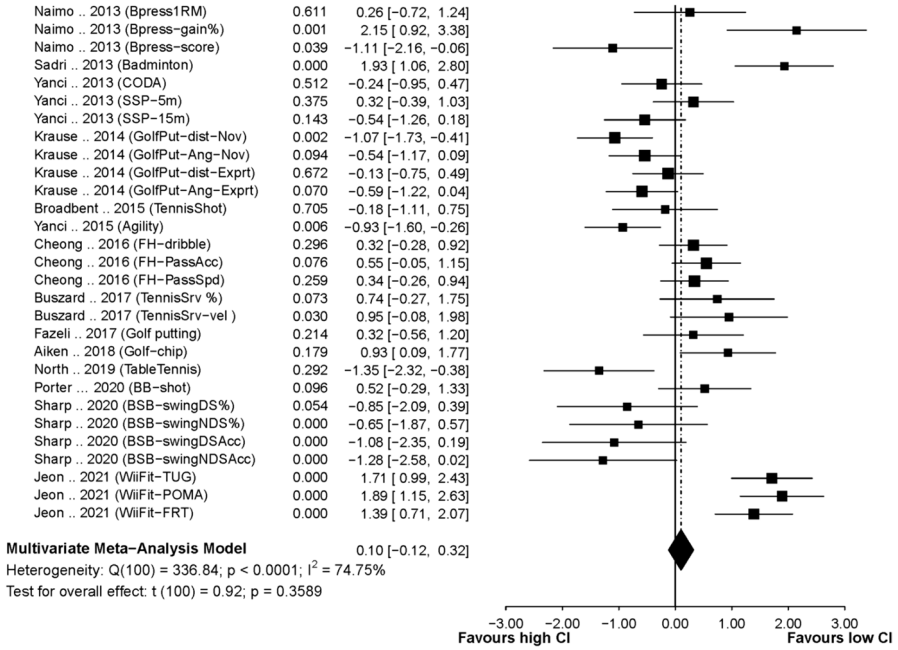

Fig. 2 (continued)

$p = 0.017$) (Table 2), However, this significant difference was in favor of high CI practice for the "Baseball (swing)" sport (ES = −1.11, $p = 0.005$) (Table 2).

Multivariate meta-regression analyses revealed a moderating effect of only female percentage ($F(1, 98) = 6.99$, $p = 0.0095$), with coefficient = 0.0084 (SE = 0.003, 95% CI = 0.002 to 0.015, $t(98) = 2.64$, and $p = 0.0095$, Fig. 3). However, intervention duration, number of training sessions, total number of acquisition trials, and duration between pre- and post-test sessions did not moderate the difference between low and high CI practice in the Δ pre-post performance (Table 3).

Publication Bias

Visual inspection of the funnel plot (Fig. 4) revealed no evidence of publication bias, a conclusion confirmed by Begg and Mazumdar's rank correlation test (Kendall's tau = −0.08, $p = 0.266$)) and the multilevel model of Egger's test ($F(1,99) = 0.09$, $p = 0.766$)). *z*-curve analysis estimated an expected discovery rate (average statistical power) of EDR = 0.09 (95% CI [0.05, 0.65]) and an observed discovery rate of ODR = 0.32 (95% CI [0.23, 0.42]) (Fig. 5). Comparing the observed discovery rate with the 95% confidence interval of the expected discovery rate revealed no discrepancy and thereby provided further evidence of the absence of reporting bias.

Table 2 Test of moderators and subgroup analysis for categorial variables of pre-post performance

Moderator Subgroups	*N* (k)	ES (95% CI)	*p*-value
Age group	QE (94) = 213.86, $p < 0.0001$; F (5, 95) = 8.83, $p < 0.0001$		
5–9Y	4 (18)	−0.22 (−0.49 to 0.04)	0.0898
10–14Y	7 (12)	0.44 (−0.09 to 0.96)	0.0975
15–19Y	6 (28)	−0.47 (−0.95 to 0.02)	0.0585
20–24Y	9 (26)	0.33 (0.9 to 0.58)	0.0092
25–32Y	8 (14)	−0.06 (−0.55 to 0.44)	0.8032
> 65	1 (3)	1.65 (0.75 to 2.55)	0.0158
Testing protocol	QE (98) = 320.51, $p < 0.0001$; F (2, 987) = 0.37, $p = 0.9165$		
Block	22 (71)	0.08 (−0.24 to 0.4)	0.618
Random	4 (11)	0.12 (−0.52 to 0.75)	0.6936
Block/random	11 (19)	0.2 (−0.08 to 0.48)	0.154
Experience	QE (99) = 327.72, $p < 0.0001$; F (1, 99) = 3.09, $p = 0.0819$		
Experienced	8 (21)	−0.27 (−0.71 to 0.18)	0.2273
Unexperienced	30 (80)	0.19 (−0.05 to 0.42)	0.119
Sports	QE (88) = 267.22, $p < 0.0001$; F (12, 88) = 1.19 $p = 0.2994$		
Australian Rules Football	1 (2)	−0.44 (−11.36 to 10.49)	0.7
Badminton	2 (5)	0.98 (−1.13 to 3.08)	0.2663
Baseball (swing)	2 (6)	−1.11 (−1.72 to 0.51)	0.0053
Basketball	2 (2)	0.66 (−2.83 to 4.16)	0.2492
Bench press	1 (3)	0.41 (−3.6 to 4.42)	0.7032
Field hockey	2 (19)	−0.004 (−0.71 to 0.7)	0.991
Football	1 (2)	−0.01 (−2.69 to 2.678)	0.97
Golf	4 (9)	0.12 (−0.63 to 0.86)	0.7275
Gymnastic/dance	2 (2)	0.72 (−2.38 to 3.82)	0.2088
Kneeling/pistol shooting	2 (5)	0.51 (0.15 to 0.86)	0.0167

Table 2 (continued)

Moderator Subgroups	*N* (k)	ES (95% CI)	*p*-value
Motor skills/throwing/balance	6 (23)	0.07 (−0.64 to 0.78)	0.837
Tennis table	5 (1211)	0.13 (−0.42 to 0.69)	0.6093
Volleyball	5 (11)	−0.07 (−0.51 to 0.37)	0.7246
Skill	QE (88) = 245.12, p < .0001; F (12, 88) = 1.66, p = 0.0902		
Bench press	1 (3)	0.41 (−3.6 to 4.42)	0.7032
Combined skills	1 (2)	−1.3 (−5.72 to 3.12)	0.1667
Dribbling	3 (6)	−0.19 (−1.22 to 0.84)	0.6516
Golf putting/chipping	4 (9)	0.12 (−0.63 to 0.86)	0.7275
Hitting	2 (12)	−0.56 (−1.25 to 0.13)	0.1001
Jumping	1 (2)	0.03 (−3.64 to 3.7)	0.9341
Kicking	3 (3)	−0.38 (−1.91 to 1.14)	0.3933
Pass	**4 (8)**	**0.4 (0.07 to 0.73)**	**0.237**
Serve/set/strocke	9 (21)	0.26 (−0.16 to 0.68)	0.2091
Shooting	7 (13)	0.27 (0.21 to 0.68)	0.2658
Sprint/running	2 (4)	−0.14 (−0.86 to 0.59)	0.5913
Throwing	3(11)	−0.09 (−0.46 to 0.28)	0.5901
Others	5 (7)	0.4 (−0.72 to 1.51)	0.4283

Positive ES means in favor of low CI and negative ES means in favor of high CI

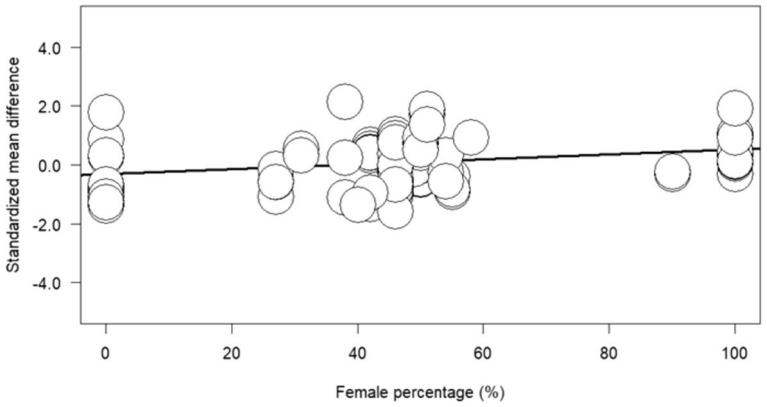

Fig. 3 Moderation effect of female % on acquisition gain

Table 3 Test of moderators for continuous variables of pre-post performance

	Original	
Moderator	Moderator test	Heterogeneity
Female percentage	F(1, 98) = 6.99, $p = 0.0095$	Q (98) = 319.27; $p < 0.0001$
Intervention durations	F(1, 98) = 0.78, $p = 0.3798$	Q (98) = 309.21; $p < 0.0001$
Number of training sessions	F(1, 99) = 0.12, $p = 0.73$	Q (99) = 330.79; $p < 0.0001$
Tot number of acquisition trials	F(1, 88) = 0.02, $p = 0.9042$	Q (88) = 299.17, $p < 0.0001$
Duration between pre-post	F(1, 98) = 0.51, $p = 04786$	Q (98) = 308.79, $p < 0.0001$

Difference Between Low and High CI Practices in the Δ of Post-Retention

Data from 27 studies investigating the difference between low and high CI practice in the Δ post-retention of performance outcomes were pooled in our MA. Outlier analysis identified five performance outcomes from studies by Menayo et al. (2010) and Pasand et al. (2016) as potential outliers due to their extreme effect sizes (|SMD| range between 3.69 and 12.02). Following outlier removal, the analysis proceeded. Given that the study by Pasand et al. (2016) investigated only one performance outcome, identified as outlier, this study was completely excluded from the multilevel meta-analysis. Accordingly, 26 included studies yielded a total of 76 ES. Based on the three-level meta-analytic model (Fig. 6), a significant difference between high and low CI practice in the Δ post-retention of performance outcomes was found (ES = −0.35, SE = 0.1, 95% CI = −0.55 to −0.15, t(75) = −3.42, $p = 0.001$). Significant heterogeneity was computed ($Q = 303.1$, df = 75, $p < 0.0001$). The estimated variance components were $\tau^2_{\text{Level 3}} = 0$ and $\tau^2_{\text{Level 2}} = 0.61$. This means that $I^2_{\text{Level 3}} = 0\%$ of the total variation can be attributed to between-cluster, and $I^2_{\text{Level 2}} = 80.01\%$ to within-cluster heterogeneity, with a total $I^2 = 80.01\%$. These results suggest that there is no variance attributed to between-cluster differences,

