



OPEN ACCESS

EDITED BY

Guy Cheron,
Université Libre de Bruxelles, Belgium

REVIEWED BY

David Sherwood,
University of Colorado Boulder, United States
Neeraj Kumar,
Indian Institute of Technology Hyderabad,
India

*CORRESPONDENCE

Stanisław H. Czyż
✉ stachu.czyz@gmail.com

RECEIVED 26 January 2024

ACCEPTED 22 July 2024

PUBLISHED 14 August 2024

CITATION

Czyż SH, Wójcik AM and Solarská P (2024)
The effect of contextual interference on
transfer in motor learning - a systematic
review and meta-analysis.
Front. Psychol. 15:1377122.
doi: 10.3389/fpsyg.2024.1377122

COPYRIGHT

© 2024 Czyż, Wójcik and Solarská. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

The effect of contextual interference on transfer in motor learning - a systematic review and meta-analysis

Stanisław H. Czyż^{1,2,3*}, Aleksandra M. Wójcik¹ and Petra Solarská²

¹Faculty of Physical Education and Sports, Wrocław University of Health and Sport Sciences, Wrocław, Poland, ²Faculty of Sport Studies, Masaryk University, Brno, Czechia, ³Physical Activity, Sport and Recreation (PhASRec), Faculty of Health Sciences, North-West University, Potchefstroom, South Africa

Since the initial study on contextual interference (CI) in 1966, research has explored how practice schedules impact retention and transfer. Apart from support from scientists and practitioners, the CI effect has also faced skepticism. Therefore, we aimed to review the existing literature on the CI effect and determine how it affects transfer in laboratory and applied settings and in different age groups. We found 1,287 articles in the following databases: Scopus, EBSCO, Web of Science, ScienceDirect, supplemented by the Google Scholar search engine and manual search. Of 300 fully screened articles, 42 studies were included in the systematic review and 34 in the quantitative analysis (meta-analysis). The overall CI effect on transfer in motor learning was medium (SMD=0.55), favoring random practice. Random practice was favored in the laboratory and applied settings. However, in laboratory studies, the medium effect size was statistically significant (SMD=0.75), whereas, in applied studies, the effect size was small and statistically non-significant (SMD=0.34). Age group analysis turned out to be significant only in adults and older adults. In both, the random practice was favored. In adults, the effect was medium (SMD=0.54), whereas in older adults was large (SMD=1.28). In young participants, the effect size was negligible (SMD=0.12).

Systematic review registration: <https://clinicaltrials.gov/>, identifier CRD42021228267.

KEYWORDS

contextual interference, practice schedule, random practice, blocked practice, transfer, motor learning

Introduction

The very first meta-analysis on contextual interference (CI) in motor learning was published in 2004 by Brady (2004). Brady asked how CI affects retention and transfer, i.e., two processes which, according to Battig (1966, 1972), are altered by CI level. Battig noted that high CI level originating from “random order,” i.e., an order (or schedule) consisting of the trials arranged randomly and rapidly changing, hindered performance and facilitated retention and transfer. On the other hand, the so-called “blocked order,” i.e., an order in which one skill

variation trials were completed before introducing the next skill, resulted in a low CI level. Low CI, unlike high CI, facilitated performance and hindered retention and transfer.

This rather unexpected finding was soon tested in motor learning, originally by [Shea and Morgan \(1979\)](#) and later by many others. In [Brady \(2004\)](#) meta-analytically studied CI effect in motor learning concluding that the general (on retention and transfer) effect size (ES) in laboratory research was medium (Cohen's $d=0.57$) and in applied research was negligible (Cohen's $d=0.19$). On the other hand, the ES in studies on transfer in adults was small, i.e., $d=0.47$ for laboratory research and $d=0.43$ for applied research. In children, the ES in applied research was negligible ($d=0.10$).

Retention and transfer, though both equally important in motor learning measure assess different constructs. Retention is a good indicator of learning, whereas transfer is a good indicator of adaptability ([Magill and Anderson, 2021](#)). One could argue that the ultimate goal of practice is to transfer skills from one context to another, i.e., generalize practiced skill to a non-practice one ([Raviv et al., 2022](#)). Consequently, transfer could be considered more reliable measure of CI effect ([Shewokis, 1997](#); [Brady, 2004](#)). It is hence crucial to analyze and discuss CI effect on transfer exclusively.

Almost 20 years have passed since Brady performed his meta-analysis. Given that the meta-analysis should be updated within 2 years ([Cumpston and Chandler, 2019](#)) and the median time for the update in Cochrane Collaboration is about 5.5 ([Shojania et al., 2007](#)), the meta-analysis on the CI effect should have been updated much earlier. Some other analyses were performed in the meantime. However, most of them focused on very specific questions, e.g., CI effect in random vs. serial practice ([Lage et al., 2021](#)), CI effect in children with cerebral palsy ([Graser et al., 2019](#)), or CI effect in students in medical and physiotherapy education ([Sattelmayer et al., 2016](#)).

Recently, in 2023, Ammar and colleagues published a meta-analysis ([Ammar et al., 2023](#)) on the CI effect on retention and transfer in sport settings. They referred to the CI effects as “a myth” because they found statistically non-significant and small effect sizes favoring random practice in transfer and retention tests (in analyses on the whole population). Unfortunately, they did not report how they treated outcomes retrieved from one sample (multiple dependent effect sizes), i.e., how they dealt with the dependency problem. Additionally, they did not screen the EBSCO databases, which include key resources like APA PsycArticles, APA PsychInfo, SPORTDiscus with Full Text, Medline, and Academic Search Complete. Furthermore, they did not preregister their study, a now common practice in research that reduces publication bias and selective reporting of outcome-related bias.

Given all of the above, we decided to perform a systematic review, and we formulated our objectives based on the ones advanced by [Brady \(2004\)](#). However, we exclusively focused on transfer due to the number of conducted analyses and included studies. The study's primary objective is to determine the overall effect size of CI on transfer in motor learning. Our secondary objectives, based on [Brady \(2004\)](#) inclusion criteria, are:

- 1) To estimate the CI effect in laboratory vs. field-based studies
- 2) To estimate the CI effect in young vs. adults vs. elderly adults
- 3) To estimate the CI effect in novice vs. experienced participants

Methods

The study was registered in PROSPERO under the number CRD42021228267. Given the number of conducted analyses and included studies as well as the page/word limit for an article, the registered review has been split into two consecutive papers: retention ([Czyż et al., 2023](#)) and transfer meta-analyses. Transfer performance was analyzed in the present study. Each consecutive step of the study was performed in accordance with the PICO guidelines ([Methley et al., 2014](#)), the PRISMA statement ([Moher et al., 2009, 2015](#); [Page and Moher, 2017](#); see [Supplementary Appendix 1](#)), and, successively *Quality Assessment Tool for Quantitative Studies* ([Thomas, 2003](#); [Thomas et al., 2004](#)).

Inclusion and exclusion criteria

The inclusion criteria were defined in accordance with PICO (*Population, Intervention, Control, Outcome*) and included:

Population: adult, young, novice, experienced. In line with Brady's inclusion criteria ([Brady, 2004](#)), only healthy participants were considered, and two variables (age and experience) were incorporated. Participants under 18 years old were labeled as *Young*. Participants aged 60 years old and more were classified as *Older Adults*, and those between 18 and 60 years old were ranked as *Adults*.

Intervention: field setting, high CI volume (random/interleaved schedule);

Control: laboratory settings, low CI volume (blocked schedule/repetitive practice);

A wide variety of motor tasks and experimental procedures were utilized in reviewed studies potentially considered for inclusion; however, it is worth mentioning that only the studies using single-task procedures were relevant for the current review.

The main classification considered for analysis was based on contextual interference volume: studies including groups with different practice schedules: blocked order (low CI), and random order (high CI) were compared.

The subgroup analysis was performed based on the age of participants: young (<18 years old), adults, and older adults (>60 years old). The subsequent subgroup analysis was the nature of the research: studies carried out in a controlled laboratory environment were matched up with studies conducted in a field setting using typical sports skills (applied research). The subsequent subgroup analysis was the experience—experienced participants vs. novices.

Outcome: transfer test results. The primary outcomes were the standardized effect sizes of CI in transfer in motor learning. The outcomes evaluating the transfer of the learned motor skill were considered selectable. Taking into account the effect of *sleep-induced consolidation of trained skills* ([Diekelmann and Born, 2010](#); [Yang et al., 2014](#)) current meta-analysis consisted of delayed-transfer results. Studies covering immediate-transfer testing were discussed in the systematic review. This approach is similar to [Brady \(2004\)](#) since he used only delayed outcomes, assuming that learning effects may be obscured in the more immediate measures.

Search methods and selection procedure

AW performed the search on the following databases: EBSCO (“contextual interference” in Title OR Abstract—no Keywords option),

Scopus (“contextual interference” in Title OR Abstract OR Keywords), ScienceDirect (“contextual interference” in Title OR Abstract OR Keywords), Web of Science (“contextual interference” in Topic), in September 2021 (2020 to 2021), and updated March 2022 (period 2021–2022) and November 2022. Relevant studies were scrutinized using the Google Scholar search engine (“contextual interference” in Title OR Abstract OR Keywords). SC searched the PsycINFO database (“contextual interference” in Title OR Abstract OR Keywords).

Studies in languages other than English and the “gray literature” (e.g., master and Ph.D. dissertations and theses, conference proceedings, non-reviewed articles, etc.) available online in the searched databases were excluded to facilitate the reliable risk-of-bias assessment. Exclusion of non-English literature does not cause systematic bias (Morrison et al., 2012) and the proper and reliable translation from other languages into English may be problematic (Jackson and Kuriyama, 2019). Moreover, Jackson and Kuriyama noticed that only 2% of the articles included in the systematic reviews were published in languages other than English (Jackson and Kuriyama, 2019). On the other hand, there was no reason to include gray literature, since the topic (CI) is not novel (Gunnell et al., 2020) and gray literature inclusion would cause unnecessary burden on the reviewers (Mahood et al., 2014). Some of the non-peer-reviewed documents have different structure, different length, some of them are not peer-reviewed at all, some only partially.

Due to a large number of retrieved studies, a method proposed by Dundar and Fleeman (2017) was applied, i.e., AW evaluated the titles, abstract, and keywords for inclusion criteria, and a random sample was cross-checked by the senior researcher (SC). Inadequate articles were excluded, and duplicates of identified studies were removed. Subsequently, the full text of each study was read by the two co-authors (AW and PS), who independently assessed the papers for final eligibility. In case of a disagreement, a senior researcher (SC) was consulted, and a consensus was reached.

Data collection and analysis

During the screening process, all relevant data were summarized by AW and PS in developed MS Excel data extraction forms. Each entry consisted of study specifications, such as the authors’ names, the study title, the year of publishing, and the journal title, in case of multiple experiments in the same study—number of experiments. The following details were based on PICO criteria:

- *Population* (age, gender, number of participants, expertise level).
- *Intervention* (testing procedure, dependent variable, nature of research, practice schedule, type of motor task).
- *Objectives/Outcomes* (immediate transfer results and delayed transfer results: extracted means and standard deviations for all groups and all measures). Consistently with our meta-analysis on retention (2023?), the results of the first block only from the transfer testing procedures were considered for the extraction. We assumed that the following blocks might enhance further learning. If the standard error of the mean (SEM) was available, we converted it into standard deviation (SD). If quartiles were available, we used the Mean Variance calculator (Luo et al., 2018; Shi et al., 2020a,b) to convert these into SD. Furthermore, study quality indicators were included, covering the following sections:

selection bias, study design, confounders, blinding, data collection methods, withdrawals and dropouts, and global rating, based on *Quality Assessment tool for Quantitative studies* (Thomas, 2003; Thomas et al., 2004).