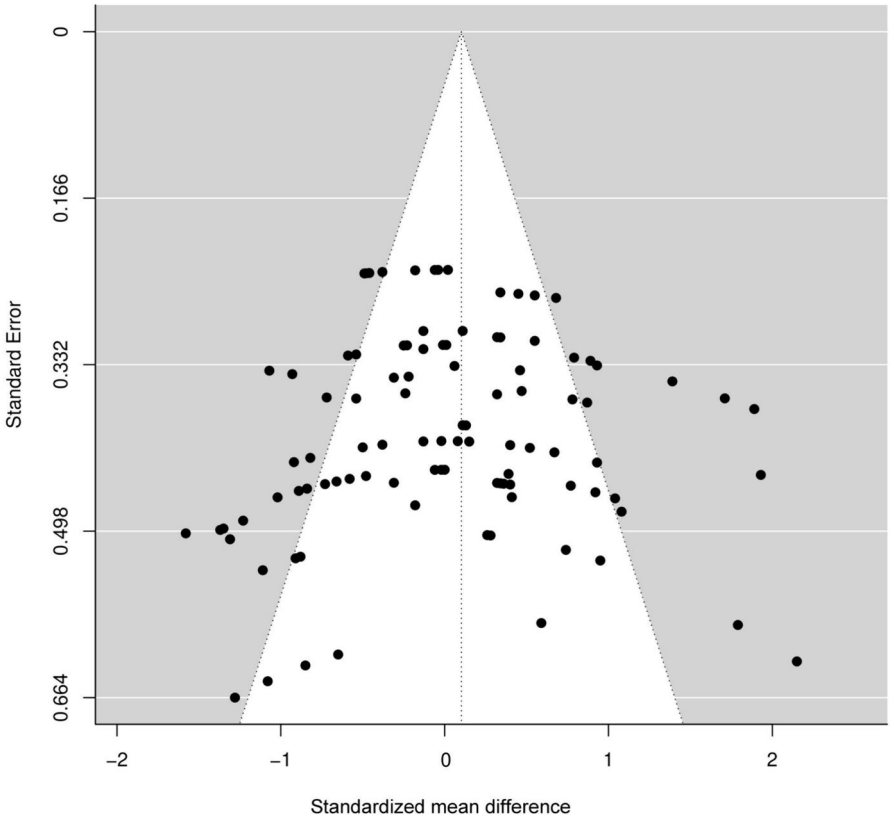

Fig. 4 Funnel plot for outcome performances at acquisition gain

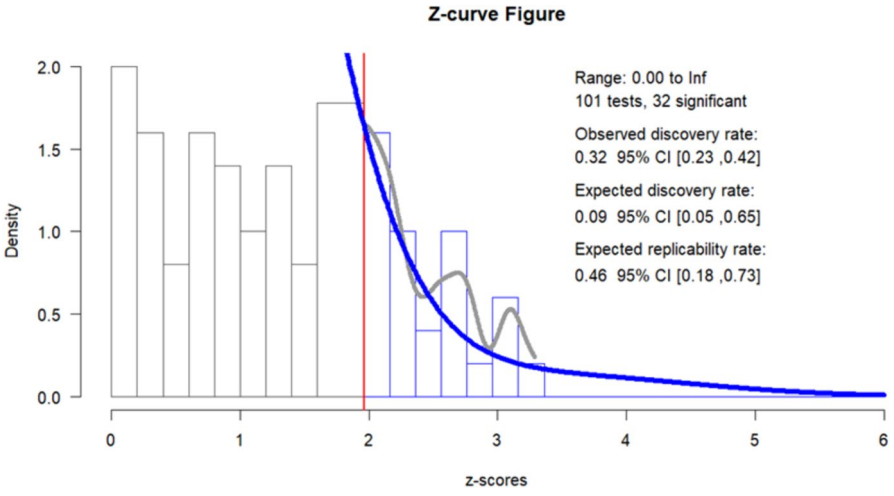

Fig. 5 *Z*-curve analysis of the included studies in the Δ pre-post

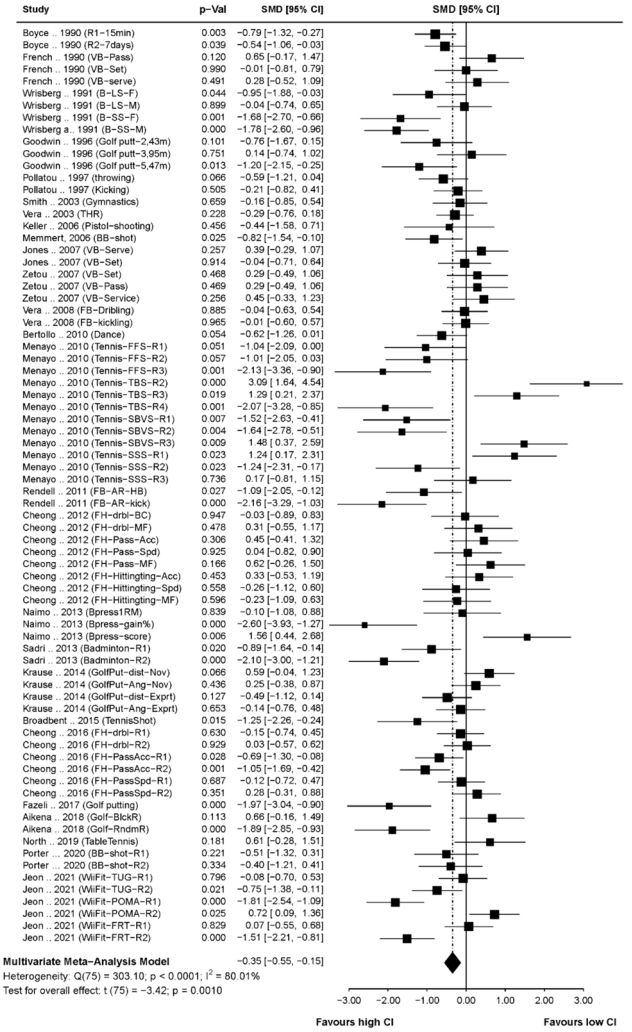

Fig. 6 Forest plot for the weighted difference between blocked and random practices at Δ post-retention. *Legends:* KS, kneeling shooting; B, badminton; LS, long serves; SS, short serves; M, male; F, female; TS, tennis strockes; LS, low skills; HS, high skills; VBS, volleyball serve; THR, throwing; VT, vertical target; HT, horizontal target; Tscore, total score; TB, tennis ball; FfB, feather fly ball; VB, volleyball; FB, football; FFS, flat forehand shot; TBS, topspin backhand shot; SBVS, slice backhand volley shot; SSS, slice serve shot; HB, handball; AR, Australian Rules; TUG, timed up and go test; POMA, Tinetti performance-oriented mobility assessment; FRT, functional reach test; FH, field hockey; PassAcc, push pass accuracy; PassSpd, push pass speed; NVS, nonvariable schedules; VS, variable schedules; drbl, dribble; BC, ball control; MF, movement form; FH1, field hockey acquisition 1; FH2, field hockey acquisition 2; Srv, serve; dis, displacement; Bpress, bench press; Uhd, underhand; CODA, change of direction ability; SSP, straight sprinting performance; BSB, baseball; DS, dominant side; NDS, non dominant side; Put, putting; dist, distance; ang, angle; NOV, novice; Exprt, experts

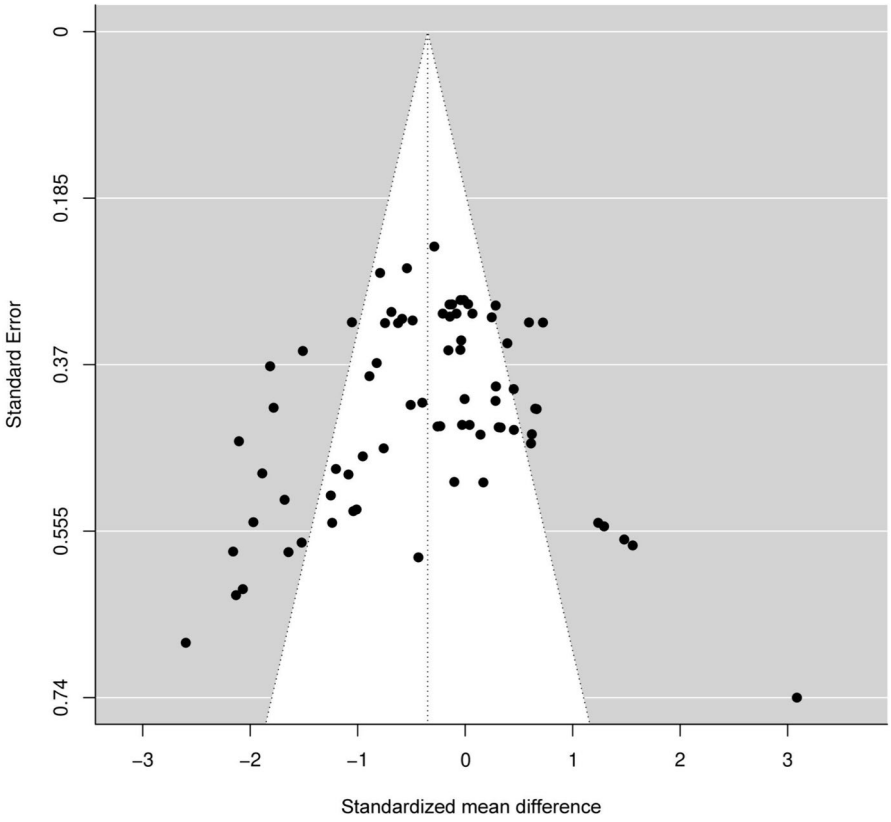

Fig. 7 Funnel plot for outcome performances at Δ post-retention

indicating that all moderators are expected to vary across studies rather than within the same cluster. Consequently, subgroup analyses and meta-regression are deemed unnecessary in this context.

Publication Bias

Visual inspection of the funnel plot (Fig. 7) revealed no evidence of publication bias, a conclusion confirmed by Begg and Mazumdar's rank correlation test (Kendall's tau = −0.11, $p = 0.15$) and the multilevel model of Egger's test ($F(1,74) = 2.23$, $p = 0.1395$). *z*-curve analysis estimated an expected discovery rate (average statistical power) of EDR = 0.41 (95% CI [0.05, 0.77]) and an observed discovery rate of ODR = 0.41 (95% CI [0.30, 0.53]) (Fig. 8). Comparing the observed discovery rate with the 95% confidence interval of the expected discovery rate revealed no discrepancy and thereby provided further evidence of the absence of reporting bias.

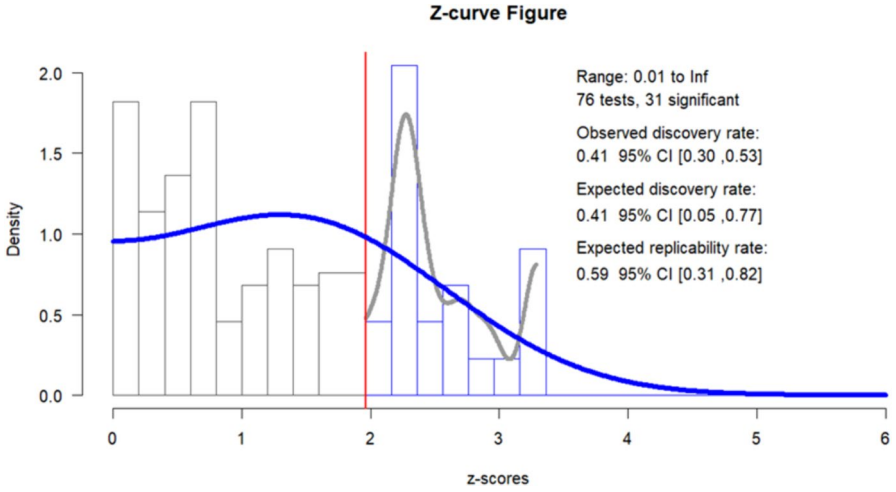

Fig. 8 *z*-curve analysis of the included studies in the Δ post-retention

Difference Between Low and High CI Practices in the Δ of Pre-Retention

Data from 28 studies investigating the difference between low and high CI practice in the Δ pre-retention of performance outcomes were pooled in our MA. Outlier analysis identified six performance outcomes from studies by Menayo et al. (2010), Rendell et al. (2011), Pasand et al. (2016), and Jeon et al. (2021) as potential outliers due to their extreme effect sizes (|SMD| range between 3.3 and 4.2). Following outlier removal, the analysis proceeded. Given that the studies by Pasand et al. (2016) investigated only one performance outcome, identified as outlier, this study was completely excluded from the multilevel meta-analysis. Accordingly, 27 included studies yielded a total of 75 ES. A small effect size based on the three-level meta-analytic model (Fig. 9), indicating a non-significant difference between low and high CI practice in the Δ pre-retention of performance outcomes (ES = −0.133, SE = 0.1, 95% CI = −0.33 to 0.06, t(74) = −1.36, $p = 0.179$). Significant heterogeneity was computed ($Q = 257.92.020$, df = 74, $p < 0.0001$). The estimated variance components were $\tau^2_{\text{Level 3}} = 0.03$ and $\tau^2_{\text{Level 2}} = 0.41$. This means that $I^2_{\text{Level 3}} = 5.32\%$ of the total variation can be attributed to between-cluster, and $I^2_{\text{Level 2}} = 69.32\%$ to within-cluster heterogeneity, with a total $I^2 = 74.64\%$. To identify potential sources of heterogeneity, subgroup analysis and meta-regression analysis were performed.

Moderator Analyses (of Multilevel Models)

Multivariate *s*ubgroup analysis revealed that "age groups" significantly moderate the difference between low and high CI practice in Δ pre-retention performance (F(5, 69) = 2.69, $p = 0.0384$). However, the "experience level," "sports," "skills," and

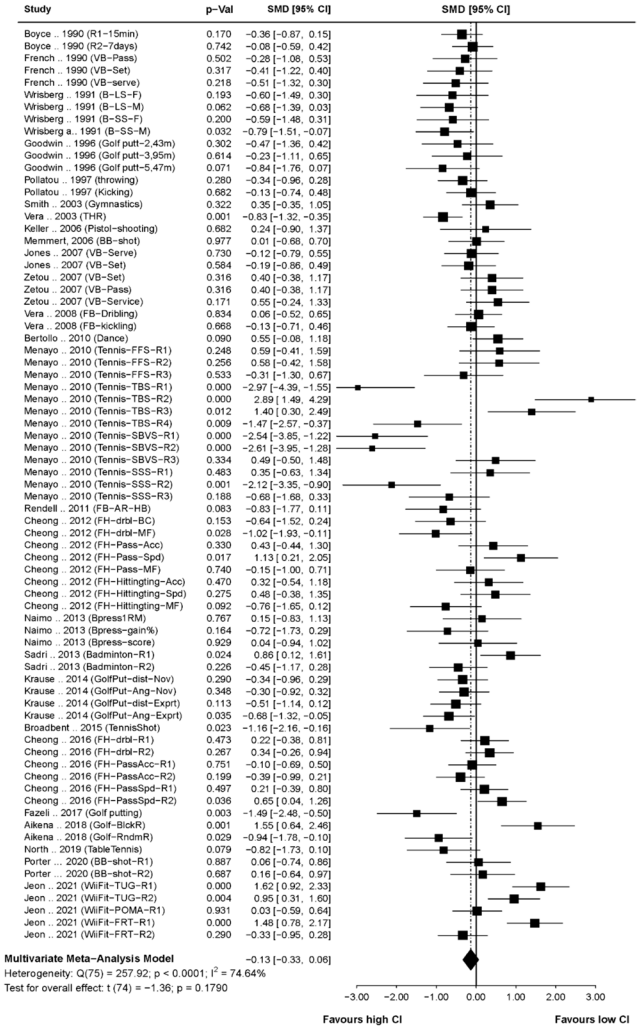

Fig. 9 Forest plot for the difference between blocked and random practices at relatively permanent gain. *Legends:* KS, kneeling shooting; B, badminton; LS, long serves; SS, short serves; M, male; F, female; TS, tennis strockes; LS, low skills; HS, high skills; VBS, volleyball serve; THR, throwing; VT, vertical target; HT, horizontal target; Tscore, Ttal score; TB, tennis ball; FfB, feather fly ball; VB, volleyball; FB, football; FFS, flat forehand shot; TBS, topspin backhand shot; SBVS, slice backhand volley shot; SSS, slice serve shot; HB, handball; AR, Australian Rules; TUG, timed up and go test; POMA, Tinetti performance-oriented mobility assessment; FRT, functional reach test; FH, field hockey; PassAcc, push pass accuracy; PassSpd, push pass speed; NVS, nonvariable schedules; VS, variable schedules; drbl, dribble; BC, ball control; MF, movement form; FH1, field hockey acquisition 1; FH2, field hockey acquisition 2; Srv, serve; dis, displacement; Bpress, bench press; Uhd, underhand; CODA, change of direction ability; SSP, straight sprinting performance; BSB, baseball; DS, dominant side; NDS, non dominant side; Put, putting; dist, distance; ang, angle; NOV, novice; Exprt, experts

Table 4 Test of moderators and subgroup analysis for categorial variables of pre-retention performance

Moderator Subgroups	*N* (k)	ES (95% CI)	*p*-value
Age group	QE (69) = 211.88, $p < 0.0001$; F (5, 69) = 2.69, $p = 0.0384$		
5–9Y	1 (1)	-	-
10–14Y	4 (8)	0.04 (−0.51 to 0.6)	0.8558
15–19Y	4 (14)	−0.08 (−0.4 to 0.25)	0.6202
20–24Y	8 (32)	−0.22 (−0.58 to 0.13)	0.2152
25–32Y	8 (14)	−0.42 (−0.74 to −0.08)	0.0175
> 65	1 (5)	0.74 (−0.35 to 1.82)	0.1316
Testing protocol	QE (72) = 246.55, $p < 0.0001$; F (2, 72) = 2.11, $p = 0.1293$		
Block	15 (39)	0.02 (−0.25 to 0.3)	0.8596
Random	4 (20)	−0.47 (−1.08 to 0.15)	0.1286
Block/random	7 (15)	−0.39 (−0.8 to 0.03)	0.0637
Experience	QE (73) = 248.56, $p < 0.0001$; F (1, 73) = 3.29, $p = 0.0739$		
Experienced	3 (16)	−0.53 (−1.33 to 0.28)	0.1815
Unexperienced	24 (59)	−0.09 (−028 to 0.1)	0.3613
Sports	QE (63) = 225.33, $p < 0.0001$; F (11, 63) = 0.9, $p = 0.5088$		
Australian Rules Football	1 (1)	-	-
Badminton	2 (6)	−0.27 (−1.38 to 0.84)	0.5613
Basketball	2 (3)	0.07 (−0.89 to 1.03)	0.7818
Bench press	1 (3)	−0.16 (−1.4 to 1.09)	0.6298
Field hockey	2 (14)	0.07 (−0.23 to 0.37)	0.6333
Football	1 (2)	−0.03 (−2.71 to 2.65)	0.9019
Golf	4 (10)	−0.42 (−0.93 to 0.09)	0.0926
Gymnastic/dance	2 (2)	0.46 (−2.58 to 3.49)	0.3052
Kneeling/pistol shooting	2 (3)	−0.19 (−0.93 to 0.57)	0.4152
Motor skill/throwing/balance	3 (8)	0.05 (−1.06 to 1.15)	0.9234

Table 4 (continued)

Moderator Subgroups	*N* (k)	ES (95% CI)	*p*-value
Tennis table	3 (15)	− 0.54 (− 1.42 to 0.34)	0.208
Volleyball	3 (8)	− 0.03 (− 0.63 to 0.57)	0.9055
Skill	QE (65) = 204.76, $p < 0.0001$; F (9, 65) = 1.47, $p = 0.1779$		
Bench press	1 (3)	− 0.16 (− 1.42 to 1.09)	0.6298
Dribbling	3 (5)	− 0.13 (− 1.05 to 0.78)	0.711
Golf putting/chipping	4 (10)	− 0.42 (− 0.93 to − 0.09)	0.0926
Hitting	1 (4)	− 0.02 (− 0.9 to 0.85)	0.9433
Kicking	2 (2)	− 0.13 (− 2.87 to 2.62)	0.6592
Pass	5 (9)	0.12 (− 0.38 to 0.61)	0.5979
Serve/set/volley/bump	6 (12)	− 0.14 (− 0.61 to 0.32)	0.5102
Shooting	7 (21)	− 0.37 (− 0.95 to 0.22)	0.2045
Throwing	2 (2)	− 0.62 (− 3.71 to 2.46)	0.2364
Others	3 (7)	0.65 (− 0.02 to 1.32)	0.0551

Positive ES means in favor of low CI and negative ES means in favor of high CI

Table 5 Test of moderators for continuous variables of pre-retention performance

Moderator	Moderator test	Heterogeneity
Female percentage	$F(1, 73) = 3.3$, $p = 0.0732$	$Q(73) = 247.9$; $p < 0.0001$
Intervention durations	$F(1, 71) = 0.62$, $p = 0.4354$	$Q(71) = 249.02$; $p < 0.0001$
Number of training sessions	$F(1, 73) = 0.72$, $p = 0.3978$	$Q(73) = 255.51$; $p < 0.0001$
Tot Number of acquisition trials	$F(1, 64) = 0.003$, $p = 0.955$	$Q(64) = 241.59$, $p < 0.0001$
Duration between pre-post	$F(1, 72) = 0.17$, $p = 0.6786$	$Q(72) = 256.79$, $p < 0.0001$

"testing protocol" did not moderate the difference between low and high CI practice in the Δ pre-retention performance (Table 4). No significant differences between low and high CI practices in Δ pre-retention performances were detected for the majority of analyzed subgroups in each category ($p > 0.05$). Better Δ pre-retention performance in favor of high CI practice was apparent for the age groups "25–32 years old" (ES = −0.42, $p = 0.018$) (Table 4).

Multivariate meta-regression analyses did not identify a moderating effect for any of the variables tested (i.e., female percentage, intervention duration, number of training sessions, total number of acquisition trials, or duration between post- and retention-test sessions (Table 5).

Publication Bias

Visual inspection of the funnel plot (Fig. 10) revealed no evidence of publication bias, a conclusion confirmed by Begg and Mazumdar's rank correlation test (Kendall's tau = −0.15, $p = 0.06$) but not by the multilevel model of Egger's test ($F(1, 73) = 4.86$, $p = 0.031$). *z*-curve analysis estimated an expected discovery rate (average statistical power) of EDR = 0.23 (95% CI [0.05, 0.76]) and an observed discovery rate of ODR = 0.28 (95% CI [0.19, 0.40]) (Fig. 11). Comparing the observed discovery rate with the 95% confidence interval of the expected discovery rate revealed no discrepancy and thereby provided further evidence of the absence of reporting bias.

Methodological Quality of Studies

Overall, the study quality was deemed to be fair to good (Table 6). The PEDro scale revealed a score of five to seven for all included studies with a mean PEDro score of 5.89 ± 0.66 (mean ± SD). Of the 36 included studies, 26 investigations received a good score of six (20 trials) or seven (6 trials) out of 10 (i.e., a double-blind but not triple-blind trial). The remaining studies (10 out of 37) received a fair score of 5 as they failed to randomize the trial, conceal the allocation, and/ or blind subjects, therapists and assessors.

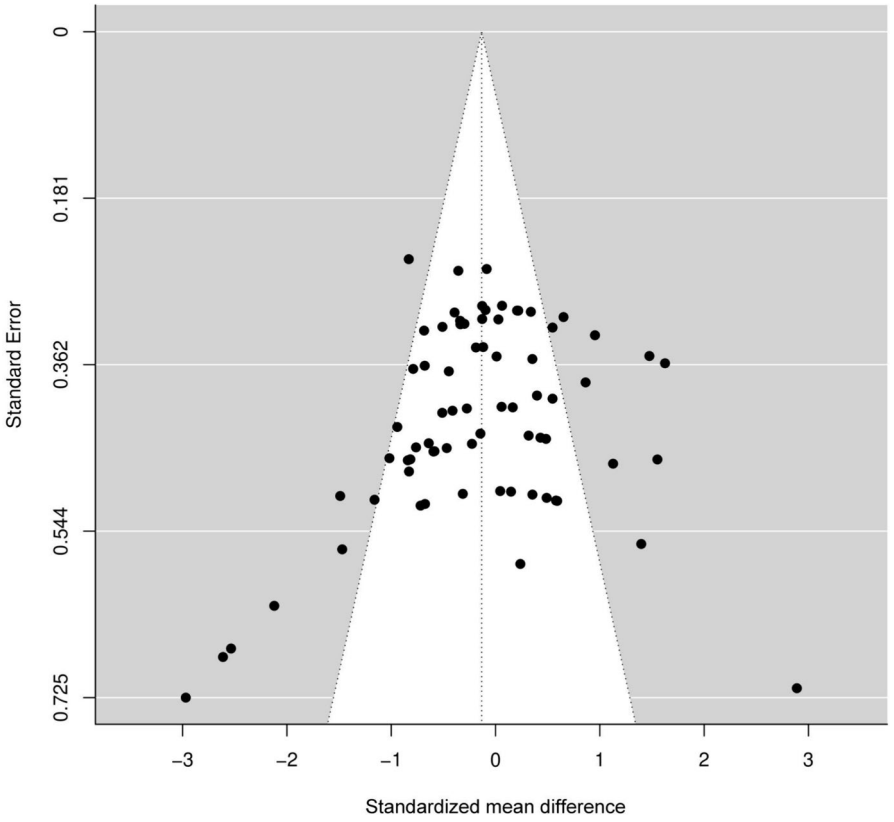

Fig. 10 Funnel plot for outcome performances at relatively permanent gain

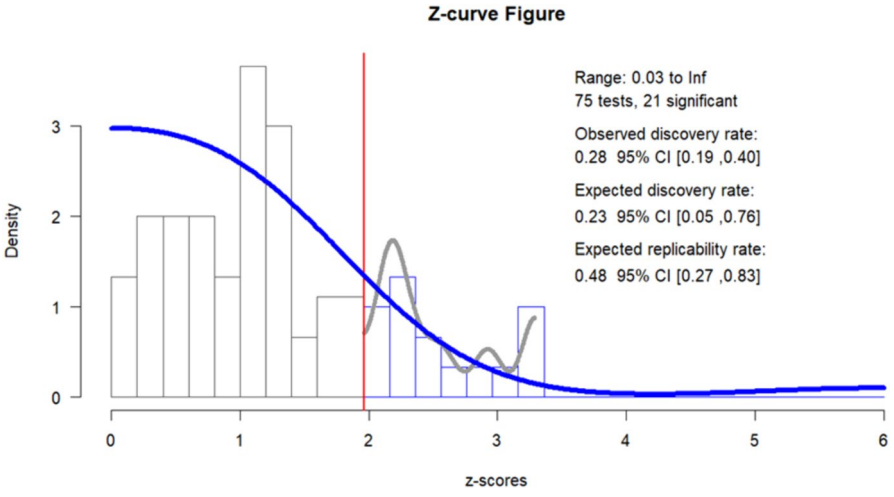

Fig. 11 *Z*-curve analysis of the included studies in the Δ pre-retention

Table 6 Methodological quality of the included studies assessed with the Pedro scale

Included study	Eligibility criteria	Random allocation	Concealed allocation	Baseline similarity	Blinding of subjects	Blinding of therapists	Blinding of assessors	Outcome obtained from > 85% of allocated subject	Intention to treat analysis	Between group comparisons	Point measure and variability	Score
Boyce & Del Rey., 1990	1	1	0	1	1	0	0	1	1	1	1	**7**
French et al., 1990	0	1	0	1	0	0	0	1	1	1	1	**6**
Wrisberg and Zhan Liu. 1991	0	0	0	1	0	0	0	1	1	1	1	**5**
Hall et al., 1994	1	1	0	1	1	0	0	1	1	1	1	**7**
Goodwin & Meeuwsen, 1996	1	1	0	1	0	0	0	1	1	1	1	**6**
Hebert et al., 1996	1	1	0	1	0	0	0	1	1	1	1	**6**
Pollatou et al., 1997	1	1	0	1	0	0	0	1	1	1	1	**6**
Bortoli et al., 2001	1	1	0	1	0	0	0	1	1	1	1	**6**
Meira. And Tani 2003	1	0	0	1	0	0	0	1	1	1	1	**5**
Smith et al., 2003	1	1	0	1	1	0	0	1	1	1	1	**7**
Vera & Montilla., 2003	1	0	0	1	0	0	0	1	1	1	1	**5**
Keller et al., 2006	1	1	0	1	1	0	0	1	1	1	1	**7**
Memmert, 2006	1	0	0	1	0	0	0	1	1	1	1	**5**
Jones & French, 2007	1	1	0	1	0	0	0	0	1	1	1	**5**

Table 6 (continued)

Included study	Eligibility criteria	Random allocation	Concealed allocation	Baseline similarity	Blinding of subjects	Blinding of therapists	Blinding of assessors	Outcome obtained from > 85% of allocated subject	Intention to treat analysis	Between group comparisons	Point measure and variability	Score
Zetou et al., 2007	1	1	0	1	1	0	0	1	1	1	1	**7**
Vera et al., 2008	0	1	0	1	0	0	0	1	1	1	1	**6**
Bertollo et al., 2010	1	1	0	1	0	0	0	1	1	1	1	**6**
Menayo et al., 2010	1	1	0	1	0	0	0	1	1	1	1	**6**
Travlos, 2010	1	1	0	1	0	0	0	1	1	1	1	**6**
Rendell et al., 2011	1	0	0	1	0	0	0	1	1	1	1	**5**
Cheong et al., 2012	1	1	0	1	0	0	0	1	1	1	1	**6**
Naimo et al., 2013	1	0	0	1	0	0	0	1	1	1	1	**5**
Sadri et al., 2013	0	1	0	1	0	0	0	1	1	1	1	**6**
Yanci et al., 2013	1	1	0	1	0	0	0	1	1	1	1	**6**
Krause et al., 2014	1	1	0	1	0	0	0	1	1	1	1	**6**
Broadbent et al., 2015	1	0	0	1	0	0	0	1	1	1	1	**5**
Yanci et al., 2015	0	1	0	1	0	0	0	1	1	1	1	**6**
Cheong et al., 2016	1	0	0	1	0	0	0	1	1	1	1	**5**
Pasand et al., 2016	1	1	0	1	0	0	0	1	1	1	1	**6**
Buszard et al., 2017	1	1	0	1	0	0	0	1	1	1	1	**6**

Table 6 (continued)

Included study	Eligibility criteria	Random allocation	Concealed allocation	Baseline similarity	Blinding of subjects	Blinding of therapists	Blinding of assessors	Outcome obtained from > 85% of allocated subject	Intention to treat analysis	Between group comparisons	Point measure and variability	Score
Fazeli et al., 2017	1	1	0	1	0	0	0	1	1	1	1	**6**
Aiken & Genter., 2018	0	1	1	1	0	0	0	1	1	1	1	**7**
North et al., 2019	1	0	0	1	0	0	0	1	1	1	1	**5**
Porter et al., 2020	1	1	0	1	0	0	0	1	1	1	1	**6**
Sharp et al., 2020	1	1	0	1	0	0	0	1	1	1	1	**6**
Jeon et al., 2021	1	1	0	1	0	0	0	1	1	1	1	**6**

Discussion

The paradoxical effects of CI assume high CI practices with a random schedule hinder performances during the acquisition phase (i.e., "cognitive load theory," CLT Gentile, 1998), while providing more permanent enhancement during the retention phase (i.e., the "elaborative" (Shea & Zimny, 1983) and the "forgetting/reconstruction" hypotheses, Lee & Magill, 1985). Considering Schmidt and Lee's (2019) definition of motor learning, the generalization of CI effects on motor abilities implies a lower acquisition gain and a greater relatively permanent gain in the capability of skilled performance, following high versus low CI practice.