Since included studies utilized different motor skills (tasks) and the transfer was measured with different units (meters, seconds, number of cycles etc.) and scores (percentages, numbers), we used standardized mean differences (SMDs) effect sizes, i.e., Hedges’ (adjusted) g , very similar to Cohen’s d , but it includes an adjustment for small sample bias. The I^2 statistic was used to evaluate the heterogeneity among the studies. The interpretation of I^2 is as follows: 30–60% represent moderate heterogeneity; 50–90%—substantial heterogeneity; and 75–100%—considerable heterogeneity (Higgins et al., 2021). Nevertheless, interpretation thresholds can be misleading (Deeks et al., 2019).

In line with the Cohen’s recommendation of interpreting the magnitude of SMD in the social sciences (Cohen, 1988), we applied the following guidelines: small (SMD = 0.2); medium (SMD = 0.5); and large (SMD = 0.8).

We computed a 3-level mixed model which uses (restricted) maximum likelihood procedures (Cheung, 2014; Assink and Wibbelink, 2016). The advantage of that model is that it takes into account the potential dependence among the effect sizes, i.e., when there are multiple outcomes (effect sizes) from a single study. The model assumes that the random effects at different levels and the sampling error are independent. Three levels of the model refer to variance between effects sizes among participants (level 1), outcomes, i.e., effect sizes extracted from the same study (level 2; within-cluster variance), and studies (level 3; between-clusters variance) (Assink and Wibbelink, 2016). There is no need to know the correlations between outcomes extracted from one study for estimating the covariance matrix of the effect sizes since the second level in the model accounts for sampling covariation (Assink and Wibbelink, 2016).

Sensitivity analysis was performed using Cook’s D distances. Outcomes further than $4/n$ (where n was the number of outcomes) were removed to assess how these outliers influence the pooled effect.

Meta-analyses were performed with RStudio (version 2023.06.0) and the following packages “metaphor,” “dmetar,” “tidyverse,” “readxl,” and “ggplot.”

Assessment of risk of bias/quality assessment of included studies

We followed the guidelines of the Effective Public Health Practice Project (EPHPP) *Quality Assessment Tool for Quantitative Studies* (Thomas, 2003) while assessing the risk of bias in included studies. The checklist elements (sample selection, study design, identification of confounders, blinding, reliability and validity of data collection methods, withdrawals, and dropouts) could be rated strong, moderate, or weak. According to a standardized guide and dictionary, the comprehensive evaluation is determined by assessing six rating aspects. Studies with two or more weak ratings are considered weak, those with less than four strong ratings and one weak rating are considered moderate, and, subsequently, studies with no weak ratings and at least four strong ratings are regarded as strong (Thomas et al., 2004). AW and PS independently assessed the level of evidence and

methodological quality of the eligible studies. In case of discrepancy, the authors discussed until a consensus was reached. Any issues regarding the quality of the study was discussed with the senior researcher (SC).

Results

Results of the search

The primary search in the databases identified 2,161 potential records. After removing duplicates, the titles and abstracts of 1,287 articles were screened according to PICO criteria, including 8 records identified manually (see [Supplementary Appendix 2](#) for more details). Nine hundred eighty-seven articles (987) were excluded due to study design issues, not relevant topics, and population. The detailed evaluation process is highlighted in the PRISMA Flow Diagram (Moher et al., 2009; [Figure 1](#)).

The remaining 300 studies were evaluated, and 258 were excluded (comprehensive reasons for exclusion are listed in [Supplementary Appendix 3](#)). In case of missing data, the authors were contacted via e-mail and/or the ResearchGate platform. Finally, 42 articles were included in the present systematic review. The quantitative analysis covered 34 studies. Transfer tests conducted up to 24 h after the acquisition were defined as *immediate transfer testing*, and consequently, testing procedures performed after 24 h were defined as a *delayed transfer*. Eight studies described immediate transfer testing; therefore, the meta-analysis did not include these results. The summary of all included studies is provided in [Supplementary Table S1](#). The summary of studies characteristics is presented in [Supplementary Table S2](#).

Reasons for exclusion of individual experiments or particular groups of participants

Occasionally, more than one experiment was presented in a single paper. There were cases where the authors did not report data on all of them, or, similarly, some studies did not report statistically non-significant data on specific variables. In such situations, we contacted authors; however, the authors did not provide the data in a few cases—the main reported reason was that their studies were performed a long time ago.

An article by [Ste-Marie et al. \(2010\)](#) consisted of three experiments where the authors examined the CI effect on handwriting skills in young participants. Unfortunately, it was not possible to obtain data from the first experiment. In the second experiment, data on the scores measure, and in the third experiment, data on the time variable were unavailable. In the paper by [Porter and Magill \(2010\)](#) covering two experiments, the results of the first one were available. The second experiment was excluded from the analysis due to group characteristics not compliant with PICO: group of ratio-feedback and segment-feedback. In the study by [Chua et al. \(2019\)](#) three experiments were conducted; however, the results of the first one were excluded from our review as these were describing constant practice group instead of blocked practice.

Some of the included studies covered more than two (random and blocked) groups of participants. In line with PICO criteria, in that situation, only the results of groups with blocked and random/interleaved schedules were regarded as appropriate for the analysis. Included primary studies consisting of more than two groups are listed below. In the study by [Goodwin and Meeuwse \(1996\)](#), participants were randomly assigned to three groups: random, blocked/random, and blocked. Transfer results of blocked and random groups were included in the meta-analysis. Participants of the study by [Porter and Beckerman \(2016\)](#) were randomly allocated to three groups with random, increasing, or blocked schedules. The random and blocked group transfer data were considered applicable for the current analysis. In the study by [Beik et al. \(2021\)](#), participants were randomly assigned into six groups of blocked-similar, algorithm-similar, random-similar, random-dissimilar, blocked-dissimilar, and algorithm-dissimilar. Out of these groups, transfer results of four (blocked-similar, blocked-dissimilar, random-similar, random-dissimilar) were considered appropriate according to PICO. A similar situation occurred in the study by [Beik and Fazeli \(2021\)](#), where participants were randomly allocated to one of six groups of blocked-similar, blocked-dissimilar, learner-adapted-similar, random-similar, random-dissimilar and learner-adapted-dissimilar. Four groups were considered appropriate for analysis (random-similar, random-dissimilar, blocked-similar, blocked-dissimilar).

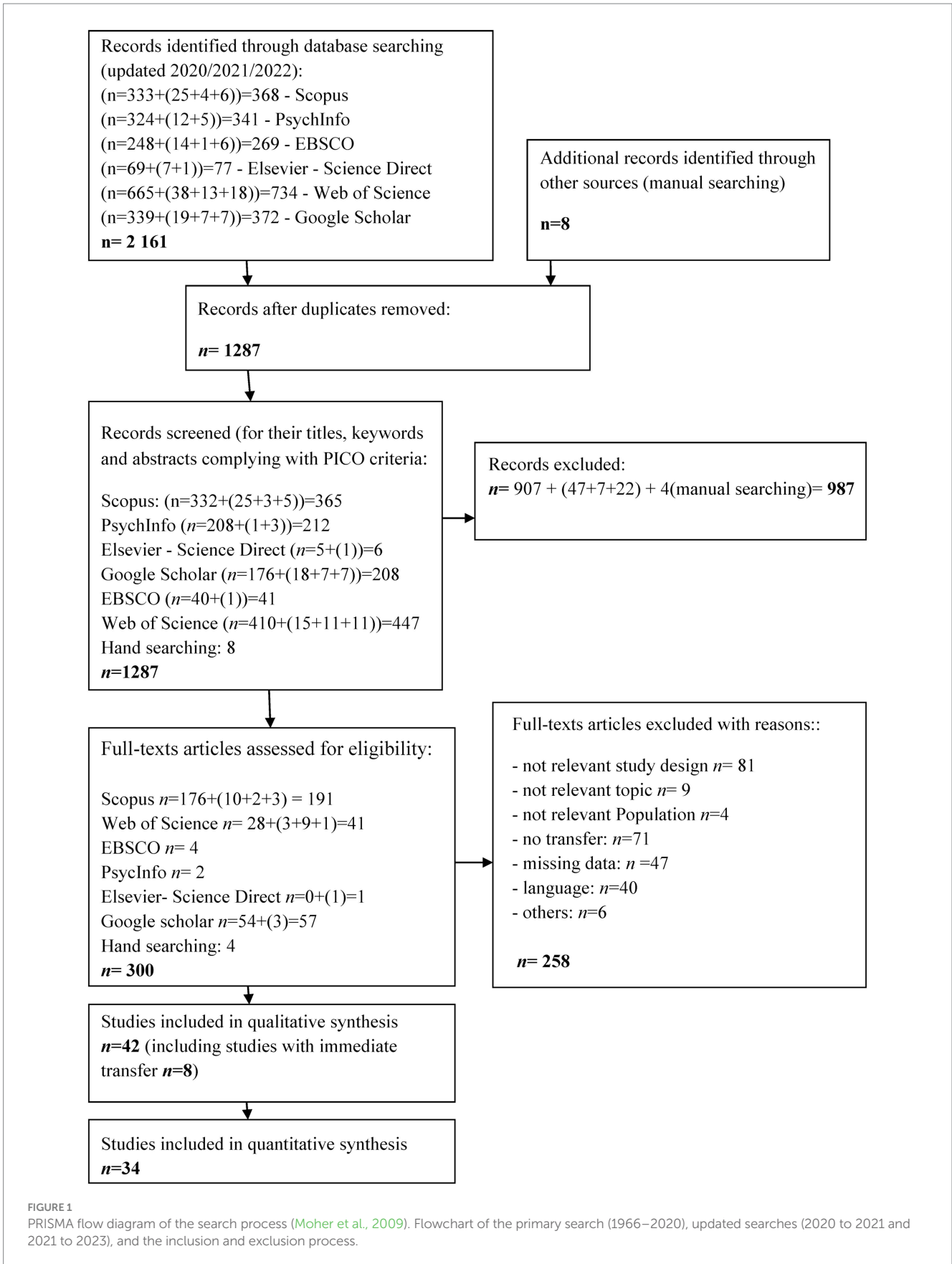
Results of quality assessment of included studies

The detailed methodological assessment of the included studies is presented in [Supplementary Appendix 4](#). None of the included studies was assessed as strong. Only two articles ([Ste-Marie et al., 2010](#); [Johnson et al., 2022](#)) presented moderate methodological quality according to the *Quality Assessment Tool for Quantitative studies* ([Thomas, 2003](#)). The primary studies failed mainly on the latter criteria: 26 studies scored weak in Selection Bias, more than 21 articles scored weak in the *Confounders* section, 30 received weak ratings in the *Withdrawals and dropouts* section. The applied assessment tool specification might explain such relatively strict evaluation: two weak ratings were enough to automatically determine a weak classification of a study in its global rating for all six determinants of the checklist.

According to [Thomas et al. \(2004\)](#), only articles rated as moderate or strong should be included in the meta-analysis. Nevertheless, excluding all articles with *weak* global ratings would make the current analysis relatively doubtful (with only two studies included). For this reason, we have decided to include 34 studies in the meta-analysis. The impact of this decision on heterogeneity was taken into account.

Findings

As aforementioned, only delayed transfer results were considered for the current meta-analysis, yielding 86 effect sizes. Outcomes of 34 studies were included, resulting in testing 1,421 participants. A wide range of variables was involved: time (absolute error time, decision time, variable time, response time, reaction time, completion time), the number of performed movements, distance (absolute error



distance, accuracy error distance, median pathway traveled), accuracy (proficiency percentage, accuracy scores). Evaluation of transfer was presented by the mean of various outcome measures' units: meters, seconds, percentages, or scores.

Motor skills utilized in primary studies varied in many ways. They were presented in different arrangements: discrete and continuous motor skills, open and closed motor skills, or fine and gross motor skills. Motor tasks were associated with volleyball (Bortoli et al., 1992; Meira and Tani, 2003; Fialho et al., 2006; Travlos, 2010; Pasand et al., 2016), golf (Goodwin and Meeuwse, 1996; Porter and Magill, 2010; Chua et al., 2019), tennis (Broadbent et al., 2015), hockey (Cheong et al., 2016), throwing darts (Meira and Tani, 2001; Frömer et al., 2016), hopping (Parab et al., 2018), basketball (Porter et al., 2020), baseball (Hall et al., 1994), throwing (Vera and Montilla, 2003; Chua et al., 2019), learning of the Pawlata roll (Smith and Davies, 2008) and rifle shooting (Moretto et al., 2018). Non-sports skills included laparoscopic skills (Shewokis et al., 2017; Johnson et al., 2022) and handwriting (Ste-Marie et al., 2010), and Wii-Fit dynamic balance task (Jeon et al., 2020).

The other motor tasks applied in the primary studies covered: serial reaction time tasks (Lin et al., 2018), key pressing tasks (Del Rey et al., 1994; Shewokis, 1997; Shea et al., 2001; Meira et al., 2015; Beik et al., 2021), pursuit tracking tasks (Dunham et al., 1991; Porter and Beckerman, 2016), positioning tasks (Perez et al., 2005; Lage et al., 2006) and reversal or rapid movements on manipulandum (Green and Sherwood, 2000; Herzog et al., 2022).

Laboratory versus applied setting—comparison characteristics

Laboratory vs. applied meta-analytic comparison included 34 studies. Seventeen studies, covering 46 effect sizes in total, were describing experiments conducted in laboratory settings. There were 738 participants in laboratory studies, of which 121 were older adults, and 542 were adults. The age of participants in the aforementioned studies ranged from 10 ± 0.6 years (Perez et al., 2005) to 82 years (Jeon et al., 2020). It is worth mentioning that only 75 participants from the laboratory studies were less than 18 years old (Perez et al., 2005; Broadbent et al., 2015). Motor skills utilized in the laboratory setting were: Wii-Fit dynamic balance task (Jeon et al., 2020), rapid movements on manipulandum (Green and Sherwood, 2000; Herzog et al., 2022), pursuit tracking tasks (Dunham et al., 1991; Porter and Beckerman, 2016), key pressing tasks (Shewokis, 1997; Shea et al., 2001; Meira et al., 2015; Beik et al., 2021), serial reaction time task (Lin et al., 2018), positioning tasks (Perez et al., 2005; Lage et al., 2006). In the study of Broadbent (Broadbent et al., 2015), the acquisition of tennis skills was conducted in laboratory settings—similar to the study by Chua et al. (2019), where the acquisition and testing of throwing skills and golf were assessed in the laboratory environment. An article by Frömer et al. (2016) described virtual darts throwing in the laboratory environment.

Applied studies were performed in natural environments (during physical education classes or game-based). Seventeen articles described experiments conducted in an applied setting, covering 40 effect sizes. Six hundred eighty-three (683) participants were included in these studies, of which 393 were under 18 years old, and 290 were adults. An article by De Souza et al. (2015) was the only study describing the transfer of motor skills in applied settings in older

adults (65–80 years old). The motor task implemented in the study consisted of throwing a bocchia ball to three targets. However, this study was excluded from the meta-analytic analysis due to missing data. The age of participants in the aforementioned studies ranged from 5.5 years (Ste-Marie et al., 2010) to 35 years (Thomas et al., 2021).

The following motor skills were practiced and examined in the included applied studies: golf skills (Goodwin and Meeuwse, 1996; Porter and Saemi, 2010), volleyball skills (Bortoli et al., 1992; Meira and Tani, 2003; Fialho et al., 2006; Travlos, 2010; Pasand et al., 2016), hockey (Cheong et al., 2016), baseball (Hall et al., 1994), throwing skills (Vera and Montilla, 2003), rifle shooting (Moretto et al., 2018). In the study by Moretto et al. (2018) acquisition and testing of rifle shooting were conducted in indoor laboratories; however, all adjustments, including the position target height, followed the *Olympic and International Shooting Sport Federation* standards; and for this reason results were included in the *applied setting* comparison. Pawlata roll (Smith and Davies, 2008) learning and testing procedures took place in the indoor pool. CI effect on handwriting skills in children was tested by Ste-Marie et al. (2010) in the school setting. Laparoscopic skills in the study by Shewokis et al. (2017) were performed by medical students and post-graduate residents using a virtual reality simulator (LapSim® VR simulator), mimicking the regular laparoscopic tasks. In the study by Johnson and colleagues (Johnson et al., 2022), laparoscopic skills were practiced and tested with the use of *Fundamentals of Laparoscopic Surgery* (FLS) box trainer (VTI Medical, Waltham, MA). The laparoscopic tasks were acquired and tested in both articles according to the FLS program.

The CI effect in youth vs. adults vs. elderly adults—comparison characteristics

The age of participants in the studies included in the present review ranged from 5.5 years (Ste-Marie et al., 2010) to 82 years (Jeon et al., 2020). The age subgroup analyses were classified as follows: young (up to 18 years old), adults (18 years old to 59 years old), and older adults (60 years and older). The eight articles reporting immediate transfer covered the results of 80 children (Del Rey et al., 1983a), 10 adolescents (Fialho et al., 2006), and 318 adults (Del Rey et al., 1983b; Smith and Rudisill, 1993; Del Rey et al., 1994; Meira and Tani, 2001; Lage et al., 2006; Sherwood and Duffel, 2010). In the delayed transfer studies, there were 468 young participants, 812 adults, and 121 older adults. The results of 20 participants aged 17–21 described in the article by Hall and colleagues (Hall et al., 1994) were not classified in any of the age subgroups analysis. According to specific age group criteria applied in the present study: young (up to 18 years old), adults (18 years old to 59 years old), and older adults (60 years and older), the participants from the aforementioned study shall be included in both groups young and adults simultaneously. Therefore, we refrained from including their results.

The CI effect in novice vs. experienced participants—comparison characteristics

In his meta-analytic study on CI, Brady (2004), among the other subgroup analyses, compared the CI effect between novice and skilled participants. While classifying their skill levels, Brady was guided by how the authors labeled the participants' experience. We utilized the same

rule in our review. In the immediate transfer articles, 90 participants were described as skilled (Del Rey et al., 1983a; Fialho et al., 2006). Out of the studies on delayed transfer, we classified 62 participants as experienced (Hall et al., 1994; Broadbent et al., 2015; Porter et al., 2020). Participants of these studies were characterized as follows: “intermediate-level junior tennis players” (Broadbent et al., 2015, p. 1245), “baseball players from a junior college baseball team (...) with an average of 9.5 years of experience in competitive baseball” (Hall et al., 1994, p. 837–838), participants “had less than two years’ basketball playing experience (1.1 ± 1.3 years) and no representative level basketball playing experience” (Porter et al., 2020, p. 7).

The primary studies in the current review referred to different inclusion standards for a skilled group. Classifying skill (experience) levels correctly could lead our review in a different direction, still not warranting there will be no confusion or doubts. Therefore, we decided not to conduct the subgroup analysis of *skilled* versus *novice*.

Meta-analysis: results

The analysis of the CI effect on delayed transfer (Figure 2) covered 34 studies, yielding 86 effect sizes and resulting in testing 1,421 participants.

The pooled effect size based on the three-level meta-analytic model was medium $SMD = 0.55$ (95% CI: 0.25, 0.86; $p < 0.001$). The estimated variance components (tau squared) were $\tau_3^2 = 0.488$ and $\tau_2^2 = 0.403$ for the level 3 and level 2 components, respectively. This means that $I^2_3 = 47\%$ of the total variation can be attributed to between-cluster, and $I^2_2 = 39\%$ to within-cluster heterogeneity. Total $I^2 = 86\%$. We found that the three-level model provided a significantly better fit compared to a two-level model with level 3 heterogeneity constrained to zero ($\chi^2 = 9.99$; $p < 0.001$): for three level model ($df = 3$) $AIC = 246$ while for the 2-level model $AIC = 254$. Test of moderators: $F(1, 84) = 1.34$, $p = 0.25$ suggesting the subgroup analyses are not necessary, though, we decided to perform them analogically to Brady (2004).

Sensitivity analysis revealed that there were 5 outcomes which were further than $4/n$ threshold (Figure 2). These were outcomes from Beik et al. (2021) and Beik and Fazeli (2021)—one outcome, Green and Sherwood (2000) and Parab et al. (2018). After having removed the outliers (Figure 3), the pooled effect size was small $SMD = 0.40$ (95% CI: 0.18, 0.62; $p < 0.001$). The estimated variance components (tau squared) were $\tau_3^2 = 0.19$ and $\tau_2^2 = 0.23$; $I^2_3 = 34\%$ and $I^2_2 = 41\%$; respectively. The outliers had a substantial effect on pooled effect size, i.e., SMD decreased from 0.55 (with outliers included) to 0.40 (without outliers).

Laboratory vs. field-based (applied) studies

The primary studies were divided into those carried out in a laboratory setting ($n = 17$), including 738 participants (46 effect sizes), and the remaining ($n = 17$) conducted in an applied setting (40 effect sizes), including 683 participants.

A subgroup analysis of the CI effect in laboratory studies was performed (Figure 4). The pooled effect size based on the three-level meta-analytic model was medium $SMD = 0.75$ (95% CI: 0.26, 1.25; $p = 0.004$). The estimated variance components (tau squared) were $\tau_3^2 = 0.68$ and $\tau_2^2 = 0.52$ for the level 3 and level 2 components, respectively. Heterogeneity was high, $I^2_3 = 50\%$ and $I^2_2 = 39\%$; total $I^2 = 88.96\%$.

Sensitivity analysis revealed that after removing two outcomes, i.e., Green and Sherwood (2000) and Lin et al. (2018), the pooled effect size dropped $SMD = 0.62$ (95% CI: 0.24, 1.01; $p = 0.002$).

Analogously, a subgroup analysis of the CI effect in applied studies was conducted (Figure 5). The pooled effect size based on the three-level meta-analytic model was small $SMD = 0.37$ (95% CI: -0.02 , 0.69; $p = 0.06$). The estimated variance components (tau squared) were $\tau_3^2 = 0.27$ and $\tau_2^2 = 0.31$ for the level 3 and level 2 components, respectively. Heterogeneity was high, $I^2_3 = 37\%$ and $I^2_2 = 44\%$; total $I^2 = 81.35\%$.

Sensitivity analysis revealed that after removing two outcomes, i.e., Travlos (2010) and Pasand et al. (2016), the pooled effect size decreased to negligible $SMD = 0.11$ (95% CI: -0.13 , 0.34; $p = 0.36$).

The CI effect in young vs. adults vs. elderly adults

Thirty-three articles were included in a meta-analytic comparison of the CI effect in three age groups (Figure 6), resulting in the testing of 468 young participants, 812 adults, and 121 older adults. As aforementioned, the results of 20 participants aged 17–21 from the study by Hall et al. (1994) were not included in the current subgroup analysis. This analysis yielded 2,632 measurements in total and elicited: 21 effect sizes for young participants, 55 effect sizes for adults, and eight effect sizes for the group of older adults.

Firstly, we performed an analysis for the subgroups of young participants. The pooled effect size based on the three-level meta-analytic model was negligible $SMD = 0.12$ (95% CI: -0.28 , 0.53; $p = 0.54$). The estimated variance components (tau squared) were $\tau_3^2 = 0.10$ and $\tau_2^2 = 0.31$ for the level 3 and level 2 components, respectively. Heterogeneity was high, $I^2_3 = 18\%$ and $I^2_2 = 60\%$; total $I^2 = 78.42\%$.

Sensitivity analysis revealed that after removing one outcome, i.e., Travlos (2010), the pooled effect size decreased to $SMD = 0.02$ (95% CI: -0.36 , 0.39; $p = 0.36$).

Secondly, we analyzed the adult’s subgroup (Figure 7). The pooled effect size based on the three-level meta-analytic model was medium $SMD = 0.54$ (95% CI: 0.16, 0.92; $p = 0.54$). The estimated variance components (tau squared) were $\tau_3^2 = 0.42$ and $\tau_2^2 = 0.53$ for the level 3 and level 2 components, respectively. Heterogeneity was high, $I^2_3 = 39\%$ and $I^2_2 = 48\%$; total $I^2 = 86.82\%$.