The results of the current systematic review and meta-analysis indicate that neither the CI effects nor the associated hypotheses can be generalized to gross/sport motor skills. There is no statistical evidence that high CI practice leads to systematic lower acquisition gain (pre-post %change = 28.32 ± 58.41% and 28.08 ± 53.21% for low and high CI schedules, respectively) and/or greater relatively permanent gain (pre-retention % change = 23.09 ± 60.46% and 24.40 ± 93.91% for low and high CI schedules, respectively) in PE/sports settings.

Additionally, Δ in only 37 out of 183 overall pooled performance outcomes (20%) agreed with the paradoxical CI effects on the acquisition (22 out of 102 for Δ pre-post) and the relatively permanent (15 out of 81 for Δ pre-retention) gains (Table 1). Consequently, no statistically significant overall difference between low and high CI practices was detected for any of these gains. Interestingly, the 20% consensus is of the same order of magnitude or even smaller as found in replication studies in the fields of medicine and psychology (Baker, 2016).

It should be emphasized that an overall significant difference in favor of high CI practice, in coherence with the CI theory, has been detected in Δ post-retention (post-retention % change = − 2.39 ± 22.29% and − 0.33 ± 25.09% for low and high CI schedules, respectively). However, this difference was not large enough (ES = − 0.35) to produce an overall greater gain (i.e., Δ pre-retention) following high CI practice.

In line with these overall findings, the subgroup analyses and meta-regressions revealed that only in extremely limited "conditions" can the CI paradoxical effects on the acquisition (i.e., conditions: age groups 20–24 and > 65 years old, female participant, and kneeling/pistol shooting sports) or the relatively permanent gains (i.e., conditions: age group 25–32 years old) be partially corroborated. Importantly, no paradoxical CI effects were concurrently identified for both gains in any of the tested subgroups.

Moreover, only four out of the twenty-six studies (15%) focusing on both acquisition and retention phases approved the paradoxical CI effects on both acquisition and relatively permanent gains in some performance outcomes (1 out of 4 outcomes, Wrisberg and Liu (1991) and Menayo et al. (2010); 1 out of 2 outcomes, Aiken and Genter (2018); and 1 out of 1 outcome, Pasand et al. (2016)) (Table 1). These findings further support the limited paradoxical CI effects in a PE- and/or sports-related context.

Since the majority of sports skills involve whole-body motion and require coordinated movements within several limbs in parallel, increased cognitive and physical loads are generated (Gebkenjans et al., 2007), and intra-task interference can be created within repetitive skills (Albaret & Thon, 1998). A previous report suggested that this intra-task interference could supersede the inter-task interference generated via CI, thereby limiting the CI effects in gross/sports skills (Albaret & Thon, 1998). Likewise, high CI practice has been suggested to overload the learners' working memory during the acquisition phase, engaging them in deeper elaborative and/or reconstruction processes (Lee & Magill, 1983, 1985; Shea & Zimny, 1983). The combination of high CI training with complex and highly demanding motor tasks (e.g., sport skills) should therefore result in the learner's cognitive abilities being overloaded even more or earlier, thereby further decreasing the likely benefit of a high CI schedule, resulting in reduced Δ pre-post performance gains. While CI-related studies of fine motor skills have primarily examined sequential movements with dominant visual-spatial input, for which the working memory model was originally developed (Baddeley, 1992, 2018), most CI studies in sport involve movements with highly parallel tasks and dominant proprioceptive and kinaesthetic sensations. Similarly, the CLT was developed for phenomena related to cognitive learning tasks (Sweller, 1988), and especially related to secondary biological knowledge (Sweller, 2023), with similar characteristics as in the working memory model of Baddeley.

Differences Between Low and High CI Practices in the Acquisition Gain

The first hypothesis of the CI paradoxical phenomenon suggests that low CI practice better facilitates the acquisition of skilled performance, while practicing under high CI schedules leads to inferior acquisition (Battig, 1966; Lee & Magill, 1985; Schmidt et al., 2011; Shea & Morgan, 1979). This first CI-phenomenon is typically explained by an increased demand on working memory due to a random schedule compared to a blocked schedule (Magill & Hall 1990). According to the CLT, applying a random schedule, particularly during the early stage of complex skill learning, may lead to system overload given the high attention, memory, and motor demands, therefore hindering performance during the acquisition phase (Gentile, 1998).

Since the working memory of children and beginners is assumed to become overloaded more quickly, as they either have a smaller working memory than adults (Reynolds et al., 2022) or the exercises contain even more subjective information due to their novelty (Amico & Schaefer, 2022; Schöllhorn et al., 2022), this first CI-effect should logically be more pronounced in children and beginners (at least in some studies). The present subgroup analysis contradicts this hypothesis and shows no difference between low and high CI in both 5–9 (small ES = −0.2) and 10–14 (medium ES = 0.6) years old age groups.

The fact that beginners and children do not respond according to the CI model may be due to two conditions more in line with the differential learning (DL) model (Ammar et al., 2024; Apidogo et al., 2021). Firstly, children (and beginners)

show increased variability in the execution of movements during repetitions (e.g., block of repetition in low CI), which in the case of large motor movements can reach a level corresponding to that which is also produced in high CI. Similar to the training principle of artificial neural networks, it has been shown switching between different movements significantly increases the variance of the individual movements compared to a blocked sequence, which leads to more stable responses after acquisition (Apidogo et al., 2024; Gebkenjans et al., 2007; Schöllhorn, 2016). Evidence for the dependence of learning on the degree of variance was provided by Wu et al. (2014). If a low and a high CI sequence leads to a comparable degree of variance in the to-be-learned movements, then similar learning progress can be expected and hardly any differences can be recognized. Secondly, the working memory of children can be assumed to have less capacity or, at least, be structured differently than that of adults, which may result in a different processing of extensive information. Alternatively, instead of a quantitative overload, a qualitative change in the load of information processing within the working memory, similar to the chunking process surpassing the threshold of 7 ± 2 stimuli (Miller, 1956), has been suggested as a possible reason for the different effects (Schöllhorn et al., 2022).

This is consistent with the phenomena observed in DL studies, where a greater number of variants are performed compared to CI related studies, leading to regularly increased acquisition and learning rates, even in adults (e.g., Apidogo et al., 2024). In DL, there are not only 3–6 prescribed exercises repeated at different distances, but each exercise is varied, resulting in a complete overload of a classical working memory. Thus, in comparison to blocked learning, a significantly lower performance gain during the acquisition phase is to be expected, if the cognitive load theory would be generalized to physical tasks. However, at least the same performance increases are observed (e.g., Ammar et al., 2024; Apidogo et al., 2024), further increasing the subsequent learning phase (Schöllhorn et al., 2014, 2015). In summary, studies on DL provide additional evidence for the CLT's ensuing uncertainty, thereby prompting us to reconsider the explanations used before proceeding with drawing contradictory conclusions and invalid generalizations.

Another approach for explaining the missing CI-effect in children may relate to different evaluation strategies of errors in children and adults, providing a further argument in favor of DL. Indeed, in addition to the variety of exercises, DL is characterized by the absence of augmented feedback (Schöllhorn, 2000), which is argued to be ineffective when sufficient exercise variation is provided (Silverman et al., 1992; Ramirez Ruiz et al., 2022; Wu et al., 2011) or can even be misleading (Buekers et al., 1992). The lack of augmented feedback might lead to different error processing strategies, which Broadbent et al. (2017) suggested could reduce cognitive overload in randomized learning sequences by minimizing the focus on error correction. Since DL results in high error rates at least similar to those in randomized CI, but without the observed overload effects post-acquisition, the way deviations are assessed and processed could crucially influence learning effectiveness. The differences in error evaluation between children and adults, and how these evaluations vary by learning method, warrant further investigation.

More recent studies have used neurophysiological measurements to evaluate the CI effects and/or the CLT following different levels of CI practices in motor skill

acquisition. The majority of them were able to show that high CI practice was associated with higher cognitive load in fine motor skills (e.g., finger sequence production), as evidenced by the involvement of a wide range of motor-skill-related neural regions (Hardwick et al., 2013; Penhune & Steele, 2012), activation of the default or scaffolding network (Cross et al., 2007; Wymbs & Grafton, 2009), an increased frontal–parietal blood oxygenation signal, and higher motor cortex (M1) excitability during the acquisition phase (Lin et al., 2011). However, there is little evidence on the applicability of the CI first hypothesis and the CLT when practicing sport/gross movements. Existing literature reports altered electrical brain activity with high CI practice being coupled with lower left temporal alpha prior to golf putting (Gallicchio and Ring 2019), stronger reduction of the P3 component during target throwing training (Frömer et al., 2016), and increased prefrontal gamma—associated with stress—during the acquisition of a badminton serve (Henz et al., 2018).

Aside from the limited neurophysiological evidence for the generalization of the CLT to sport settings, the existing findings have revealed numerous inconsistencies regarding the CI effects on sports skill acquisition and performance. When comparing the difference between low and high CI practices in terms of post-test performance, some studies show superior acquisition performance for the low CI group (Boyce & Rey, 1990; Pasand et al., 2016; Sadri et al., 2013), while others show no significant differences between the two practices (Bortoli et al., 2001; Broadbent et al., 2015; Fazeli et al., 2017; Goodwin & Meeuwsen, 1996; Meira & Tani, 2003; Memmert, 2006; Pollatou et al., 1997; Porter et al., 2020; Zetou et al., 2007); sometimes even an inverted first CI effect with higher acquisition performance is observed following random practice (Aiken & Genter, 2018; Hall et al., 1994; North et al., 2019; Sharp et al., 2020). Equivocal conclusions have been reached in previous review studies, with some of them supporting the CI effects (Brady, 1998, 2004), while others reported the difficulty of generalizing the CI effects in applied sport settings (Barreiros et al., 2007; Shewokis, 1997; Wang et al., 2022). Similar to original investigations, these review papers have mainly focused on the difference between high and low CI practices in retention-test performance, with little attention given to the gain during the interference part after the acquisition phase. Additionally, these reviews did not exclude studies failing to guarantee equivalent baselines between high and low CI conditions, calling their conclusions into question.

Focusing on the performance gains during the acquisition phase in studies that succeed in guaranteeing an equivalent baseline in low and high CI conditions, the current meta-analysis provides statistical evidence supporting previous reviews highlighting the problem of generalizing the CI effect on sport performance during the acquisition phase. Only 20 performance outcomes (out of 101 pooled outcomes = 19.8%) demonstrated superior gain after low vs. high CI practice, which agrees with the CI effect during the acquisition phase, whereas more than four times as many (80.2%) outcomes disprove this first CI effect and show either non-significant differences (69 outcomes) or even a reversed effect (12 outcomes). Overall, there was no statistically significant difference in performance gain following low vs. high CI practices after the acquisition phase, according to the

current meta-analysis. Therefore, the recommendation to start with blocked training during the early stage of complex skill learning cannot be derived logically from the present meta-analysis data.

Differences Between Low and High CI Practices in the Relatively Permanent Gain

The second hypothesis of the CI paradoxical phenomenon suggests an overcompensation of the short-term disadvantage of CI during the acquisition phase with a relatively permanent gain in the retention phase (Battig, 1966; Lee & Magill, 1985; Shea & Morgan, 1979). The suggested long-term advantages in the subsequent learning phase were most often assigned to either the "elaborative" or the "forgetting/reconstruction" theoretical positions. By using various encoding strategies, the elaborative hypothesis suggests that a high CI schedule engages the learner in a deeper and more elaborative process, resulting in a more distinct and sophisticated task representation (Shea & Zimny, 1983). While the forgetting/reconstruction theoretical position suggests that changing the characteristics of a task between trials results in forgetting motor processing, which engages the learner in a plan-of-action reconstruction process and necessitates extra retrieval practice to complete each task (Lee & Magill, 1983, 1985). Although controversial, both theoretical positions assume a more detailed and permanent mental representation of the task, which can be beneficial for the retention performance following high CI practice. More recent studies examining the effects of CI on motor learning have used neurophysiological measurements to assess the evidence for these early theoretical positions.

However, there is scarce indirect evidence for the emergence of a specific behavioral or neural signature or profile during high-CI schedules versus low-CI schedules (Wright et al., 2016), and if any, it is indirect, with numerous additional assumptions. More surprisingly, contradicting the early hypotheses of more elaborative and/or extra reconstruction processes following high CI practice, Wright et al. (2004) reported that during extensive acquisition under high CI practice, the learner can use similar motor programming strategies, but perform very differently during the retention testing phases. More unexpectedly, when focusing on fine motor skill (a finger-tapping task), Lin et al. (2011) reported a decreased BOLD signal in frontal–parietal regions (important motor regions for training) during the retention test following interleaved practice.

The limited neurophysiological evidence supporting the "elaborative" and the "forgetting/reconstruction" theoretical positions calls into question the generalization of the second CI hypothesis, suggesting a longer-term advantage of high CI practice, particularly in complex sport performances. This uncertainty can also be derived from numerous inconsistencies about the CI effects on the retention performance of sports skills. When comparing the difference between high and low CI practices in terms of retention-test performance, some studies reveal that the high CI group demonstrated superior retention performance (Hajihosseini, 2016; Kalkhoran & Shariati, 2015; Memmert, 2006; Pasand et al., 2016), while others showed no significant differences between the two practices (Naimo et al., 2013;

Porter et al., 2020; Sharp et al., 2020; Yanci et al., 2013; Zetou et al., 2007). In some cases, an inverted first CI effect was observed, with higher acquisition performance following random practice (Farrow & Maschetie, 1997; Hall et al., 1994; Zipp & Gentile, 2010). The contradictory responses to high CI designs become plausible if, in addition to the pure exercise sequence, the age and specific conditions of the movements to be learned are also taken into account. The results confirming the first CI effect were exclusively observed in adults performing movements with high visual components (e.g., pistol shooting, volleyball passes and playing at a target, throwing at a basket, and baseball hitting balls with varying curved trajectories). Post-test results contradicting the first CI effect were typically found in children or adolescents acquiring movements with dominant kinesthetic and proprioceptive components (e.g., bench press, Frisbee throwing, sprinting). An exception in this category is the basketball study by Porter et al. (2020), which adhere to Miller's (1956) chunking threshold (7 ± 2) by involving throws from 8 different locations for the high CI group, largely exceeding the 4 locations in Memmert's (2006) study, and constant location for the control group. The assumption that the acquisition impairment is a prerequisite for subsequent increased learning rates was not supported by studies on differential learning, where increased acquisition rates were followed by equally increased learning rates. In line with the original findings, recent systematic reviews also noted the difficulty of generalizing the CI phenomenon to applied sports skills (Ammar et al., 2023; Wang et al., 2022).

Despite the fact that performance gain, particularly the relatively permanent one, is considered a major component in the field of motor learning (Schmidt & Lee, 2019), these previous studies (original and review) have evaluated the second CI effect only by assessing the difference between both practices in terms of performance in the retention test, with little attention given to the relatively permanent gain.

The current meta-analysis provides additional statistical evidence on the myth of CI effects and possible generalizations in a sport context by focusing on the relatively permanent performance gain, as reflected in the difference between baseline (pre-test) and retention performances. Only 25 out of 75 pooled performance outcomes (33.3%) demonstrated superior gain following random vs. blocked practice. This agrees with the second CI effect, whereas nearly twice as many ($50 = 66.6\%$) outcomes contradict the first CI effect, showing either non-significant differences (46 outcomes) or even a reversed effect (4 outcomes). Overall, the present meta-analysis revealed no statistically significant difference between relatively permanent performance gain following high vs. low CI practices after the "acquisition + random" phases. Therefore, the recommendation to employ high CI practices to gain a longer-term performance advantage cannot be derived from the present evidence.

Moderator Variables of the CI Effects on Performance Gains

The lack of scientific consensus regarding the evidence of CI effects in sport practice has been attributed to various factors, such as participants' age and

experience levels, complexity and nature of the sports-related skill, and duration of the acquisition phase (Barreiros et al., 2007; Brady, 1998, 2008; Graser et al., 2019; Magill & Hall, 1990; Merbah & Meulemans, 2011; Wang et al., 2022). The various moderators of the CI hypothesis have historically explained the lack of differences in skilled performances measured at post- and/or retention testing, with little attention paid to the moderation effects on possible gains (e.g., acquisition gain and/or relatively permanent gain) generalized by each amount of CI. The present meta-analysis with subgroup and meta-regression analysis addressed this gap, revealing that most of the studied variables including the learner's experience level, as well as the acquisition protocol (i.e., practice duration, number of training sessions and acquisition trials, duration between testing sessions, and testing protocol), did not moderate the difference between low and high CI practices in terms of acquisition and/or relatively permanent gain.

Indeed, no significant moderation effect has been identified for any of these variables for the CI effect on the acquisition (Δ pre-post performances) and relatively permanent gain (Δ pre-retention performances), while only age group significantly moderates both CI effects (acquisition and relatively permanent gains) and female % significantly moderates the CI effect only on acquisition gains.

In terms of participants' age, pre-post Δ performances were in favor of low CI practice in older adults (> 65 years old, $p = 0.02$) with a large ES of 1.65, and in population aged between 20 and 24 years old ($p = 0.01$) with a small ES of 0.33. Furthermore, the Δ pre-retention performances were in favor of high CI practice in studies involving participants aged between 25 and 32 years old ($p = 0.02$ with a small ES of − 0.42).

The significant moderation effects of age on both CI effects support previous evidence for a significant interaction of participant age with the CI effect (Jeon et al., 2021; Pollock & Lee, 1997). Particularly, the greater relatively permanent gain using the high compared to the low CI practice, only observed in populations aged between 25 and 32, aligns with findings from previous reports indicating that CI effects seem to exist more in young adults than in children (Pollock & Lee, 1997) or older adults (Jeon et al., 2021). One possible explanation for the moderating effect of age is age-related changes in cognitive function. In this context, Lin et al. (2012) found that the brain regions associated with the main activities seem to differ after high CI practice at different ages. More recently, the same research groups revealed psychophysiological interactions for functional MRI data between age group and high vs. low CI practices of a motor skill. Indeed, following a high CI practice, younger adults showed efficient brain network topology, stronger functional segregation, and a significant association between higher network centrality and better learning, while older adults did not exhibit those favorable network properties (Lin et al., 2016).

The greater pre-post difference (Δ) observed with low compared to high CI practice in the older adult population (aged over 65 years, as studied by Jeon et al., 2021) may be attributed to their familiarity with learning through repetition, characteristic of blocked practice (i.e., a series of repeated movements). This approach aligns with the repetitive learning (RL) model, grounded in early learning theories like behaviorism, and emphasizes the importance of repeating correct trials

to enable the learner to execute specific skills with enhanced precision and accuracy. For older adults, the RL model was a foundational educational approach, preceding more recent models like variability of practice (VP) and high CI practice. These recent models focus on either varying the variable parameters (e.g., absolute timing, absolute forces), but not the invariant ones, as in the case of VP (Adams, 1987; Schmidt, 1975), or varying both the invariants (represented as generalized motor programs (GMPs)) and the schedule (i.e., the temporal structuring) of the learning process, as in the case of CI (Battig, 1979; Shea & Morgan, 1979). Both models were developed in the last quarter of the twentieth century (Schöllhorn et al., 2022), but have taken time to be integrated into educational systems.

Nevertheless, the benefits of low over high CI practice during the acquisition phase in older adults should be interpreted with caution. The present age-related subgroup analysis was based on data from only three pooled outcomes, all extracted from a single study by Jeon et al., 2021, which had a PEDro score of 6, indicating that it barely meets the criteria for good quality. Moreover, the computed confidence interval for this subgroup's ES was large (0.75 to 2.55). The limited scope of this data may not comprehensively reflect the impact of the CI's practice models on the older population.

In terms of sex's moderation effects, the present result showed a benefit of low over high CI practice in studies with higher female proportions. These results indicate the higher the proportion of females in a study, the higher the chance of getting acquisition performance gains in favor of the low CI practice. Thus, the better the chance of being in line with the first CI hypothesis. The effect of gender on the CI phenomenon has rarely been investigated. The findings of the present meta-analysis confirm the previous findings of the seminal study by Smith and Rudisill (1993), investigating the influence of proficiency level and gender on the CI effect, with an open skill and reported only females to exhibit the contextual interference effect during the acquisition phase. The author mentioned the difficulty of explaining this difference between the sexes in terms of the first CI effect, speculating that the sex difference could be due to the lower experience level of females in open skill compared to males. They also highlighted the need for further research to determine whether proficiency can account for interactions between CI and gender. Moreover, given the differences in the neurophysiological processing of language between women and men (Baxter et al., 2003), and considering that language processing occurs in the same areas responsible for sequential and visual-spatial information processing (Binding et al., 2022; Manohar et al., 2017; Meyer et al., 2005), it is plausible to suggest that most studies on CI learning have predominantly focused on very specific properties of the medio-lateral temporal (MLT) cortex (Cutsuridis & Yoshida, 2017; Patel et al., 2023) involved in a very specific form of movements, indicating a narrow scope of investigation. The extent to which these differences lead to women holding on to a mnemonic strategy for longer due to their different language processing, even when learning sequential movement forms due to possibly greater capacities, before switching to modified chunk-like processing and thus triggering clearer CI effects, requires extensive new and interdisciplinary research. However, the close relationship between bilingualism and episodic memory benefits in older adults (Schroeder & Marian, 2012), and the

close link of episodic memory and MLT (Cutsuridis & Yoshida, 2017), provides a new opportunity to create alternative approaches to training these brain areas.

A similar explanation of the current findings is related to gender differences in terms of basic cognitive resources and their associated neuroanatomical functioning, particularly working memory (Clements et al., 2006; Gur et al., 2000; Harness et al., 2008; Hill et al., 2014; Saylik et al., 2018; Shaywitz et al., 1995). Evidence for specific gender-related working memory networks suggest that females consistently activate more limbic (e.g., amygdala and hippocampus) and prefrontal structures (e.g., right inferior frontal gyrus), and males activate a more distributed network including more parietal regions (Hill et al., 2014). On the other hand, many studies have previously shown that, during the acquisition of motor skills, high CI practice is associated with greater cognitive engagement compared to low CI practice (Cross et al., 2007; Hardwick et al., 2013; Henz et al., 2018; Lin et al., 2011; Penhune & Steele, 2012; Wymbs & Grafton, 2009). This was evidenced by a greater level of activation in frontal areas and increased working memory demand due to non-systematic variations (Lage et al., 2015). Taking into consideration a lower female ability to activate a more distributed working memory network compared to males (Hill et al., 2014), it is possible that the working memory resources of high CI practice in female learners may be more quickly overwhelmed by the magnitude of visual-spatial information required to generate multiple motor solutions (i.e., task switching) during acquisition, which may explain the observation that this population benefits more from the low CI practice in this acquisition phase.

Future research is required to demonstrate whether studies on working memory training help compensate for deficits in this area. Evidence for the trainability of the linguistic (sequential)-visual spatial domain is provided by studies in children with ADHD (Klingberg et al., 2005) and chimpanzees (Tang et al., 2019). The conditions under which this seems possible are extensively discussed (Jacoby & Ahissar, 2013); however, to what extent the capacities of specific working memories can be enlarged by training and to what extent this enlargement allows a non-specific transfer between the different types of working memory is still unclear. Interestingly, even the components of the multi-component model of working memory according to Baddeley and Hitch (1974) are limited to verbal and visuo-spatial aspects, but do not consider the proprioceptive, kinaesthetic, and tactile domains (Baddeley, 2018).

In contrast to the acquisition and relatively permanent gain, participants' age and gender, similar to their experience level and testing protocol, showed that they did not significantly moderate the CI effect on post-retention performance.

Taken together, a significant difference between high and low CI practice was only computed during the retention phase (post-retention period). However, the magnitude of this difference was not large enough to generate a significant overall difference on the relatively permanent gain (pre-retention). The present results reveal that CI effects on motor learning gain can only be confirmed under very limited conditions, challenging the extension of CI effects to the sport context in general.

Methodological Quality of the Studies

The studies included in the current MA have an overall quality ranging between 5 and 7 out of 10, which is considered to be a "fair" to "good" standard. Given the nature of motor learning experiments, these findings are reasonable. It is difficult, if not impossible, to blind assessors, therapists, and participants (when they are familiar with low and high CI practices) to the condition in which they are allocated. In order to enhance the overall methodological quality of future studies, more effort should be put into addressing the other risk of bias items (i.e., PEDro items 1–4 and 8–11) in future studies. Furthermore, the earliest papers included in this systematic review were published in the 1990s, while the Standardized Reporting of Trials (SORT) statement, including the checklists for interventional trials, was published in 1994. This may further explain the fair score in some of the previous studies.

Methodological Considerations and Recommendations

After carefully screening the results section of each study, 36 studies were excluded as they did not report precise results (e.g., mean and SD or a related statistical measure) at one or more time points of the test session, making it impossible to compute Δ change and SDM, or they failed to ensure an equivalent baseline between control and intervention groups. Although such exclusion is necessary, it may impact the overall conclusions due to missing data from these studies, which could offer valuable insights into certain aspects of CI. This necessitates cautious interpretation of the current results, while also highlighting the importance of proper result reporting and ensuring an equal baseline between groups in future motor learning research.

Motor control research is characterized by a variety of measured performance outcomes. Different skills and sports movements can be tested within the same study (e.g., service, set, and pass) using different measurements (e.g., accuracy and error) that can have inverse interpretations for their changes (an increase in terms of accuracy score is beneficial, but a similar increase in terms of the number of errors is not). Including multiple and varied data points from the same studies in a meta-analysis is challenging and can induce several dependency issues. Therefore, meta-analyses conducted in the field of motor learning are recommended to address this dependency issue when running the meta-analysis while also controlling the effect direction when computing the ES. A possible way to do this is to consider the computation of a corrected ES accounting for the effect direction (see the "Calculation of effect size and variance" subsection for more detail) and the employment of multilevel meta-analysis and multivariate subgroup and meta-regression analyses, a methodology that allows for appropriately handling the hierarchical data structure (Assink & Wibbelink, 2016; Jukic et al., 2023) (see the "Statistical Synthesis of Effect Sizes" and the "Statistical Heterogeneity and Moderators Analysis" subsection for more detail).

Additionally, it is noteworthy to highlight the need for conducting outlier detection to exclude performance outcomes with extreme ES from the analysis (see the

"outlier detection" subsection for more detail), as such extreme values may affect the overall results. Furthermore, robust bias correction techniques are also warranted in such meta-analyses. For instance, in multilevel meta-analysis, the multilevel model of the Egger's test can be used, along with the Begg's rank correlation test, as robust model to detect publication bias (Fernández-Castilla et al., 2021). Alternatively, as recommended by Mckay et al. (2023), the RoBMA-PSMA method can be used to offer a clearer indication of preferred method while accounting for several forms of reporting bias. Conducting a *z*-curve is also recommended to estimate the power of the included studies and better understand the presence of publication bias (Mckay et al., 2023). In this context, it is important to note that the reported average power of the studies included in the present meta-analysis might sometimes be overlooked. Therefore, it should be interpreted with caution. Understanding this metric is crucial as it can guide the sample size of future studies in this field, thereby improving their power, quality, and replicability.

Regarding future original research in this field, investigators should aim for high study quality by fulfilling as many of the PEDro scale's items as possible. For instance, studies involving inexperienced subjects and assessors unfamiliar with the different motor learning models can satisfy all 11 items. However, in studies involving subjects and assessors with prior knowledge of models such as low and high CI, satisfying the blinding related PEDro items (items 5–7) may be challenging and even impossible. In such cases, satisfying the remaining items to achieve a score between 6 and 8 is crucial. Achieving this score requires meeting at least 6 criteria from items 2–4 and 8–11. Criteria 2–4 are met by employing random and concealed allocations and ensuring baseline comparability. Criteria 8–11 involve analyzing key outcomes for more than 85% of initially allocated subjects, using statistical comparisons, and reporting effect sizes and variability measures such as standard deviations or confidence intervals. Meeting these criteria enhances the study's quality and internal validity, facilitating its inclusion in future meta-analyses.

Although the PEDro scale proved effective in assessing studies, there is a need for a scale specifically designed for motor learning. This scale would address the aforementioned blinding issue and account for a detailed description of the training protocol and tested skill (e.g., mention of the dominant components: visual, kinesthetic, and/or proprioceptive, the degree of variability, and the noise level) as well as a detailed description of the study's subjects' characteristics (e.g., detailed experience levels, including years of experience). Such a scale could help identify other potential moderators of the motor learning model's effects.