Sensitivity analysis revealed that after removing five outcomes, i.e., Green and Sherwood (2000), Shea et al. (2001), Travlos (2010), Pasand et al. (2016), and Lin et al. (2018), the pooled effect size decreased to medium $SMD = 0.38$ (95% CI: 0.13, 0.62; $p = 0.003$).

Thirdly, an analysis for older adults was performed (Figure 8). The pooled effect size based on the three-level meta-analytic model was large $SMD = 1.28$ (95% CI: -0.34 , 2.90; $p = 0.10$). The estimated variance components (tau squared) were $\tau_3^2 = 1.24$ and $\tau_2^2 = 0.15$ for the level 3 and level 2 components, respectively. Heterogeneity was high, $I^2_3 = 77\%$ and $I^2_2 = 10\%$; total $I^2 = 87.05\%$.

There were no outcomes further than $4/n$ Cook’s D distances.

Risk of publication bias assessment

The risk of publication bias was assessed using a funnel plot (Figure 9). Given that substantial heterogeneity was present in each of the analyses we performed, we decided not to apply other statistical methods, i.e., Egger’s regression test or rank-correlation test.

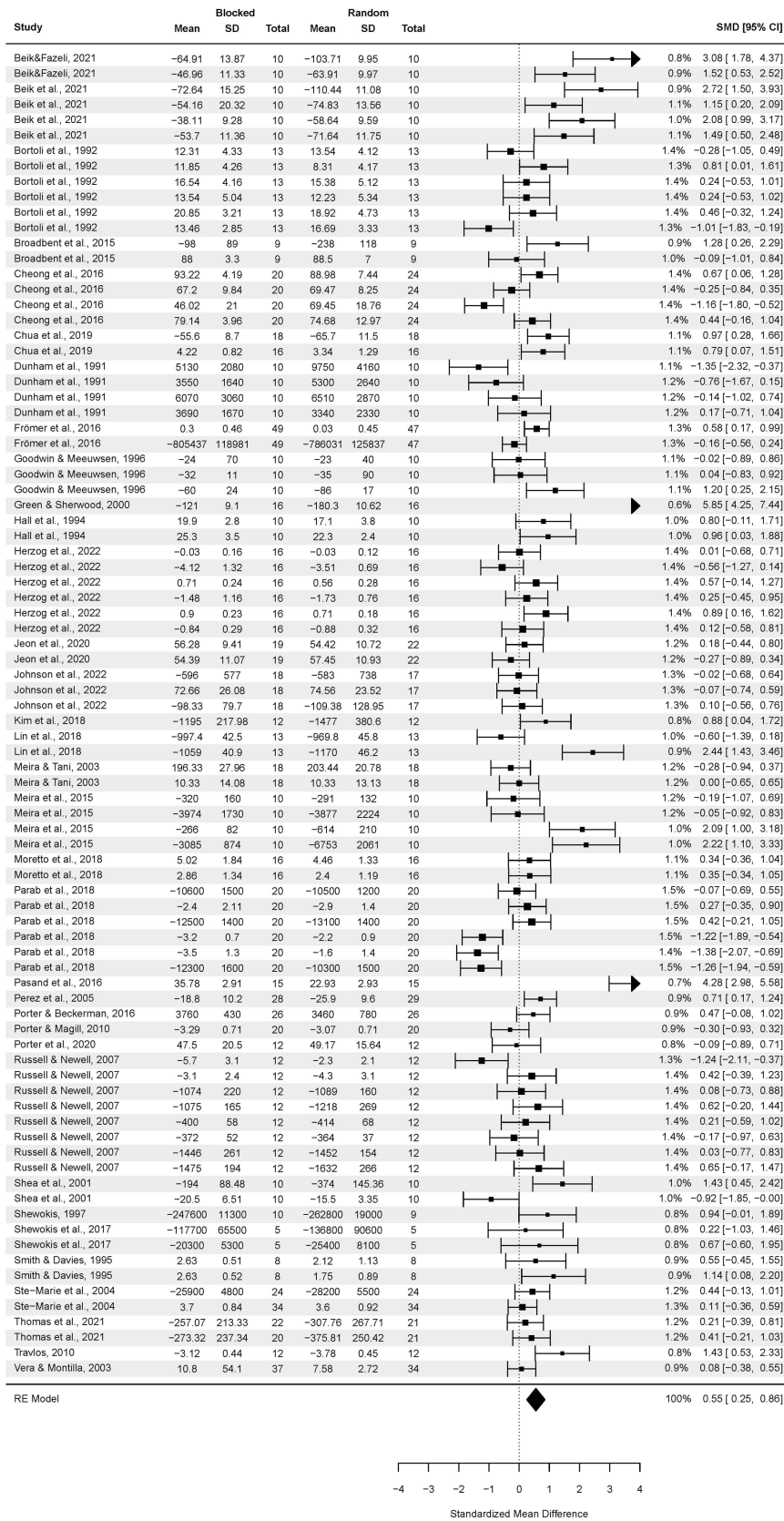


FIGURE 2 Analysis of transfer test results of random practice vs. blocked practice. The forest plot presents the results obtained by participants aged 5.5–82, including various motor tasks and different outcome measures.

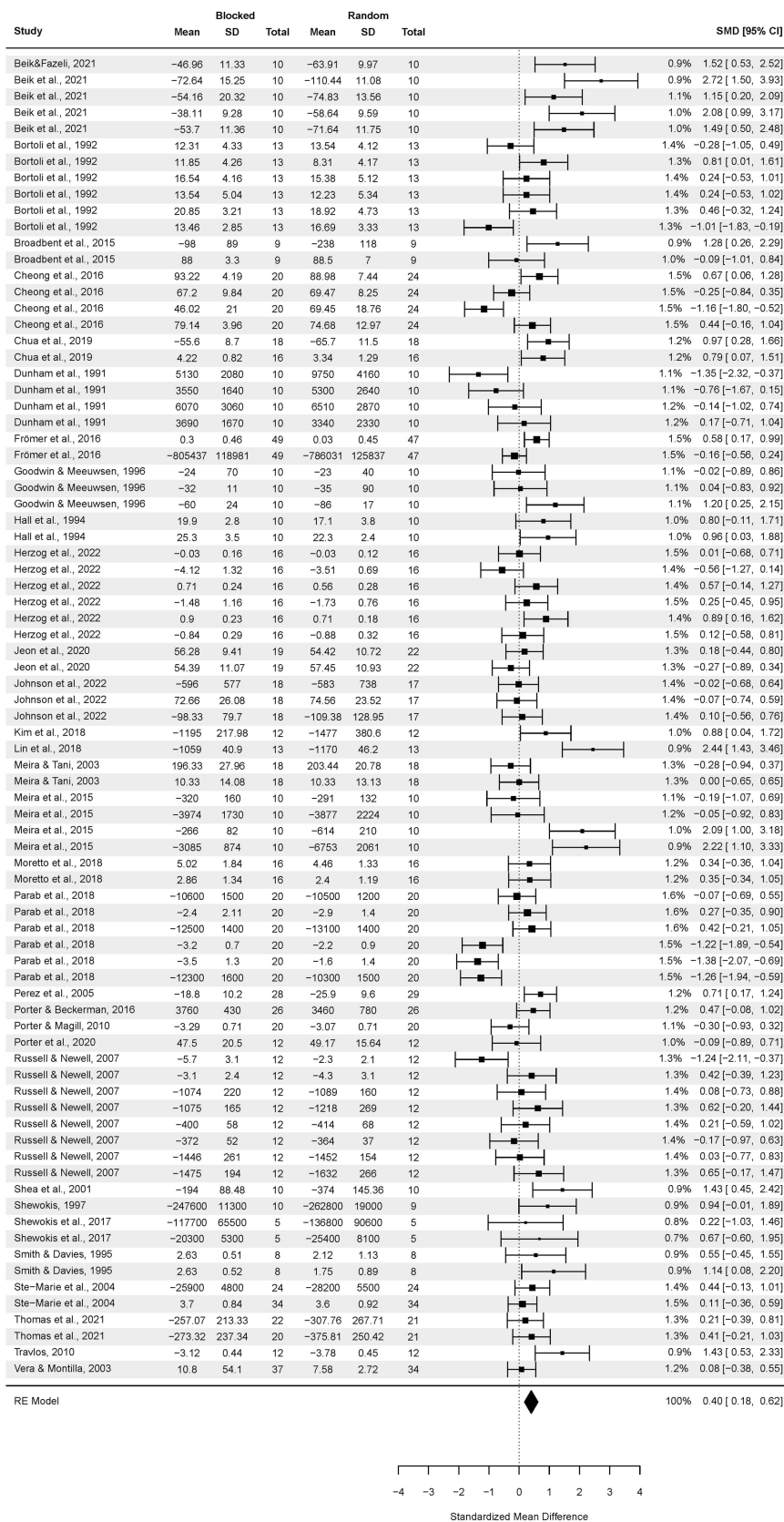


FIGURE 3 Analysis of transfer test results of random practice vs. blocked practice with no outliers (further than 4/n Cook D distances).

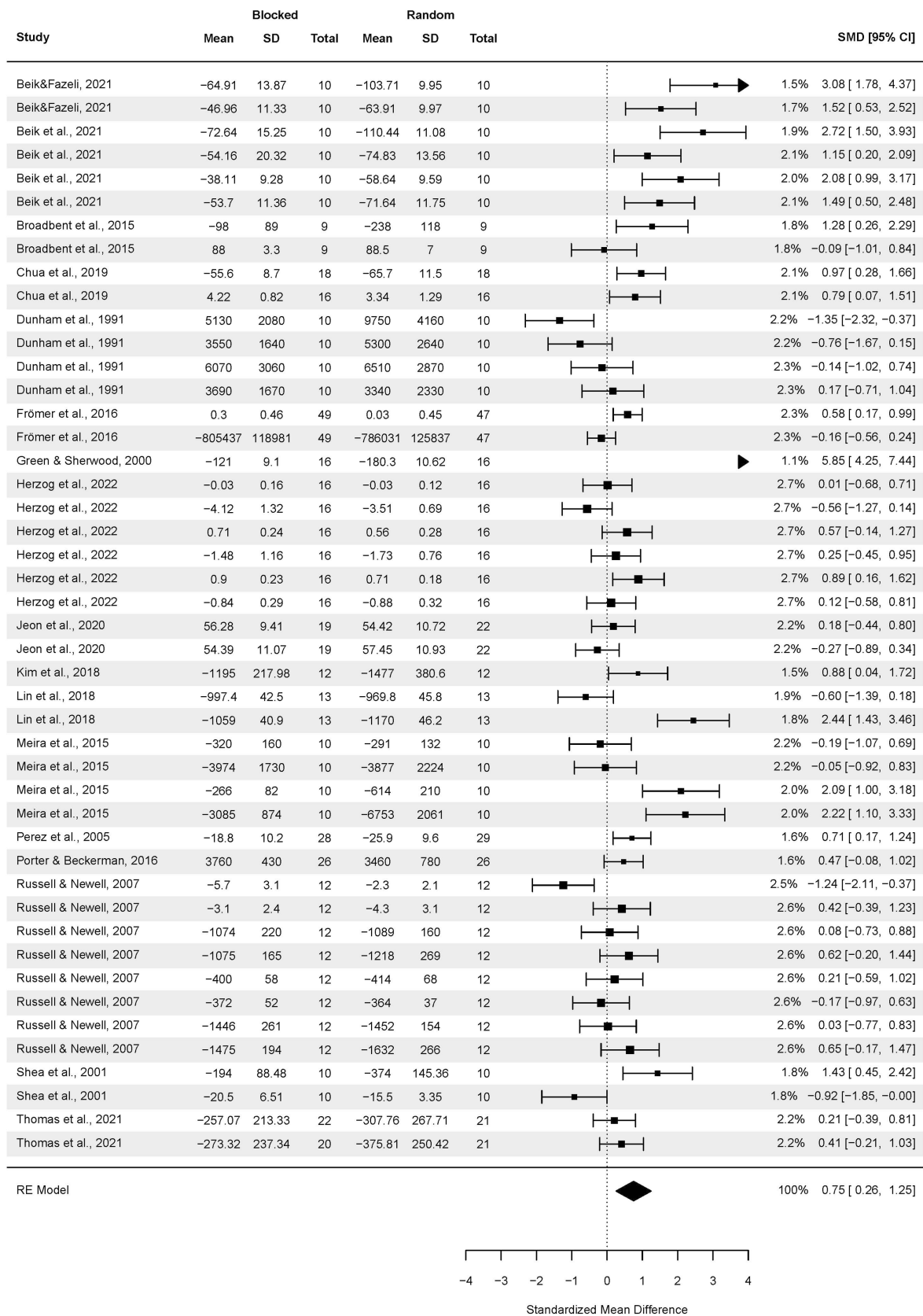


FIGURE 4 Analysis of transfer test results in a random and blocked schedule in a laboratory setting. The forest plot presents the transfer test results obtained by participants practicing in a laboratory setting, including various motor tasks and different outcome measures.

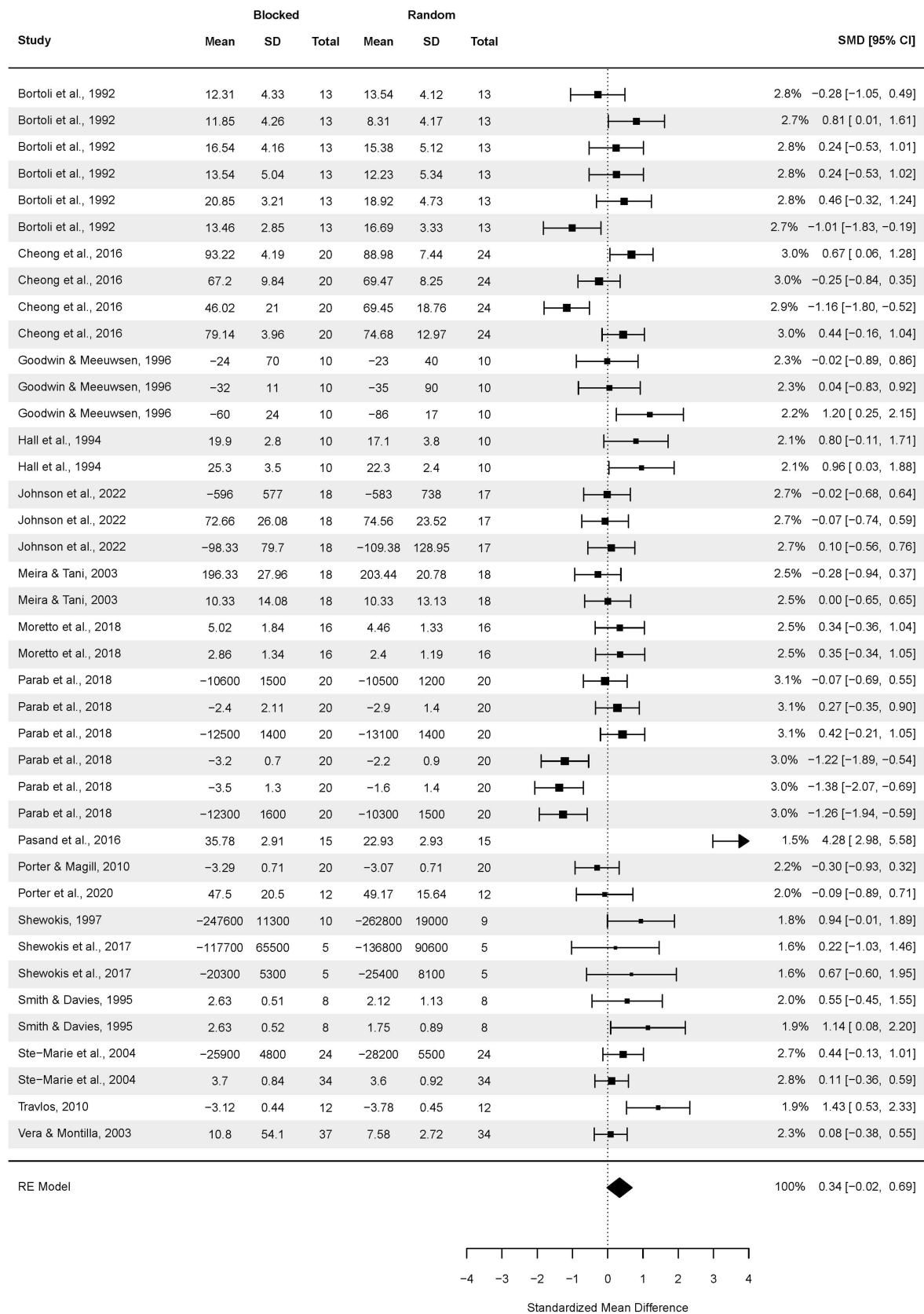


FIGURE 5 Analysis of random and blocked schedule transfer test results in an applied setting. The forest plot presents the transfer test results in an applied setting, including various motor tasks and different outcome measures.

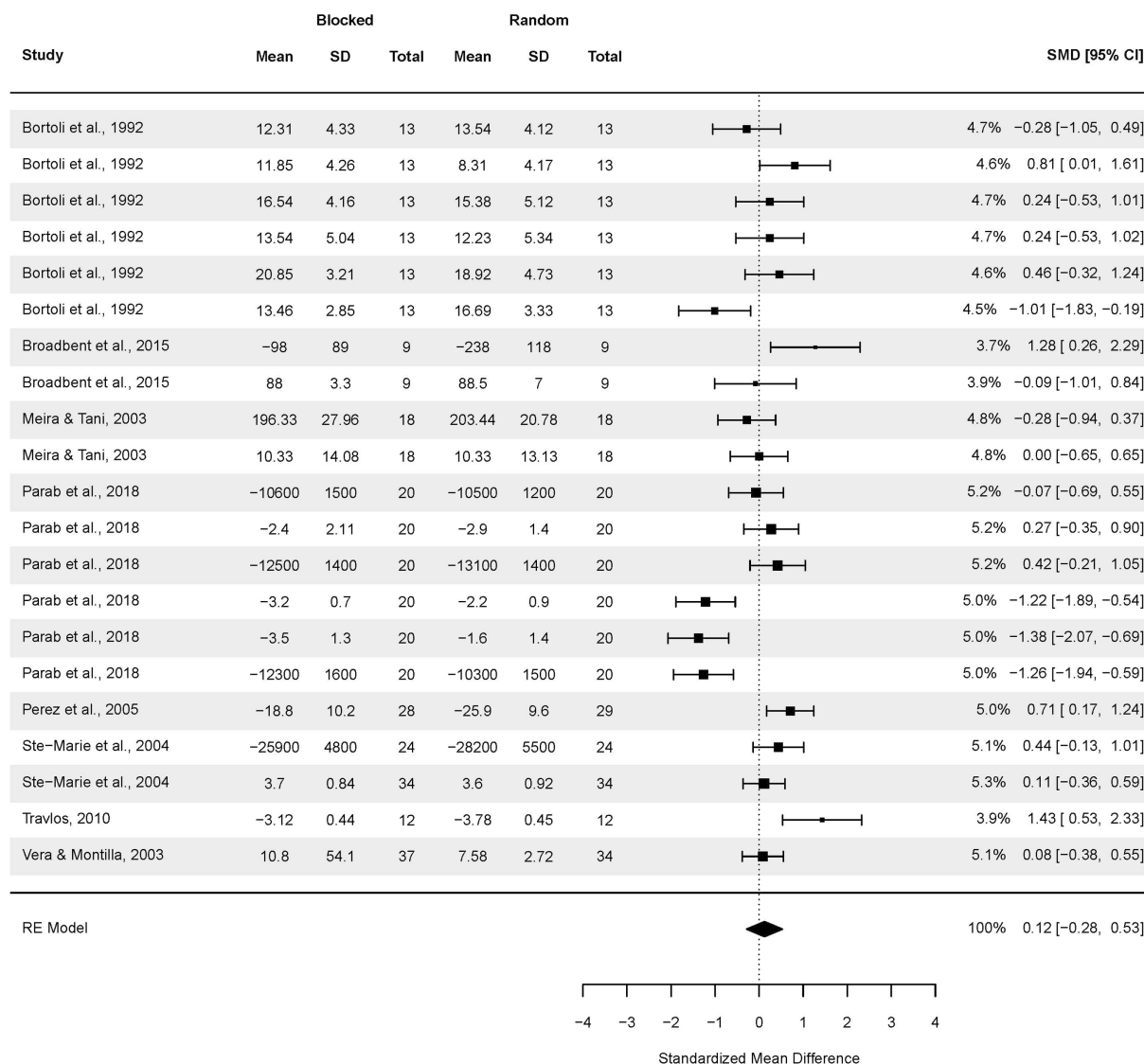


FIGURE 6 Analysis of young participants' transfer test results: random vs. blocked practice. The forest plot presents the transfer test results of participants aged 5.5–18, including various motor tasks and different outcome measures.

Discussion

The study's primary objective was to determine the overall effect size of CI on transfer in motor learning. We found that the statistically significant overall CI effect on transfer was medium (SMD = 0.55) in favor of the random practice.

Our secondary objectives were to estimate the CI effect on transfer in laboratory versus non-laboratory studies and the CI effect in different age groups. Similarly, to the overall effect, the random practice was favored in the laboratory and applied settings. However, in laboratory studies, the medium effect size was statistically significant (SMD = 0.75), whereas, in applied studies, the effect size was small and statistically non-significant (SMD = 0.34). Significant and larger effect sizes in laboratory settings may be due to well-controlled environmental variables and simpler tasks utilized in laboratories (Jeon et al., 2020). Complex tasks used in applied settings may be too challenging for information processing (Hebert et al., 1996) and deleterious for learning as a result (Wulf and Shea, 2002). However, there may be another

explanation; as Al-Mustafa stated, CI is a laboratory artifact (Al-Mustafa, 1989; Brady, 2004), i.e., CI effect is conspicuous in labs but not in real life.

Age group analysis turned out to be significant only in adults and older adults. In both age categories, random practice was favored. In adults, the effect was medium (SMD = 0.54), whereas in older adults was large (SMD = 1.28). The results in the adult group align with Brady's (Brady, 2004), who reported small effect sizes in both laboratory and applied settings. However, Brady did not recognize the older adult group; therefore, comparing the results with any previous ones is difficult. On the other hand, nonsignificant effect size in young participants was negligible (SMD = 0.12).

Comparison with Brady's and Ammar's et al. meta-analysis

Only seven studies originally included in Brady's meta-analysis, yielding 17 effect sizes, were included in our analysis.