Strengths, Limitations, and Implications for Future Research

The current meta-analysis is the first to attempt to validate the CI effects in a PE and/or sports context by focusing on the acute and delayed skilled performance's gains generalized by each amount of CI. The main findings indicated extremely limited support for the validity of the CI phenomenon in PE/sports settings as no overall significant difference between low and high CI practice was found in either the acquisition or the relatively permanent gains.

Nonetheless, the high level of heterogeneity in the included studies due to the disparate participants' characteristics, motor skills, and intervention protocols highlights the need for careful interpretation of the current findings and high-quality follow-up research considering more important covariates such as the "noise amount" of the training intervention (Schöllhorn et al., 2006, 2009) and the dominant components (i.e., visual, kinesthetic, and/or proprioceptive) of the learned skill, in combination with reliable neurophysiological measurements appropriate to sport skill assessment (e.g., EEG, HRV). The findings of subgroup analysis and meta-regression should also be interpreted with caution because of their observational nature (Spineli & Pandis, 2020).

Furthermore, future motor learning research should properly report their results' values, while using a sufficient sample size and ensuring an equal baseline in skilled performance outcomes.

Conclusions

Unlike previous meta-analyses, the present one did not focus on the comparison of discrete performances after the acquisition and retention phases, but on performance gains during these phases, with special attention to equal starting levels. The findings showed no overall significant CI effects for performance gains during these phases. Importantly, the CI phenomenon (i.e., acquisition gain in favor of low CI and relatively permanent gain in favor of high CI) was not fully confirmed in any of the investigated subgroups. Thus, the often-made generalization of CI effects on sports movements is not supported by our findings. Numerous inconsistent findings call for a reconsideration of previous explanatory approaches. This concerns to a large extent the inadmissible generalization of the working memory concept, which has so far primarily been based on cognitive learning and insufficiently differentiates motor learning with a high proprioceptive, kinaesthetic, and tactile component (Baddeley, 2018). Ambiguities also exist regarding the influence of the number of different exercises, as well as their relative similarities (Schöllhorn, 2016). Similarly, there are deficits with regard to the criteria for quantifying performance progress. So far, these have focused exclusively on discrete outcomes rather than on the quality of movement execution. As such, we recommend a change of direction toward the newly emerging substantive issues to further progress our understanding in motor learning, when the idea of CI learning originally inspired many researchers, but has since stalled.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10648-024-09892-z.

Author Contribution Conceptualization: AA, KT, WIS; methodology: AA, KT, WIS; investigation: AA, KT, MAB, WIS; formal analysis, AA, KT, MAB, AS, WIS; writing original draft: AA, KT, WIS; writing-review and editing: AA, KT, MAB, AS, OB, JMG, PZ, HAJ, HC, WIS; validation: AA, KT, WIS; supervision: WIS.

Funding Open Access funding enabled and organized by Projekt DEAL.

Data Availability The datasets generated and analyzed during the current meta-analysis are available from the corresponding author upon reasonable request.

Declarations

Competing Interests The authors declare no competing interests.

References

Adams, J. A. (1987). Historical review and appraisal of research on the learning, retention, and transfer of human motor skills. *Psychological Bulletin, 101*(1), 41.

Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods, 16*(2), 270–301. https://doi.org/10.1177/1094428112470848

*Aiken, C. A., & Genter, A. M. (2018). The effects of blocked and random practice on the learning of three variations of the golf chip shot. *International Journal of Performance Analysis in Sport, 18*(2), 339–349. https://doi.org/10.1080/24748668.2018.1475199

Albaret, J. M., & Thon, B. (1998). Differential effects of task complexity on contextual interference in a drawing task. *Acta Psychologica, 100*(1–2), 9–24. https://doi.org/10.1016/S0001-6918(98)00022-5

Amico, G., & Schaefer, S. (2022). Tennis expertise reduces costs in cognition but not in motor skills in a cognitive-motor dual-task condition. *Acta Pathologica, Microbiologica, Et Immunologica Scandinavica, 223*, 103503. https://doi.org/10.1016/j.actpsy.2022.103503

Ammar, A., Trabelsi, K., Boukhris, O., Bouaziz, B., Müller, P., Glenn, J. M., Chamari, K., Müller, N., Chtourou, H., Driss, T., & Hökelmann, A. (2020). Moderators of the impact of (poly) phenols interventions on psychomotor functions and bdnf: Insights from subgroup analysis and meta-regression. *Nutrients, 12*(9), 2872. https://doi.org/10.3390/nu12092872

Ammar, A., Trabelsi, K., Müller, P., Bouaziz, B., Boukhris, O., Glenn, J. M., Bott, N., Driss, T., Chtourou, H., Müller, N., & Hökelmann, A. (2020). The effect of (poly) phenol-rich interventions on cognitive functions and neuroprotective measures in healthy aging adults: A systematic review and meta-analysis. *Journal of Clinical Medicine, 9*(3), 835. https://doi.org/10.3390/jcm9030835

Ammar, A., Trabelsi, K., Boujelbane, M. A., Boukhris, O., Glenn, J. M., Chtourou, H., & Schöllhorn, W. I. (2023). The myth of contextual interference learning benefit in sports practice: A systematic review and meta-analysis. *Educational Research Review, 39*, 100537. https://doi.org/10.1016/j.edurev.2023.100537

Ammar, A., Boujelbane, M. A., Simak, M. L., Fraile-Fuente, I., Rizzi, N., Washif, J. A., Zmijewski, P., Jahrami, H., & Schöllhorn, W. I. (2024). Unveiling the acute neurophysiological responses to strength training: An exploratory study on novices performing weightlifting bouts with different motor learning models. *Biology of Sport, 41*(2), 249–274. https://doi.org/10.5114/biolsport.2024.133481

Ammar, A., Trabelsi, K., Boukhris, O., Bouaziz, B., Müller, P., M Glenn, J., Bott, N. T., Müller, N., Chtourou, H., Driss, T., & Hökelmann, A. (2020b). Effects of polyphenol-rich interventions on cognition and brain health in healthy young and middle-aged adults: Systematic review and meta-analysis. *Journal of Clinical Medicine, 9*(5), 1598.https://doi.org/10.3390/jcm9051598

Apidogo, J. B., Burdack, J., & Schöllhorn, W. I. (2021). Repetition without repetition or differential learning of multiple techniques in volleyball? *International Journal of Environmental Research and Public Health, 18*(19), 10499. https://doi.org/10.3390/ijerph181910499

Apidogo, J. B., Ammar, A., Salem, A., Burdack, J., & Schöllhorn, W. I. (2024). Resonance effects in variable practice for handball, basketball, and volleyball skills: A study on contextual interference and differential learning. *Sports, 12*(1), 5. https://doi.org/10.3390/sports12010005

Assink, M., & Wibbelink, C. J. M. (2016). Fitting three-level meta-analytic models in R: A step by-step tutorial. *The Quantitative Methods for Psychology, 12*(3), 154–174. https://doi.org/10.20982/tqmp.12.3.p154

Baddeley, A. (1992). Working memory. *Science, 255*(5044), 556–559. https://doi.org/10.1126/science.1736359

Baddeley, A. D., & Hitch, G. (1974). Working Memory. In: Bower, G.H., Ed., The psychology of learning and motivation: Advances in research and theory. *Academic Press, New York*, 47–89.https://doi.org/10.1016/S0079-7421(08)60452-1

Baddeley, A. D. (2018). Exploring working memory: Selected works of Alan Baddeley.

Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature, 533*(7604), 452–454. https://doi.org/10.1038/533452a

Barreiros, J., Figueiredo, T., & Godinho, M. (2007). The contextual interference effect in applied settings. *European Physical Education Review, 13*(2), 195–208. https://doi.org/10.1177/1356336X07076876

Bartoš, F., & Schimmack, U. (2022). Z-curve 2.0: Estimating replication rates and discovery rates. *Meta-Psychology, 6*. https://doi.org/10.15626/MP.2021.2720

Battig, W. F. (1966). Facilitation and interference. In Acquisition of Skill; Bilodeau, E.A., Ed.; *Academic Press: New York, NY, USA*, pp. 215–244.

Battig, W. F. (1972). Intratask interference as a source of facilitation in transfer and retention. *Topics in Learning and Performance*, 131–159.

Battig, W. F. (1979). The flexibility of human memory. *Levels of Processing in Human Memory*, 23–44.

Baxter, L. C., Saykin, A. J., Flashman, L. A., Johnson, S. C., Guerin, S. J., Babcock, D. R., & Wishart, H. A. (2003). Sex differences in semantic language processing: A functional MRI study. *Brain and Language, 84*(2), 264–272. https://doi.org/10.1016/s0093-934x(02)00549-7

*Bertollo, M., Berchicci, M., Carraro, A., Comani, S., & Robazza, C. (2010). Blocked and random practice organization in the learning of rhythmic dance step sequences. *Perceptual and Motor Skills, 110*(1), 77–84.https://doi.org/10.2466/PMS.110.1.77-84

Binding, L. P., Dasgupta, D., Giampiccolo, D., Duncan, J. S., & Vos, S. B. (2022). Structure and function of language networks in temporal lobe epilepsy. *Epilepsia, 63*(5), 1025–1040. https://doi.org/10.1111/epi.17204

*Bortoli, L., Spagolla, G., & Robazza, C. (2001). Variability effects on retention of a motor skill in elementary school children. *Perceptual and Motor Skills, 93*(1), 51–63.https://doi.org/10.2466/pms.2001.93.1.51

*Boyce, B. A., & Del Rey, P. (1990). Designing applied research in a naturalistic setting using a contextual interference paradigm. *Journal of Human Movement Studies, 18*, 189–200

Brady, F. (1998). A theoretical and empirical review of the contextual interference effect and the learning of motor skills. *Quest, 50*(3), 266–293. https://doi.org/10.1080/00336297.1998.10484285

Brady, F. (2004). Contextual interference: A meta-analytic study. *Perceptual and Motor Skills, 99*(1), 116–126. https://doi.org/10.2466/PMS.99.4.116-126

Brady, F. (2008). The contextual interference effect and sport skills. *Perceptual and Motor Skills, 106*(2), 461–472. https://doi.org/10.2466/PMS.106.2.461-472

Broadbent, D. P., Causer, J., Williams, A. M., & Ford, P. R. (2017). The role of error processing in the contextual interference effect during the training of perceptual-cognitive skills. *Journal of Experimental Psychology: Human Perception and Performance, 43*(7), 1329–1342. https://doi.org/10.1037/xhp0000375

*Broadbent, D. P., Causer, J., Ford, P. R., & Williams, A. M. (2015). Contextual interference effect on perceptual–cognitive skills training. *Medicine and Science in Sports and Exercise, 47*(6), 1243–1250.https://doi.org/10.1249/MSS.0000000000000530

Buekers, M. J. A., Magill, R. A., & Hall, K. G. (1992). The effect of erroneous knowledge of results on skill acquisition when augmented information is redundant. *The Quarterly Journal of Experimental Psychology Section A, 44*(1), 105–117. https://doi.org/10.1080/14640749208401285

*Buszard, T., Reid, M., Krause, L., Kovalchik, S., & Farrow, D. (2017). Quantifying contextual interference and its effect on skill transfer in skilled youth tennis players. *Frontiers in Psychology, 8*, 1931.https://doi.org/10.3389/fpsyg.2017.01931

Cashin, A. G., & McAuley, J. H. (2020). Clinimetrics: Physiotherapy evidence database (PEDro) scale. *Journal of Physiotherapy, 66*(1), 59. https://doi.org/10.1016/j.jphys.2019.08.005

*Cheong, J., Lay, B., Grove, J. R., Medic, N., & Razman, R. (2012). Practicing field hockey skills along the contextual interference continuum: A comparison of five practice schedules. *Journal of Sports Science and Medicine, 11*(2), 304–311.

*Cheong, J. P., Lay, B., & Razman, R. (2016). Investigating the contextual interference effect using combination sports skills in open and closed skill environments. *Journal of Sports Science and Medicine, 15*(1), 167

Cheung, M. W. (2014). Modeling dependent effect sizes with three-level meta-analyses: A structural equation modeling approach. *Psychological Methods, 19*(2), 211–229. https://doi.org/10.1037/a0032968

Cheung, M. W. (2019). A guide to conducting a meta-analysis with non-independent effect sizes. *Neuropsychology Review, 29*(4), 387–396. https://doi.org/10.1007/s11065-019-09415-6

Clements, A. M., Rimrodt, S. L., Abel, J. R., Blankner, J. G., Mostofsky, S. H., Pekar, J. J., Denckla, M. B., & Cutting, L. E. (2006). Sex differences in cerebral laterality of language and visuospatial processing. *Brain and Language, 98*(2), 150–158. https://doi.org/10.1016/j.bandl.2006.04.007

Cross, E. S., Schmitt, P. J., & Grafton, S. T. (2007). Neural substrates of contextual interference during motor learning support a model of active preparation. *Journal of Cognitive Neuroscience, 19*(11), 1854–1871. https://doi.org/10.1162/jocn.2007.19.11.1854

Cutsuridis, V., & Yoshida, M. (2017). Editorial: Memory processes in medial temporal lobe: Experimental, theoretical and computational approaches. *Frontiers in Systems Neuroscience, 11*, 19. https://doi.org/10.3389/fnsys.2017.00019

Deeks, J. J., Higgins, J. P., & Altman, D. G. (2019). Group obotCSM. Analysing data and undertaking meta-analyses. C*ochrane Handbook for Systematic Reviews of Interventions,* pp. 241–84.

Esposito, G., Vivanti, G. (2013). Gross motor skills. In: Volkmar, F.R. (eds) Encyclopedia of Autism Spectrum Disorders. *Springer, New York, NY.* https://doi.org/10.1007/978-1-4419-1698-3_179

Farrow, D., & Maschette, W. (1997). The effects of contextual interference on children learning forehand tennis groundstrokes. *Journal of Human Movement Studies, 33*, 47–67.