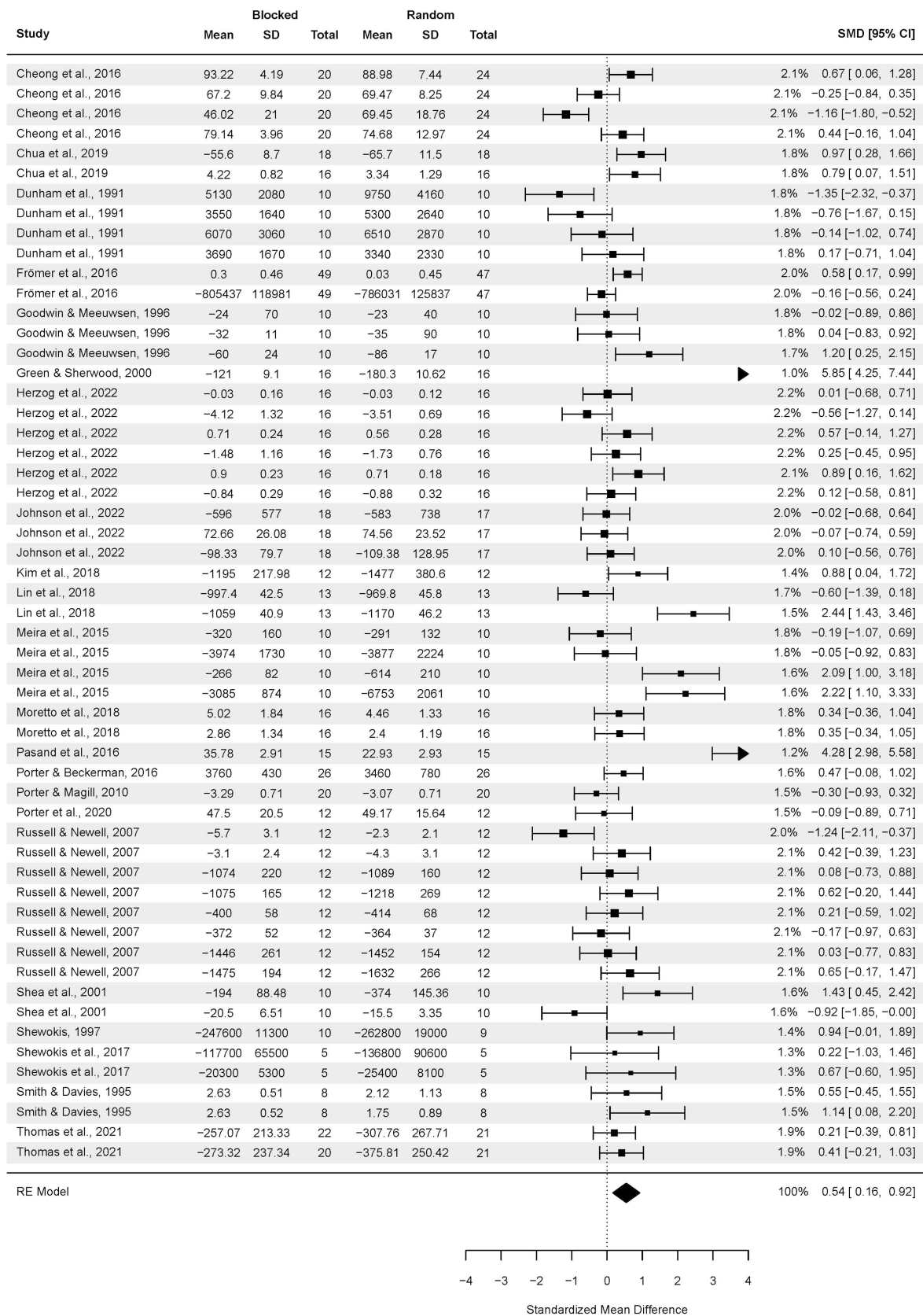


FIGURE 7 Analysis of adult participants' transfer test results: random vs. blocked practice. The forest plot presents the transfer test results obtained by participants aged 18–59, including various motor tasks and different outcome measures.

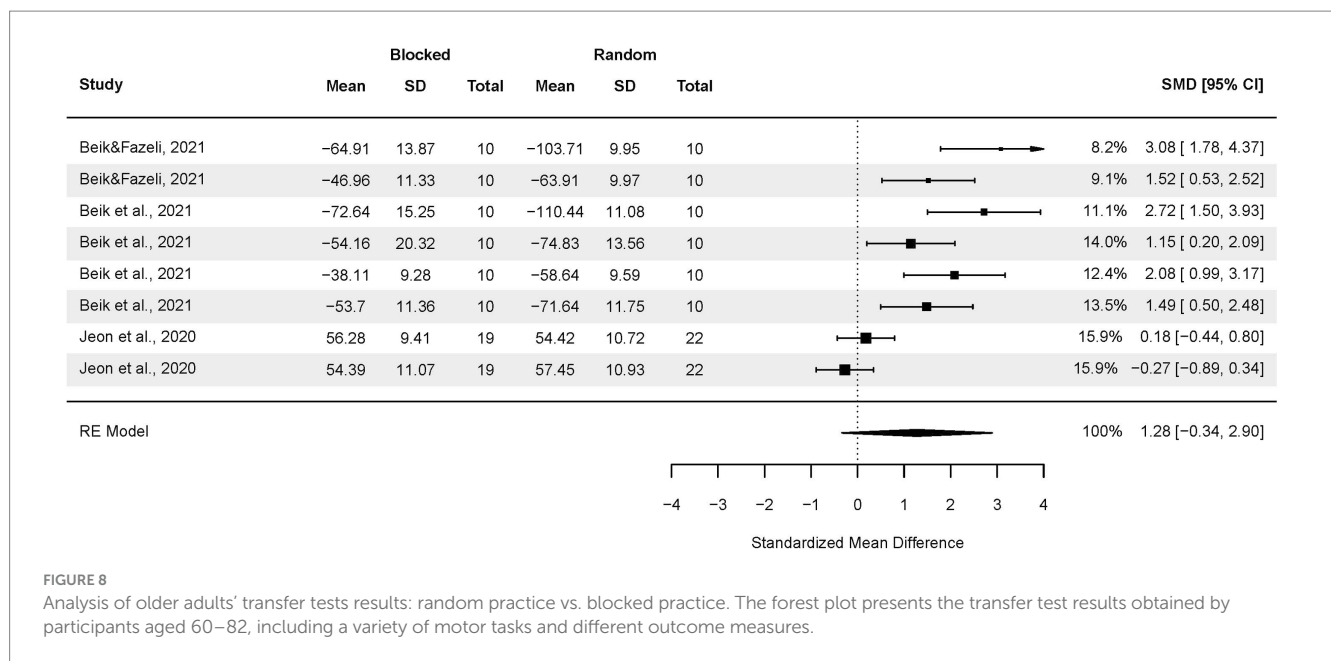


FIGURE 8 Analysis of older adults' transfer tests results: random practice vs. blocked practice. The forest plot presents the transfer test results obtained by participants aged 60–82, including a variety of motor tasks and different outcome measures.

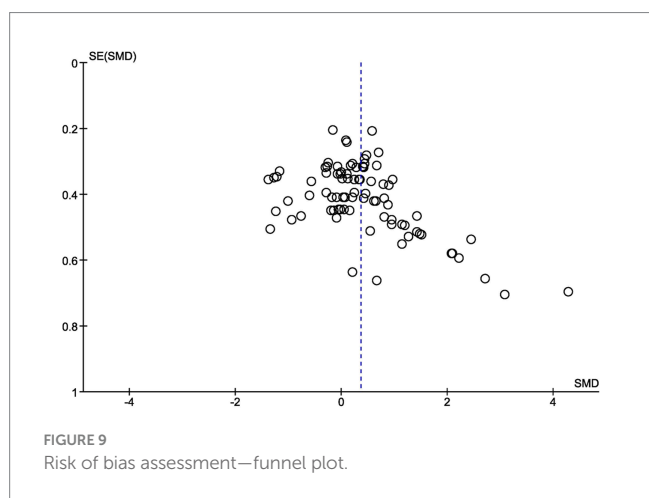


FIGURE 9 Risk of bias assessment—funnel plot.

Unfortunately, data from 23 primary studies included in his meta-analysis were unavailable. On the other hand, 27 primary studies included in our meta-analysis (yielding 69 effect sizes) were not included in Brady's.

The overall results of our review partially corresponded with those reported in the meta-analysis by Brady (2004). In line with the constantly advancing methodology of conducting meta-analyses, the inclusion criteria implemented in this review were more thoroughly detailed than those presented in Brady's. Consequently, 13 following studies included in Brady's meta-analysis (Brady, 2004) were considered irrelevant in the present review and, therefore, excluded. The primary studies by Wulf and Lee (1993), Sekiya et al. (1994, 1996), Landin and Hebert (1997), and Sekiya and Magill (2000) described serial practice order instead of random schedule.

In the studies by Wrisberg and Liu (1991), Hebert et al. (1996), Smith (2002), and Smith et al. (2003) alternating practice instead of a random schedule was presented. In the article by Hall and Magill (1995), an experiment described by Lee et al. (1992), and a study by Shea and Titzer (1993), multiple task learning was implemented. In the article by Bortoli et al. (2001), included in Brady's meta-analytic study, constant

and variable practice schedules were compared. Additionally, 15 studies did not include transfer tests. Three other studies described immediate transfer; therefore, our meta-analysis did not include the results.

Our results are different from those reported by Ammar et al. (2023). We found that the pooled effect size was medium (SMD=0.55) and statistically significant, while Ammar et al. reported small (SMD=0.243) and non-significant. In both studies, random practice was favored. Probably the differences between these studies may be attributed to the search strategies, number of studies and effects sizes included in both meta-analysis. Ammar et al. omitted EBSCO database (including APA PsycArticles, APA PsychInfo, SPORTDiscus with Full Text, Medline, and Academic Search Complete), specifying their focus differently (sport-specific). They finally included 16 studies and 38 ES referring to sport settings whereas in our meta-analysis 34 studies and 86 ES were included in general analysis and 17 studies and 40 ES in applied studies. Similarly, the differences in subgroups analysis can be explained in methods applied.

Low quality and bias problem

The studies included in our analysis were of poor quality in general. None of the included studies was assessed as strong. Quality of two articles were assessed moderate. Hence, one could re-state our question if the studies about CI effect are biased? The question is justified given 26 studies scored weak on the Selection Bias criteria. One of the possible explanations of the weak scores, in general, could be the tool we used, i.e., Quality Assessment Tool for Quantitative Studies (Thomas et al., 2004), which is rather strict. However, another explanation could be that the researchers' bias toward specific results in laboratory setting affected the final effect size favoring random practice.

One of the possibilities to decrease the effect of low-quality studies the meta-analysis result would be to exclude all studies rated as weak. This is what Thomas and colleagues postulated (Thomas et al., 2004). However, it would leave our analysis with only two papers! Therefore, we decided to include all the studies, though it could have impacted the heterogeneity statistics, increasing the I² value.

Heterogeneity problem

In all our analyses, the heterogeneity was substantial (with $I^2 > 70\%$). There may be a few explanations for why the heterogeneity was so high. Firstly, we included many studies and outcomes: more studies and outcomes, the higher the I^2 value (Rücker et al., 2008). One could decrease the I^2 value by limiting the number of studies included in the analysis and including only those with fewer participants (Schroll et al., 2011). However, this would question the validity of the presented review because a low I^2 value is not necessarily linked with a lower probability of heterogeneity but may be linked to the lower sensibility of detecting it (Schroll et al., 2011). Alba et al. (2016) noted “ I^2 can also mislead in large studies with precise results in which a low degree of inconsistency (i.e., studies report similar point estimates) can nevertheless result in high I^2 ” (p. 134). They added that we are not able to do much about it. Moreover, there is little advice for the researchers on how to do it (Schroll et al., 2011).

Secondly, we included studies of low quality (see Low quality and bias problem section). Thirdly, the source of high heterogeneity may be linked with the differences in populations, such as age and origin, followed by a variety of included motor tasks and outcome measures or different methodologies used, e.g., experiment duration.

Lastly, we used I^2 -test to assess the heterogeneity level. Unlike the Q-test (used by Brady in 2004), I^2 index not only informs about the presence or absence of heterogeneity but also quantifies its magnitude (Huedo-Medina et al., 2006).

Limitations

Limitations of our analysis have to be acknowledged. Firstly, 34 included studies yielded 86 effect sizes. As a result, some of the studies contributed multiple effect sizes. We treated them as independent, similarly to Toth et al. (2020). Thus, the combined significance values have to be interpreted with caution since they may be inflated of combined probability level.

Secondly, we tried to update Brady’s meta-analysis (Brady, 2004), nevertheless, we failed at obtaining all results (outcomes) Brady included. These not-included results could affect the overall effect of our analysis.

Thirdly, we advanced our objectives based on Brady’s. One could perform more specific analyses (more specific PICOs) that could lead to different results.

Recommendation for future research

Given that studies on CI effect on retention and transfer, were mostly of poor quality, a strong emphasis has to be put on the methodological issues.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author/s.