*Fazeli, D., Taheri, H., & Saberi Kakhki, A. (2017). Random versus blocked practice to enhance mental representation in golf putting. *Perceptual and Motor Skills, 124*(3), 674-688.https://doi.org/10.1177/0031512517704106

Fernández-Castilla, B., Declercq, L., Jamshidi, L., Beretvas, S. N., Onghena, P., & Van den Noortgate, W. (2021). Detecting selection bias in meta-analyses with multiple outcomes: A simulation study. *Journal of Experimental Education, 89*(1), 125–144. https://doi.org/10.1080/00220973.2019.1582470

*French, K. E., Rink, J. E., & Werner, P. H. (1990). Effects of contextual interference on retention of three volleyball skills. *Perceptual and motor skills, 71*(1), 179–186.https://doi.org/10.2466/PMS.71.4.179-186

Frömer, R., Stürmer, B., & Sommer, W. (2016). (Don't) Mind the effort: Effects of contextual interference on ERP indicators of motor preparation. *Psychophysiology, 53*(10), 1577–1586. https://doi.org/10.1111/psyp.12703

Gallicchio, G., & Ring, C. (2019). Don't look, don't think, just do it! Toward an understanding of alpha gating in a discrete aiming task. *Psychophysiology, 56*(3), 14. https://doi.org/10.1111/psyp.13298

Gebkenjans, F., Beckmann, H., & Schöllhorn, W. I. (2007). Does low and high contextual interference lead to different levels of noise?. In 3rd European Workshop on Movement Sciences. Book of Abstract, eds P. Beek and R. van den Langenberg (Köln: Sportverlag Strauß) (pp. 153–154).

Gentile, A. M. (1998). Movement science: Implicit and explicit processes during acquisition of functional skills. *Scandinavian Journal of Occupational Therapy, 5*(1), 7–16. https://doi.org/10.3109/11038129809035723

*Goodwin, J. E., & Meeuwsen, H. J. (1996). Investigation of the contextual interference effect in the manipulation of the motor parameter of over-all force. *Perceptual and Motor Skills, 83*(3), 735–743.https://doi.org/10.2466/pms.1996.83.3.735

Graser, J. V., Bastiaenen, C. H., & van Hedel, H. J. (2019). The role of the practice order: A systematic review about contextual interference in children. *PLoS ONE, 14*(1), e0209979. https://doi.org/10.1371/journal.pone.0209979

Gucciardi, D. F., Lines, R. L., & Ntoumanis, N. (2022). Handling effect size dependency in meta-analysis. *International Review of Sport and Exercise Psychology, 15*(1), 152–178. https://doi.org/10.1080/1750984X.2021.1946835

Gur, R. C., Alsop, D., Glahn, D., Petty, R., Swanson, C. L., Maldjian, J. A., Turetsky, B. I., Detre, J. A., Gee, J., & Gur, R. E. (2000). An fMRI study of sex differences in regional activation to a verbal and a spatial task. *Brain and Language, 74*(2), 157–170. https://doi.org/10.1006/brln.2000.2325

Hagger, M. (2022). Meta-analysis. *International Review of Sport and Exercise Psychology, 15*(1), 120–151. https://doi.org/10.1080/1750984X.2021.1966824

Hajihosseini, S. (2016). The effect of moderate contextual interference on motor-skill learning. *Biology and Medicine, 8*(7), 1. https://doi.org/10.4172/0974-8369.1000339

*Hall, K. G., Domingues, D. A., & Cavazos, R. (1994). Contextual interference effects with skilled baseball players. *Perceptual and Motor Skills, 78*(3), 835–841.https://doi.org/10.1177/003151259407800331

Hardwick, R. M., Rottschy, C., Miall, R. C., & Eickhoff, S. B. (2013). A quantitative meta-analysis and review of motor learning in the human brain. *NeuroImage, 67*, 283–297. https://doi.org/10.1016/j.neuroimage.2012.11.020

Harness, A., Jacot, L., Scherf, S., White, A., & Warnick, J. E. (2008). Sex differences in working memory. *Psychological Reports, 103*(1), 214–218. https://doi.org/10.2466/pr0.103.1.214-218

*Hebert, E. P., Landin, D., & Solmon, M. A. (1996). Practice schedule effects on the performance and learning of low-and high-skilled students: An applied study. *Research Quarterly for Exercise and Sport, 67*(1), 52–58.https://doi.org/10.1080/02701367.1996.10607925

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6*(2), 107–128. https://doi.org/10.2307/1164588

Henz, D., John, A., Merz, C., & Schöllhorn, W. I. (2018). Post-task effects on EEG brain activity differ for various differential learning and contextual interference protocols. *Frontiers in Human Neuroscience, 19*. https://doi.org/10.3389/fnhum.2018.00019

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

Higgins, J., & Green, S. (2011). Cochrane handbook for systematic reviews of interventions. *Wiley Blackwell.*

Hill, A. C., Laird, A. R., & Robinson, J. L. (2014). Gender differences in working memory networks: A BrainMap meta-analysis. *Biological Psychology, 102*, 18–29. https://doi.org/10.1016/j.biopsycho.2014.06.008

Hume, D. (1740/2007). A treatise of human nature: Being an attempt to introduce the experimental method of reasoning into moral subjects. (Norton Norton Edition). *Oxford*.

Jacoby, N., & Ahissar, M. (2013). What does it take to show that a cognitive training procedure is useful? A critical evaluation. *Progress in Brain Research, 207*, 121–140. https://doi.org/10.1016/B978-0-444-63327-9.00004-7

*Jeon, M. J., Jeon, H. S., Yi, C. H., Kwon, O. Y., You, S. H., & Park, J. H. (2021). Block and random practice: A Wii Fit dynamic balance training in older adults. *Research Quarterly for Exercise and Sport, 92*(3), 352–360.https://doi.org/10.1080/02701367.2020.1733456

*Jones, L. L., & French, K. E. (2007). Effects of contextual interference on acquisition and retention of three volleyball skills. *Perceptual and Motor Skills, 105*(3), 883–890.https://doi.org/10.2466/pms.105.3.883-890

Jukic, I., Castilla, A. P., Ramos, A. G., Van Hooren, B., McGuigan, M. R., & Helms, E. R. (2023). The acute and chronic effects of implementing velocity loss thresholds during resistance training: A systematic review, meta-analysis, and critical evaluation of the literature. *Sports medicine (Auckland, N.Z.), 53*(1), 177–214. https://doi.org/10.1007/s40279-022-01754-4

Kalkhoran, J. F., & Shariati, A. (2015). The effects of contextual interference on learning volleyball motor skills. *Journal of Sport Sciences, 6*, 12–24. https://doi.org/10.7752/jpes.2012.04081

*Keller, G. J., Li, Y., Weiss, L. W., & Relyea, G. E. (2006). Contextual interference effect on acquisition and retention of pistol-shooting skills. *Perceptual and Motor Skills, 103*(1), 241–252.https://doi.org/10.2466/pms.103.1.241-252

Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., Dahlström, K., Gillberg, C. G., Forssberg, H., & Westerberg, H. (2005). Computerized training of working memory in children with ADHD–A randomized, controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry, 44*(2), 177–186. https://doi.org/10.1097/00004583-200502000-00010

*Krause, D., Brüne, A., Fritz, S., Kramer, P., Meisterjahn, P., Schneider, M., & Sperber, A. (2014). Learning of a golf putting task with varying contextual interference levels induced by feedback schedule in novices and experts. *Perceptual and Motor Skills, 118*(2), 384–399.https://doi.org/10.2466/23.30.PMS.118k17w3

Lage, G. M., Ugrinowitsch, H., Apolinário-Souza, T., Vieira, M. M., Albuquerque, M. R., & Benda, R. N. (2015). Repetition and variation in motor practice: A review of neural correlates. *Neuroscience and Biobehavioral Reviews, 57*, 132–141. https://doi.org/10.1016/j.neubiorev.2015.08.012

Lage, G. M., Faria, L. O., Ambrósio, N. F., Borges, A. M., & Apolinário-Souza, T. (2022). What is the level of contextual interference in serial practice? A meta-analytic review. *Journal of Motor Learning and Development, 10*(1), 224–242. https://doi.org/10.1123/jmld.2021-0020

Lee, T. D., & Magill, R. A. (1985). Can forgetting facilitate skill acquisition?. *In Advances in Psychology (Vol. 27, pp. 3–22)*. North-Holland. https://doi.org/10.1016/S0166-4115(08)62528-5

Lee, T. D., & Magill, R. A. (1983). The locus of contextual interference in motor-skill acquisition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9*(4), 730–746. https://doi.org/10.1037/0278-7393.9.4.730

Lin, C. H., Knowlton, B. J., Chiang, M. C., Iacoboni, M., Udompholkul, P., & Wu, A. D. (2011). Brain-behavior correlates of optimizing learning through interleaved practice. *NeuroImage, 56*(3), 1758–1772. https://doi.org/10.1016/j.neuroimage.2011.02.066

Lin, C. H., Chiang, M. C., Wu, A. D., Iacoboni, M., Udompholkul, P., Yazdanshenas, O., & Knowlton, B. J. (2012). Age related differences in the neural substrates of motor sequence learning after interleaved and repetitive practice. *NeuroImage, 62*(3), 2007–2020. https://doi.org/10.1016/j.neuroimage.2012.05.015

Lin, C. H., Knowlton, B. J., Wu, A. D., Iacoboni, M., Yang, H. C., Ye, Y. L., Liu, K. H., & Chiang, M. C. (2016). Benefit of interleaved practice of motor skills is associated with changes in functional brain network topology that differ between younger and older adults. *Neurobiology of Aging, 42*, 189–198. https://doi.org/10.1016/j.neurobiolaging.2016.03.010

Magill, R. A., & Hall, K. G. (1990). A review of the contextual interference effect in motor skill acquisition. *Human Movement Science, 9*(3–5), 241–289. https://doi.org/10.1016/0167-9457(90)90005-X

Maher, C. G., Sherrington, C., Herbert, R. D., Moseley, A. M., & Elkins, M. (2003). Reliability of the PEDro scale for rating quality of randomized controlled trials. *Physical Therapy, 83*(8), 713–721.

Manohar, S. G., Pertzov, Y., & Husain, M. (2017). Short-term memory for spatial, sequential and duration information. *Current Opinion in Behavioral Sciences, 17*, 20–26. https://doi.org/10.1016/j.cobeha.2017.05.023

McKay, B., Bacelar, M. F., Parma, J. O., Miller, M. W., & Carter, M. J. (2023). The combination of reporting bias and underpowered study designs has substantially exaggerated the motor learning benefits of self-controlled practice and enhanced expectancies: A meta-analysis. *International Review of Sport and Exercise Psychology*, 1–21. https://doi.org/10.1080/1750984X.2023.2207255

Mechsner, F., Kerzel, D., Knoblich, G., & Prinz, W. (2001). Perceptual basis of bimanual coordination. *Nature, 414*(6859), 69–73. https://doi.org/10.1038/35102060

*Meira, C. M., & Tani, G. (2003). Contextual interference effects assessed by extended transfer trials in the acquisition of the volleyball serve. *Journal of Human Movement Studies, 45*(5), 449–468.

*Memmert, D. (2006). Long-term effects of type of practice on the learning and transfer of a complex motor skill. *Perceptual and Motor Skills, 103*(3), 912–916.https://doi.org/10.2466/pms.103.3.912-916

*Menayo, R., Sabido, R., Fuentes, J. P., Moreno, F. J., & Garcia, J. A. (2010). Simultaneous treatment effects in learning four tennis shots in contextual interference conditions. *Perceptual and Motor Skills, 110*(2), 661–673https://doi.org/10.2466/PMS.110.2.661-673

Merbah, S., & Meulemans, T. (2011). Learning a motor skill: Effects of blocked versus random practice: A review. *Psychologica Belgica, 51*, 15–48. https://doi.org/10.5334/pb-51-1-15

Meyer, P., Mecklinger, A., Grunwald, T., Fell, J., Elger, C. E., & Friederici, A. D. (2005). Language processing within the human medial temporal lobe. *Hippocampus, 15*(4), 451–459. https://doi.org/10.1002/hipo.20070

Miller, G. A. (1956). The magical number seven plus minus two: Some limits on our capacity for processing information. *Psychological Review, 29*, 106–112.

Morris, S. B. (2008). Estimating effect sizes from pretest-posttest-control group designs. *Organizational Research Methods, 11*(2), 364–386. https://doi.org/10.1177/1094428106291059

*Naimo, M. A., Zourdos, M. C., Wilson, J. M., Kim, J. S., Ward, E. G., Eccles, D. W., & Panton, L. B. (2013). Contextual interference effects on the acquisition of skill and strength of the bench press. *Human Movement Science, 32*(3), 472–484.https://doi.org/10.1016/j.humov.2013.02.002

*North, J. S., Bezodis, N. E., Murphy, C. P., Runswick, O. R., Pocock, C., & Roca, A. (2019). The effect of consistent and varied follow-through practice schedules on learning a table tennis backhand. *Journal of Sports Sciences, 37*(6), 613–620.https://doi.org/10.1080/02640414.2018.1522683

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *Systematic Reviews, 10*(1), 1–11. https://doi.org/10.1186/s13643-021-01626-4

Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., et al. (2021a). PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *British Medical Journal, 372*. https://doi.org/10.1136/bmj.n160

*Pasand, F., Fooladiyanzadeh, H., & Nazemzadegan, G. (2016). The effect of gradual increase in contextual interference on acquisition, retention and transfer of volleyball skills. *International Journal of Kinesiology and Sports Science, 4*(2), 72-77.https://doi.org/10.7575/aiac.ijkss.v.4n.2p.72

Patel, A., Biso, G. M. N. R., & Fowler, J. B. (2023). Neuroanatomy, Temporal lobe. *In StatPearls. StatPearls Publishing.* Available from: https://www.ncbi.nlm.nih.gov/books/NBK519512/. Accessed 1 Sept 2023.