Author contributions

SC: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. AW: Data curation, Formal analysis, Investigation, Resources, Validation, Writing – original draft, Writing – review & editing, Methodology. PS: Data curation, Investigation, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The work was supported by the grant project with registration number MUNI/A/1470/2023 at Masaryk University Brno, Faculty of Sports Studies.

Acknowledgments

We would like to thank all the authors who responded to our e-mails or messages and provided us with more information about their studies. We would like to thank Piotr Szulc, who helped us with the R analysis.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2024.1377122/full#supplementary-material>

All appendices are available at: https://osf.io/tqchf/?view_only=663162ef579347509b632b971b503eba

SUPPLEMENTARY TABLE S1
Summary of the included studies.

SUPPLEMENTARY TABLE S2
Characteristics of the tasks, participants, tests, and outcomes of the included studies.

References

- Alba, A. C., Alexander, P. E., Chang, J., Macisaac, J., Defry, S., and Guyatt, G. H. (2016). High statistical heterogeneity is more frequent in meta-analysis of continuous than binary outcomes. *J. Clin. Epidemiol.* 70, 129–135. doi: 10.1016/j.jclinepi.2015.09.005
- Al-Mustafa, A. A. (1989). Contextual interference: laboratory artifact or sport skill learning related. Unpublished dissertation: University of Pittsburgh.
- Ammar, A., Trabelsi, K., Boujelbane, M. A., Boukhris, O., Glenn, J. M., Chtourou, H., et al. (2023). The myth of contextual interference learning benefit in sports practice: a systematic review and meta-analysis. *Educ. Res. Rev.* 39:100537. doi: 10.1016/j.edurev.2023.100537
- Assink, M., and Wibbelink, C. J. M. (2016). Fitting three-level meta-analytic models in R: a step-by-step tutorial. *Quant. Methods Psychol.* 12, 154–174. doi: 10.20982/tqmp.12.3.p154
- Battig, W. F. (1966). Facilitation and interference. In *Acquisition of skill*. ed. E. A. Bilodeau (New York: Academic Press), 215–244.
- Battig, W. F. (1972). "Intraskill interference as a source of facilitation in transfer and retention" in *Topics in learning and performance*. eds. R. F. Thompson and J. F. Voss (New York: Academic Press), 131–159.
- Beik, M., and Fazeli, D. (2021). The effect of learner-adapted practice schedule and task similarity on motivation and motor learning in older adults. *Psychol. Sport Exerc.* 54:101911. doi: 10.1016/j.psychsport.2021.101911
- Beik, M., Taheri, H., Saberi Kakhki, A., and Ghoshuni, M. (2021). Algorithm-based practice schedule and task similarity enhance motor learning in older adults. *J. Mot. Behav.* 53, 458–470. doi: 10.1080/00222895.2020.1797620
- Bortoli, L., Robazza, C., Durigon, V., and Carra, C. (1992). Effects of contextual interference on learning technical sports skills. *Percept. Mot. Skills* 75, 555–562. doi: 10.2466/pms.1992.75.2.555
- Bortoli, L., Spagolla, G., and Robazza, C. (2001). Variability effects on retention of a motor skill in elementary school children. *Percept. Mot. Skills* 93, 51–63. doi: 10.2466/pms.2001.93.1.51
- Brady, F. (2004). Contextual interference: a meta-analytic study. *Percept. Mot. Skills* 99, 116–126. doi: 10.2466/pms.99.1.116-126
- Broadbent, D. P., Causer, J., Ford, P. R., and Williams, A. M. (2015). Contextual interference effect on perceptual-cognitive skills training. *Med. Sci. Sports Exerc.* 47, 1243–1250. doi: 10.1249/MSS.0000000000000530
- Cheong, J. P. G., Lay, B., and Razman, R. (2016). Investigating the contextual interference effect using combination sports skills in open and closed skill environments. *J. Sports Sci. Med.* 15, 167–175
- Cheung, M. W. L. (2014). Modeling dependent effect sizes with three-level meta-analyses: a structural equation modeling approach. *Psychol. Methods* 19, 211–229. doi: 10.1037/A0032968
- Chua, L. K., Dimapilis, M. K., Iwatsuki, T., Abdollahipour, R., Lewthwaite, R., and Wulf, G. (2019). Practice variability promotes an external focus of attention and enhances motor skill learning. *Hum. Mov. Sci.* 64, 307–319. doi: 10.1016/j.humov.2019.02.015
- Cohen, J. (1988). "Statistical power analysis for the behavioral sciences" in *Statistical power analysis for the behavioral sciences*. 2nd ed (New York: Routledge).
- Cumpston, M., and Chandler, J. (2019). "Chapter IV: updating a review" in *Cochrane handbook for systematic reviews of interventions version 6.0 (version 6)*. Cochrane. eds. J. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li and M. Page, et al. Available at: <https://training.cochrane.org/handbook/current/chapter-iv>
- Czyż, S. H., Wójcik, A. M., Solarská, P., and Kiper, P. (2023). High contextual interference improves retention in motor learning: systematic review and meta-analysis. *Scientific Reports*. 14:15974.
- De Souza, M. G. T. X., Nunes, M. E. S., Corrêa, U. C., and Dos Santos, S. (2015). The contextual interference effect on sport-specific motor learning in older adults. *Hum. Mov.* 16, 112–118. doi: 10.1515/HUMO-2015-0036
- Deeks, J. J., Higgins, J. P. T., and Altman, D. G. (2019). "Analysing data and undertaking meta-analyses" in *Cochrane handbook for systematic reviews of Interventions*. eds. J. P. T. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li and M. J. Page et al. (Wiley Online Library), 241–284.
- Del Rey, P., Liu, X., and Simpson, K. J. (1994). Does retroactive inhibition influence contextual interference effects? *Res. Q. Exerc. Sport* 65, 120–126. doi: 10.1080/02701367.1994.10607606
- Del Rey, P., Whitehurst, M., and Wood, J. M. (1983a). Effects of experience and contextual interference on learning and transfer by boys and girls. *Percept. Mot. Skills* 56, 581–582. doi: 10.2466/PMS.1983.56.2.581
- Del Rey, P., Whitehurst, M., Wughalter, E., and Barnwell, J. (1983b). Contextual interference and experience in acquisition and transfer. *Percept. Mo* 57, 241–242. doi: 10.2466/PMS.1983.57.1.241
- Diekelmann, S., and Born, J. (2010). The memory function of sleep. *Nat. Rev. Neurosci.* 11, 114–126. doi: 10.1038/nrn2762
- Dundar, Y., and Fleeman, N. (2017). "Applying inclusion and exclusion criteria" in *Doing a systematic review. a student's guide*. eds. A. Boland, G. M. Cherry and R. Dickson (SAGE), 79–91.
- Dunham, P. Jr., Lemke, M., and Moran, P. (1991). Effect of equal and random amounts of varied practice on transfer task performance. *Percept. Mot. Skills* 73, 673–674. doi: 10.2466/PMS.1991.73.2.673
- Fialho, J. V. A. P., Benda, R. N., and Ugrinotitsch, H. (2006). The contextual interference effect in a serve skill acquisition with experienced volleyball players. *J. Hum. Mov. Stud.* 50, 65–77.
- Frömer, R., Stürmer, B., and Sommer, W. (2016). Come to think of it: contributions of reasoning abilities and training schedule to skill acquisition in a virtual throwing task. *Acta Psychol.* 170, 58–65. doi: 10.1016/j.actpsy.2016.06.010
- Goodwin, J. E., and Meeuwse, H. J. (1996). Investigation of the contextual interference effect in the manipulation of the motor parameter of over-all force. *Percept. Mot. Skills* 83, 735–743. doi: 10.2466/pms.1996.83.3.735
- Graser, J. V., Bastiaenen, C. H. G., and van Hedel, H. J. A. (2019). The role of the practice order: a systematic review about contextual interference in children. *PLoS One* 14:e0209979. doi: 10.1371/journal.pone.0209979
- Green, S., and Sherwood, D. E. (2000). The benefits of random variable practice for accuracy and temporal error detection in a rapid aiming task. *Res. Q. Exerc. Sport* 71, 398–402. doi: 10.1080/02701367.2000.10608922
- Gunnell, K., Poitras, V. J., and Tod, D. (2020). Questions and answers about conducting systematic reviews in sport and exercise psychology. *Int. Rev. Sport Exerc. Psychol.* 13, 297–318. doi: 10.1080/1750984X.2019.1695141
- Hall, K. G., Domingues, D. A., and Cavazos, R. (1994). Contextual interference effects with skilled baseball players. *Percept. Mot. Skills* 78, 835–841. doi: 10.1177/003151259407800331
- Hall, K. G., and Magill, R. A. (1995). Variability of practice and contextual interference in motor skill learning. *J. Mot. Behav.* 27, 299–309. doi: 10.1080/00222895.1995.9941719
- Hebert, E. P., Landin, D., and Solmon, M. A. (1996). Practice schedule effects on the performance and learning of low- and high-skilled students: an applied study. *Res. Q. Exerc. Sport* 67, 52–58. doi: 10.1080/02701367.1996.10607925
- Herzog, M., Focke, A., Maurus, P., Thüerer, B., and Stein, T. (2022). Random practice enhances retention and spatial transfer in force field adaptation. *Front. Hum. Neurosci.* 16:816197. doi: 10.3389/fnhum.2022.816197
- Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., et al. (2021). "Cochrane handbook for systematic reviews of interventions[Cochrane training" in *Cochrane handbook for systematic reviews of interventions version 6.2 (updated February 2021)* (Cochrane).
- Huedo-Medina, T., Sanchez-Meca, J., Marin-Martinez, F., and Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I2 index? *Psychol. Methods* 11, 193–206. doi: 10.1037/1082-989X.11.2.193
- Jackson, J. L., and Kuriyama, A. (2019). How often do systematic reviews exclude articles not published in English? *J. Gen. Intern. Med.* 34, 1388–1389. doi: 10.1007/s11606-019-04976-x
- Jeon, M. J., Jeon, H. S., Yi, C. H., Kwon, O. Y., You, S. H., and Park, J. H. (2020). Block and random practice: a Wii fit dynamic balance training in older adults. *Res. Q. Exerc. Sport* 92, 352–360. doi: 10.1080/02701367.2020.1733456
- Johnson, G. G. R. J., Park, J., Vergis, A., Gillman, L. M., and Rivard, J. D. (2022). Contextual interference for skills development and transfer in laparoscopic surgery: a randomized controlled trial. *Surg. Endosc.* 36, 6377–6386. doi: 10.1007/S00464-021-08946-5
- Lage, G. M., Faria, L. O., Ambrósio, N. F. A., Borges, A. M. P., and Apolinário-Souza, T. (2021). What is the level of contextual interference in serial practice? A meta-analytic review. *J. Motor Learn. Dev.* 10, 224–242. doi: 10.1123/jmld.2021-0020
- Lage, G. M., Vieira, M. M., Palhares, L. R., Ugrinowitsch, H., and Benda, R. N. (2006). Practice schedules and number of skills as contextual interference factors in the learning of positioning timing tasks. *J. Hum. Mov. Stud.* 50, 185–200.
- Landin, D., and Hebert, E. P. (1997). A comparison of three practice schedules along the contextual interference continuum. *Res. Q. Exerc. Sport* 68, 357–361. doi: 10.1080/02701367.1997.10608017
- Lee, T. D., Wulf, G., and Schmidt, R. A. (1992). Contextual interference in motor learning: dissociated effects due to the nature of task variations. *Q. J. Exp. Psychol. A* 44, 627–644. doi: 10.1080/14640749208401303
- Lin, C. H., Yang, H. C., Knowlton, B. J., Wu, A. D., Iacoboni, M., Ye, Y. L., et al. (2018). Contextual interference enhances motor learning through increased resting brain connectivity during memory consolidation. *Neuroimage* 181, 1–15. doi: 10.1016/j.neuroimage.2018.06.081
- Luo, D., Wan, X., Liu, J., and Tong, T. (2018). Optimally estimating the sample mean from the sample size, median, mid-range, and/or mid-quartile range. *Stat. Methods Med. Res.* 27, 1785–1805. doi: 10.1177/0962280216669183
- Magill, R., and Anderson, D. (2021). *Motor learning and control: concepts and applications*. 12th Edn. New York: McGraw-Hill.

- Mahood, Q., Van Eerd, D., and Irvin, E. (2014). Searching for grey literature for systematic reviews: challenges and benefits. *Res. Synth. Methods* 5, 221–234. doi: 10.1002/JRSM.1106
- Meira, C. M., Fairbrother, J. T., and Perez, C. R. (2015). Contextual interference and introversion/extraversion in motor learning. *Percept. Mot. Skills* 121, 447–460. doi: 10.2466/23.PMS.121C20X6
- Meira, C. M., and Tani, G. (2001). The contextual interference effect in Acquisition of Dart-Throwing Skill Tested on a transfer test with extended trials. *Percept. Mot. Skills* 92, 910–918. doi: 10.2466/pms.2001.92.3.910
- Meira, C. M., and Tani, G. (2003). Contextual interference effects assessed by extended transfer trials in the acquisition of volleyball serve. *J. Hum. Mov. Stud.* 45, 449–468.
- Methley, A. M., Campbell, S., Chew-Graham, C., McNally, R., and Cheraghi-Sohi, S. (2014). PICO, PICOS and SPIDER: a comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. *BMC Health Serv. Res.* 14, 1–10. doi: 10.1186/s12913-014-0579-0
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., Antes, G., Atkins, D., et al. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann. Intern. Med.* 151, 264–269. doi: 10.7326/0003-4819-151-4-200908180-00135
- Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., et al. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Rev. Espan. Nutr. Hum. Diet.* 4, 148–160. doi: 10.1186/2046-4053-4-1
- Moretto, N. A., Marcori, A. J., and Okazaki, V. H. A. (2018). Contextual interference effects on motor skill acquisition, retention and transfer in sport rifle shooting. *Hum. Mov.* 19, 99–104. doi: 10.5114/hm.2018.74065
- Morrison, A., Polisen, J., Huseraue, D., Moulton, K., Clark, M., Fiander, M., et al. (2012). The effect of English-language restriction on systematic review-based meta-analyses: a systematic review of empirical studies. *Int. J. Technol. Assess. Health Care* 28, 138–144. doi: 10.1017/S0266462312000086
- Page, M. J., and Moher, D. (2017). Evaluations of the uptake and impact of the preferred reporting items for systematic reviews and Meta-analyses (PRISMA) statement and extensions: a scoping review. *Syst. Rev.* 6:263. doi: 10.1186/s13643-017-0663-8
- Parab, S., Bose, M., and Ganesan, S. (2018). Influence of random and blocked practice schedules on motor learning in children aged 6–12 years. *Crit. Rev. Phys. Rehabil. Med.* 30, 239–254. doi: 10.1615/CRITREVPHYSREHABILMED.2018027737
- Pasand, F., Fooladiyanzadeh, H., and Nazemzadegan, G. (2016). The effect of gradual increase in contextual interference on acquisition, retention and transfer of volleyball Skills on acquisition, retention and transfer of volleyball skills. *Int. J. Kinesiol. Sports Sci.* 4, 72–77. doi: 10.7575/aiac.ijks.v.4n.2p.72
- Perez, C. R., Meira, C. M., and Tani, G. O. (2005). Does the contextual interference effect last over extended transfer trials? *Percept. Mot. Skills* 100, 58–60. doi: 10.2466/PMS.100.1.58-60
- Porter, J. M., and Beckerman, T. (2016). Practicing with gradual increases in contextual interference enhances visuomotor learning. *Kinesiology* 48, 244–250. doi: 10.26582/K.48.2.5
- Porter, C., Greenwood, D., Panchuk, D., and Pepping, G. J. (2020). Learner-adapted practice promotes skill transfer in unskilled adults learning the basketball set shot. *Eur. J. Sport Sci.* 20, 61–71. doi: 10.1080/17461391.2019.1611931
- Porter, J. M., and Magill, R. A. (2010). Systematically increasing contextual interference is beneficial for learning sport skills. *J. Sports Sci.* 28, 1277–1285. doi: 10.1080/02640414.2010.502946
- Porter, J. M., and Saemi, E. (2010). Moderately skilled learners benefit by practicing with systematic increases in contextual interference. *Int. J. Coach. Sci.* 4, 61–71.
- Raviv, L., Lupyan, G., and Green, S. C. (2022). How variability shapes learning and generalization. *Trends Cogn. Sci.* 26, 462–483. doi: 10.1016/J.TICS.2022.03.007
- Rücker, G., Schwarzer, G., Carpenter, J. R., and Schumacher, M. (2008). Undue reliance on I2 in assessing heterogeneity may mislead. *BMC Med. Res. Methodol.* 8, 1–9. doi: 10.1186/1471-2288-8-79/TABLES/3
- Sattelmayer, M., Elsig, S., Hilfiker, R., and Baer, G. (2016). A systematic review and meta-analysis of selected motor learning principles in physiotherapy and medical education. *BMC Med. Educ.* 16:15. doi: 10.1186/s12909-016-0538-z
- Schroll, J. B., Moustgaard, R., and Gotzsche, P. C. (2011). Dealing with substantial heterogeneity in Cochrane reviews. Cross-sectional study. *BMC Med. Res. Methodol.* 11, 1–8. doi: 10.1186/1471-2288-11-22/FIGURES/5
- Sekiya, H., and Magill, R. A. (2000). The contextual interference effect in learning force and timing parameters of the same generalized program. *J. Hum. Mov. Stud.* 39, 45–71.
- Sekiya, H., Magill, R. A., and Anderson, D. I. (1996). The contextual interference effect in parameter modifications of the same generalized motor program. *Res. Q. Exerc. Sport* 67, 59–68. doi: 10.1080/02701367.1996.10607926
- Sekiya, H., Magill, R. A., Sidaway, B., and Anderson, D. I. (1994). The contextual interference effect for skill variations from the same and different generalized motor programs. *Res. Q. Exerc. Sport* 65, 330–338. doi: 10.1080/02701367.1994.10607637
- Shea, C. H., Lai, Q., Wright, D. L., Immink, M., and Black, C. (2001). Consistent and variable practice conditions: effects on relative and absolute timing. *J. Mot. Behav.* 33, 139–152. doi: 10.1080/00222890109603146
- Shea, J. B., and Morgan, R. L. (1979). Contextual interference effects on the acquisition, retention, and transfer of a motor skill. *J. Exp. Psychol. Hum. Learn. Mem.* 5, 179–187. doi: 10.1037/0278-7393.5.2.179
- Shea, J. B., and Titzer, R. C. (1993). The influence of reminder trials on contextual interference effects. *J. Mot. Behav.* 25, 264–274. doi: 10.1080/00222895.1993.9941647
- Sherwood, D. E., and Duffel, B. (2010). Concurrent visual feedback, practice organization, and spatial aiming accuracy in rapid movement sequences. *Int. J. Exerc. Sci.* 3, 78–91
- Shewokis, P. A. (1997). Is the contextual interference effect generalizable to computer games? *Percept. Mot. Skills* 84, 3–15. doi: 10.2466/PMS.1997.84.1.3
- Shewokis, P. A., Shariff, F. U., Liu, Y., Ayaz, H., Castellanos, A., and Lind, D. S. (2017). Acquisition, retention and transfer of simulated laparoscopic tasks using fNIR and a contextual interference paradigm. *Am. J. Surg.* 213, 336–345. doi: 10.1016/J.AMJSURG.2016.11.043
- Shi, J., Luo, D., Wan, X., Liu, Y., Liu, J., Bian, Z., et al. (2020a). Detecting the skewness of data from the sample size and the five-number summary. *ArXiv*. doi: 10.48550/arxiv.2010.05749
- Shi, J., Luo, D., Weng, H., Zeng, X.-T., Lin, L., Chu, H., et al. (2020b). Optimally estimating the sample standard deviation from the five-number summary. *Res. Synth. Methods* 11, 641–654. doi: 10.1002/jrsm.1429
- Shojania, K. G., Sampson, M., Ansari, M. T., Ji, J., Doucette, S., and Moher, D. (2007). How quickly do systematic reviews go out of date? A survival analysis. *Ann. Intern. Med.* 147, 224–233. doi: 10.7326/0003-4819-147-4-200708210-00179
- Smith, P. J. K. (2002). Applying contextual interference to snowboarding skills. *Percept. Mot. Skills* 95, 999–1005. doi: 10.2466/pms.2002.95.3.999
- Smith, P. J. K., and Davies, M. (2008). Applying contextual interference to the Pawlata roll. *J. Sport Sci.* 13, 455–462. doi: 10.1080/02640419508732262
- Smith, P. J. K., Gregory, S. K., and Davies, M. (2003). Alternating versus blocked practice in learning a cartwheel. *Percept. Mot. Skills* 96, 1255–1264. doi: 10.2466/pms.2003.96.3c.1255
- Smith, P. J. K., and Rudisill, M. E. (1993). The influence of proficiency level, transfer distance, and gender on the contextual interference effect. *Res. Q. Exerc. Sport* 64, 151–157. doi: 10.1080/02701367.1993.10608792
- Ste-Marie, D. M., Clark, S. E., Findlay, L. C., and Latimer, A. E. (2010). High levels of contextual interference enhance handwriting skill acquisition. *J. Mot. Behav.* 36, 115–126. doi: 10.3200/JMBR.36.1.115-126
- Thomas, H. (2003). Quality assessment tool for quantitative studies: effective public health practice project. Hamilton, ON: McMaster University.
- Thomas, H., Ciliska, D., Dobbins, M., and Micucci, S. (2004). A process for systematically reviewing the literature: providing the research evidence for public health nursing interventions. *Worldviews Evid.-Based Nurs.* 1, 176–184. doi: 10.1111/J.1524-475X.2004.04006.X
- Thomas, J. L., Fawver, B., Taylor, S., Miller, M. W., Williams, A. M., and Lohse, K. R. (2021). Using error-estimation to probe the psychological processes underlying contextual interference effects. *Hum. Mov. Sci.* 79:102854. doi: 10.1016/J.HUMOV.2021.102854
- Toth, A. J., McNeill, E., Hayes, K., Moran, A. P., and Campbell, M. (2020). Does mental practice still enhance performance? A 24 year follow-up and meta-analytic replication and extension. *Psychol. Sport Exerc.* 48:101672. doi: 10.1016/J.PSYCHSPORT.2020.101672
- Travlos, A. K. (2010). Specificity and variability of practice, and contextual interference in acquisition and transfer of an underhand volleyball serve. *Percept. Mot. Skills* 110, 298–312. doi: 10.2466/PMS.110.1.298-312
- Vera, J. G., and Montilla, M. M. (2003). Practice schedule and acquisition, retention, and transfer of a throwing task in 6-YR.-old children. *Percept. Mot. Skills* 96, 1015–1024. doi: 10.2466/PMS.2003.96.3.1015
- Wrisberg, C. A., and Liu, Z. (1991). The effect of contextual variety on the practice, retention, and transfer of an applied motor skill. *Res. Q. Exerc. Sport* 62, 406–412. doi: 10.1080/02701367.1991.10607541
- Wulf, G., and Lee, T. D. (1993). Contextual interference in movements of the same class: differential effects on program and parameter learning. *J. Mot. Behav.* 25, 254–263. doi: 10.1080/00222895.1993.9941646
- Wulf, G., and Shea, C. H. (2002). Principles derived from the study of simple skills do not generalize to complex skill learning. *Psychon. Bull. Rev.* 9, 185–211. doi: 10.3758/BF03196276
- Yang, G., Lai, C. S. W., Cichon, J., Ma, L., Li, W., and Gan, W. B. (2014). Sleep promotes branch-specific formation of dendritic spines after learning. *Science* 344, 1173–1178. doi: 10.1126/SCIENCE.1249098