Penhune, V. B., & Steele, C. J. (2012). Parallel contributions of cerebellar, striatal and M1 mechanisms to motor sequence learning. *Behavioural Brain Research, 226*(2), 579–591. https://doi.org/10.1016/j.bbr.2011.09.044

*Pollatou, E., Kioumourtzoglou, E., Agelousis, N., & Mavromatis, G. (1997). Contextual interference effects in learning novel motor skills. *Perceptual and Motor Skills, 84*(2), 487–496.https://doi.org/10.2466/pms.1997.84.2.487

Pollock, B. J., & Lee, T. D. (1997). Dissociated contextual interference effects in children and adults. *Perceptual and Motor Skills, 84*(3), 851–858. https://doi.org/10.2466/pms.1997.84.3.851

*Porter, C., Greenwood, D., Panchuk, D., & Pepping, G. J. (2020). Learner-adapted practice promotes skill transfer in unskilled adults learning the basketball set shot. *European Journal of Sport Science, 20*(1), 61–71.https://doi.org/10.1080/17461391.2019.1611931

Ramírez-Ruiz, J., Grytskyy, D., & Moreno-Bote, R. (2022). Seeking entropy: Complex behavior from intrinsic motivation to occupy action-state path space. *arXiv preprint* arXiv:2205.10316. Accessed 1 Sept 2023.

*Rendell, M. A., Masters, R. S., Farrow, D., & Morris, T. (2011). An implicit basis for the retention benefits of random practice. *Journal of Motor Behavior, 43*(1), 1–13.https://doi.org/10.1080/00222895.2010.530304

Reynolds, M. R., Niileksela, C. R., Gignac, G. E., & Sevillano, C. N. (2022). Working memory capacity development through childhood: A longitudinal analysis. *Developmental Psychology, 58*(7), 1254–1263. https://doi.org/10.1037/dev0001360

Rico-González, M., Pino-Ortega, J., Clemente, F., & Los Arcos, A. (2022). Guidelines for performing systematic reviews in sports science. *Biology of Sport, 39*(2). https://doi.org/10.5114/biolsport.2022.106386

*Sadri, K., Mohommadzadeh, H., & Khani, M. (2013). The effect of contextual interference on acquisition and learning badminton skills among children aged from 10 to 12. *Annals of Applied Sport Science, 1*(3), 39–46.

Saylik, R., Raman, E., & Szameitat, A. J. (2018). Sex differences in emotion recognition and working memory tasks. *Frontiers in Psychology, 9*, 1072. https://doi.org/10.3389/fpsyg.2018.01072

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

Schmidt, R. A., Lee, T. D., Winstein, C. J., Wulf, G., & Zelaznik, H. N. (2011). *Motor control and learning: A behavioral emphasis*. Human Kinetics.

Schmidt, R., & Lee, T. (2019). Motor learning and performance 6th edition with web study guide-loose-leaf edition: From principles to application. *Human Kinetics Publishers*.

Schöllhorn, W. I. (2000) Practical concequences of systems dynamic approach to technique and strength training. *Acta Academiae Olympique Estoniat*, *8*, 25–37.

Schöllhorn, W. I. (2016). Invited commentary: Differential learning is different from contextual interference learning. *Human Movement Science, 47*, 240–245. https://doi.org/10.1016/j.humov.2015.11.018

Schöllhorn, W. I., Beckmann, H., Michelbrink, M., Sechelmann, M., Trockel, M., & Davids, K. (2006). Does noise provide a basis for the unification of motor learning theories? *International Journal of Sport Psychology, 37*(2/3), 186.

Schöllhorn, W. I., Mayer-Kress, G., Newell, K. M., & Michelbrink, M. (2009). Time scales of adaptive behavior and motor learning in the presence of stochastic perturbations. *Human Movement Science, 28*(3), 319–333. https://doi.org/10.1016/j.humov.2008.10.005

Schöllhorn, W. I., Hegen, P., & Eekhoff, A. (2014). Differenzielles Lernen und andere motorische Lerntheorien. *Spectrum Der Sportwissenschaft, 2*, 35–55.

Schöllhorn, W. I., Eekhoff, A., & Hegen, P. (2015). System dynamics and differential learning. *Sportwissenschaft, 45*, 127–137. https://doi.org/10.1007/s12662-015-0366-z

Schöllhorn, W. I., Rizzi, N., Slapšinskaitė-Dackevičienė, A., & Leite, N. (2022). Always pay attention to which model of motor learning you are using. *International Journal of Environmental Research and Public Health, 19*(2), 711. https://doi.org/10.3390/ijerph19020711

Schroeder, S. R., & Marian, V. (2012). A bilingual advantage for episodic memory in older adults. *Journal of Cognitive Psychology (hove, England), 24*(5), 591–601. https://doi.org/10.1080/20445911.2012.669367

*Sharp, M. H., Gheith, R. H., Reber, D. A., Stefan, M. W., LoDuca, S., Lowery, R. P., & Wilson, J. M. (2020). The effect of blocked versus random practice on dominant and non-dominant baseball swing. *Journal of Sport and Human Performance, 8*(1). https://doi.org/10.12922/jshp.v8i1.163

Shaywitz, B. A., Shaywitz, S. E., Pugh, K. R., Constable, R. T., Skudlarski, P., Fulbright, R. K., Bronen, R. A., Fletcher, J. M., Shankweiler, D. P., & Katz, L. (1995). Sex differences in the functional organization of the brain for language. *Nature, 373*(6515), 607–609. https://doi.org/10.1038/373607a0

Shea, J. B., & Morgan, R. L. (1979). Contextual interference effects on the acquisition, retention, and transfer of a motor skill. *Journal of Experimental Psychology: Human Learning and Memory, 5*(2), 179–187. https://doi.org/10.1037/0278-7393.5.2.179

Shea, J. B., & Zimny, S. T. (1983). Context effects in memory and learning movement information. *Advances in Psychology, 12*, 345–366. https://doi.org/10.1016/S0166-4115(08)61998-6

Shewokis, P. A. (1997). Is the contextual interference effect generalizable to computer games? *Perceptual and Motor Skills, 84*(1), 3–15. https://doi.org/10.2466/pms.1997.84.1.3

Silverman, S., Tyson, L. A., & Krampitz, J. (1992). Teacher feedback and achievement in physical education: Interaction with student practice. *Teaching and Teacher Education, 8*(4), 333–344. https://doi.org/10.1016/0742-051X(92)90060-G

*Smith, P. J., Gregory, S. K., & Davies, M. (2003). Alternating versus blocked practice in learning a cartwheel. *Perceptual and Motor Skills, 96*(3), 1255–1264.https://doi.org/10.2466/pms.2003.96.3c.1255

Smith, P. J., & Rudisill, M. E. (1993). The influence of proficiency level, transfer distality, and gender on the contextual interference effect. *Research Quarterly for Exercise and Sport, 64*(2), 151–157. https://doi.org/10.1080/02701367.1993.10608792

Spineli, L. M., & Pandis, N. (2020). Problems and pitfalls in subgroup analysis and meta-regression. *American Journal of Orthodontics and Dentofacial Orthopedics : Official Publication of the American Association of Orthodontists, Its Constituent Societies, and the American Board of Orthodontics, 158*(6), 901–904. https://doi.org/10.1016/j.ajodo.2020.09.001

Stevens, J. P. (1984). Outliers and influential data points in regression analysis. *Psychological Bulletin, 95*(2), 334–344. https://doi.org/10.1037/0033-2909.95.2.334

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*(2), 257–285.

Sweller, J. (2023). The development of cognitive load theory: Replication crises and incorporation of other theories can lead to theory expansion. *Educational Psychology Review, 35*, 95. https://doi.org/10.1007/s10648-023-09817-2

Tang, H., Qi, X. L., Riley, M. R., & Constantinidis, C. (2019). Working memory capacity is enhanced by distributed prefrontal activation and invariant temporal dynamics. *Proceedings of the National Academy of Sciences of the United States of America, 116*(14), 7095–7100. https://doi.org/10.1073/pnas.1817278116

*Travlos, A. K. (2010). Specificity and variability of practice, and contextual interference in acquisition and transfer of an underhand volleyball serve. *Perceptual and Motor Skills, 110*(1), 298–312. https://doi.org/10.2466/PMS.110.1.298-312

Van den Noortgate, W., López-López, J. A., Marín-Martínez, F., & Sánchez-Meca, J. (2013). Three-level meta-analysis of dependent effect sizes. *Behavior Research Methods, 45*(2), 576–594. https://doi.org/10.3758/s13428-012-0261-6

*Vera, J. G., & Montilla, M. M. (2003). Practice schedule and acquisition, retention, and transfer of a throwing task in 6-yr.-old children. *Perceptual and Motor Skills, 96*(3), 1015–1024.https://doi.org/10.2466/pms.2003.96.3.1015

*Vera, J. G., Alvarez, J. C. B., & Medina, M. M. (2008). Effects of different practice conditions on acquisition, retention, and transfer of soccer skills by 9-year-old schoolchildren. *Perceptual and Motor Skills, 106*(2), 447–460.https://doi.org/10.2466/pms.106.2.447-460

Verhagen, A. P., De Vet, H. C., De Bie, R. A., Kessels, A. G., Boers, M., Bouter, L. M., & Knipschild, P. G. (1998). The Delphi list: A criteria list for quality assessment of randomized clinical trials for conducting systematic reviews developed by Delphi consensus. *Journal of Clinical Epidemiology, 51*(12), 1235–1241. https://doi.org/10.1016/s0895-4356(98)00131-0

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

Wang, X., Abdullah, B., & Samsudin, S. (2022). The effect of contextual interference on motor learning among healthy adolescents: A systematic review. *Journal of Positive School Psychology, 6*(7), 4545–4580.

Wright, D. L., Li, Y., & Whitacre, C. (1992). The contribution of elaborative processing to the contextual interference effect. *Research Quarterly for Exercise and Sport, 63*(1), 30–37. https://doi.org/10.1080/02701367.1992.10607554

Wright, D. L., Brueckner, S., Black, C. B., Magnuson, C., & Immink, M. A. (2004). Long-term motor programming improvements occur via concatenation of movement sequences during random but not during blocked practice. *Journal of Motor Behavior, 36*(1), 39–50. https://doi.org/10.3200/Jmbr.36.1.39-50

Wright, D., Verwey, W., Buchanen, J., Chen, J., Rhee, J., & Immink, M. (2016). Consolidating behavioral and neurophysiologic findings to explain the influence of contextual interference during motor sequence learning. *Psychonomic Bulletin & Review, 23*(1), 1–21. https://doi.org/10.3758/s13423-015-0887-3

*Wrisberg, C. A., & Liu, Z. (1991). The effect of contextual variety on the practice, retention, and transfer of an applied motor skill. *Research Quarterly for Exercise and Sport, 62*(4), 406–412.https://doi.org/10.1080/02701367.1991.10607541

Wu, W. F., Young, D. E., Schandler, S. L., Meir, G., Judy, R. L., Perez, J., & Cohen, M. J. (2011). Contextual interference and augmented feedback: Is there an additive effect for motor learning? *Human Movement Science, 30*(6), 1092–101. https://doi.org/10.1016/j.humov.2011.02.004

Wu, H. G., Miyamoto, Y. R., Castro, L. N. G., Ölveczky, B. P., & Smith, M. A. (2014). Temporal structure of motor variability is dynamically regulated and predicts motor learning ability. *Nature Neuroscience, 17*(2), 312–321.

Wymbs, N. F., & Grafton, S. T. (2009). Neural substrates of practice structure that support future off-line learning. *Journal of Neurophysiology, 102*(4), 2462–2476. https://doi.org/10.1152/jn.00315.2009

*Yanci, J., Reina, R., Los Arcos, A., & Cámara, J. (2013). Effects of different contextual interference training programs on straight sprinting and agility performance of primary school students. *Journal of Sports Science and Medicine, 12*(3), 601.

*Yanci, J., Los Arcos, A., Salinero, J. J., Mendiguchia, J., Gil, E., Sanesteban, D., & Grande, I. (2015). Effects of different agility training programs among first-grade elementary school students. *Collegium Antropologicum, 39*(1), 87–92.

*Zetou, E., Michalopoulou, M., Giazitzi, K., & Kioumourtzoglou, E. (2007). Contextual interference effects in learning volleyball skills. *Perceptual and Motor Skills, 104*(3), 995–1004.https://doi.org/10.2466/pms.104.3.995-1004

Zipp, G. P., & Gentile, A. M. (2010). Practice schedule and the learning of motor skills in children and adults: teaching implications. *Journal of College Teaching and Learning (TLC), 7*(2). https://doi.org/10.19030/tlc.v7i2.87

Authors and Affiliations

Achraf Ammar[1,2,3,4] · Khaled Trabelsi[4,5] · Mohamed Ali Boujelbane[1,4] · Atef Salem[1,4] · Omar Boukhris[6,7] · Jordan M. Glenn[8] · Piotr Zmijewski[9] · Haitham A. Jahrami[10,11] · Hamdi Chtourou[4,12] · Wolfgang I. Schöllhorn[1]

✉ Achraf Ammar
ammar.achraf@ymail.com; acammar@uni-mainz.de

1 Department of Training and Movement Science, Institute of Sport Science, Johannes Gutenberg-University Mainz, 55128 Mainz, Germany

2 Interdisciplinary Laboratory in Neurosciences, Physiology and Psychology: Physical Activity, Health and Learning (LINP2), UFR STAPS (Faculty of Sport Sciences), UPL, Paris Nanterre University, Nanterre, France

3 Research Laboratory, Molecular Bases of Human Pathology, LR19ES13, Faculty of Medicine of Sfax, University of Sfax, Sfax, Tunisia

4 High Institute of Sport and Physical Education, University of Sfax, Sfax, Tunisia

5 Research Laboratory, Education, Motricity, Sport, and Health (EM2S), LR15JS01, High Institute of Sport and Physical Education, University of Sfax, Sfax, Tunisia

6 SIESTA Research Group, School of Allied Health, Human Services and Sport, La Trobe University, Melbourne, VIC 3086, Australia

7 Sport, Performance, and Nutrition Research Group, School of Allied Health, Human Services and Sport, La Trobe University, Melbourne, VIC 3086, Australia

8 Department of Health, Exercise Science Research Center Human Performance and Recreation, University of Arkansas, Fayetteville, AR 72701, USA

9 Jozef Pilsudski University of Physical Education in Warsaw, 00-809 Warsaw, Poland

10 Ministry of Health, Manama, Bahrain

11 Department of Psychiatry, College of Medicine and Medical Sciences, Arabian Gulf University, Manama, Bahrain

12 Research Unit: “Physical Activity, Sport, and Health”, National Observatory of Sport, UR18JS011003 Tunis, Tunisia