# Rebuttal to Ammar and Schöllhorn on "*Comment on Czyż et al. (2024) on Contextual Interference in Motor Learning*"

**Stanisław H. Czyż [1,2,3*], Aleksandra M. Wójcik[1]**

[1] Faculty of Physical Education and Sports, Wroclaw University of Health and Sport Sciences, Wroclaw, Poland

[2] Faculty of Sport Studies, Masaryk University, Brno, Czechia

[3] Physical Activity, Sport and Recreation (PhASRec), Faculty of Health Sciences, North-West University, South Africa

*Corresponding author: Stanisław H. Czyż, stachu.czyz@gmail.com

**Example citation:**
Czyż, S.H., Wójcik, A.M, (2024). Rebuttal to Ammar and Schöllhorn Comment on Czyż et al. (2024) on Contextual Interference in Motor Learning. Preprint DOI: …

**THIS A PREPRINT**

**Abstract:**
Given that the contextual interference phenomenon is one of the topics covered the most in motor learning studies, books, and guides for practitioners, it was not surprising that the meta-analysis on this subject would raise discussions and doubts. In this paper, we address the comments raised by Ammar and Schöllhorn, who asked us about specific issues related to our paper. We address their comments and shortly compare our meta-analysis with the one conducted by Ammar et al. (Ammar et al., 2023).

**Key words:** contextual interferences, retention, motor learning, meta-analysis, systematic review.

**Introduction**

We were very thankful to see the comments by Ammar and Schöllhorn (Ammar & Schöllhorn, 2024) on our manuscript about the effect of high contextual interferences on retention in motor learning (Czyż et al., 2024). They made us re-think our reasoning once more. Given that critical thinking and skepticism are inherent in science, we read the critical comments with attention and humility. We address all of them below.

We are firmly convinced that cognizant readers will formulate their opinions by reading Ammar et al. (2023) and our systematic reviews, their comments, and our answers.

# REBUTTAL

## COMMENT:

- Title

In the abstract, the authors report: "We found that the random practice schedule in laboratory settings effectively improved motor skills retention. On the contrary, in the applied setting, the beneficial effect of random practice on retention was almost negligible." However, the title of the paper, "High contextual interference improves retention in motor learning: systematic review and meta-analysis," uses an affirmative and generalized statement. This is misleading. In its current form, the title suggests that high CI improves retention in motor learning in general, encompassing both laboratory and applied settings. Nevertheless, the results of the MA do not support this statement.

## RESPONSE:

According to both analyses we applied, high CI has a statistically significant moderate beneficial effect on retention in the entire population (see Figures 2 and 3). Moreover, in all subgroups meta-analyses, random practice was favored. Perhaps the authors of this comment are considering the extent of the CI benefits. Indeed, random practice benefitted retention to varying degrees; in some cases, the effect size was large, medium, small, or negligible in others. However, random order was always favored. For example, in applied studies, SMD = 0.23 (three-level mixed model) and SMD = 0.35 (random-effects model with averaged SMDs). Blocked practice did not favor retention at all. Our title does not refer to the extent of the CI effect and is, therefore, valid.

## COMMENT:

- Information related to Graser et al. Ammar et al. publications

In the introduction, the authors stated, "the meta-analyses on CI conducted by Graser22 and Sattelmayer23 provided the inclusion criteria." However, it was raised with us that the study by Graser et al. (2019) was a systematic review and did not include a MA.

## RESPONSE:

Though a thought shortcut used when referring to other meta-analyses, this statement is justified since the (Graser et al., 2019) study was planned as a meta-analysis (eventually not performed).

Graser et al. (Graser et al., 2019) planned to perform a meta-analysis *"We had planned to pool data when studies were comparable regarding populations, interventions, outcomes, and types of studies*" (see their section Data extraction and analysis).

They finally did not pool the effect sizes since "*Combining the study results by pooling the data in a meta-analysis was not appropriate since the studies were too heterogeneous considering the populations, types of motor tasks, intensities, time points (e.g. retention after five minutes, 24 hours or three weeks), and outcome measures. We also refrained from a subgroup analysis due to the low methodological quality and too small sample sizes of studies with sufficient relevant similarities."* (see their section Risk of bias).

This thought shortcut does not change anything since we do no elaborate more on Graser et al. study (Graser et al., 2019) in our paper and mentioned their study as an example of studies that provided clear inclusion criteria.

# COMMENT:

> Additionally, in regard to Ammar et al: "In 2023, Ammar and colleagues published their meta-analysis. Unfortunately, it was poorly performed. For example, they searched the Taylor and Francis database, though it is a publisher base, not a scientific one. At the same time, they did not screen the EBSCO database, which consists, among others, of APA PsycArticles, APA PsychInfo, SPORTDiscus with Full Text, Medline, and Academic Search Complete. Their review was not preregistered, which is a standard procedure these days. Given these methodological flaws, the review of Ammar et al. cannot be considered reliable and valid." However, it was raised with us that this argument is misleading and inaccurate, as Taylor and Francis was one of five databases searched (fully: PubMed, Web of Science, Scopus, Taylor and Francis, and SciELO). Furthermore, a multi-database approach is in line with typical practice suggesting that searching should include at least two databases, with multiple databases recommended. For example: a recent meta-research study by Ewad et al. (2022) concluded that searching two or more databases improves coverage and recall and decreases the risk of missing eligible studies.

# RESPONSE:

The number of databases searched is essential, but so is their relevance to the topic.

Theoretically, one can argue that they searched two databases while working on the CI effect in motor learning, for example, the Philosopher's Index with Full Text (EBSCO) and Hospitality & Tourism Complete (EBSCO). Formally, this satisfies the criterion of searching at least two databases. However, these databases probably do not include any papers on CI. Most studies on CI are published in psychological or sport-related journals. This theoretical search did not include the most relevant databases, such as SPORTDiscus with Full Text, APA PsycArticles, and APA PsycInfo. Omitting these databases is, in our opinion, a serious limitation of Ammar et al.'s study (Achraf Ammar et al., 2023).

- "*In terms of database selection, the specific database chosen can have important implications on the search breadth*." (Harari et al., 2020)
- "*However, make sure that you search only those (*databases- SC*) that are the most relevant to your review question and topic area*" (p. 66) (Dundar & Fleeman, 2017)

We intended to mention the limitations of Ammar et al.'s study (Achraf Ammar et al., 2023) though we did not want to list all of them. This is why we started our statement, "*For example, they searched the Taylor and Francis database …"* still, we never claimed they searched only one database. Our paper was not about limitations of Ammar et al. study.

## The longer list of Ammar' et al. (2023) limitations include e.g.:

1 – omitting EBSCO databases, with APA PsycArticles, APA PsychInfo, SPORTDiscus with Full Text, Medline, and Academic Search Complete in their search (discussed above)

2 – the review was not preregistered (discussed under the next Comment)

3 –the problem of multiple outcomes from one sample (dependency problem) was not addressed:

- "*A failure to handle within-study dependence can negatively impact the resulting inferences*" (p.152)(Park & Beretvas, 2019)
- "*Dependent effects sizes are less informative than independent effect sizes. Suppose that two outcome variables are perfectly correlated. Essentially, this means that both outcomes refer to the same latent variable and that effect sizes calculated for both outcomes will give exactly the same information. If, in a meta-analysis, both of these effect sizes are included as independent effect sizes, the same information therefore is used twice. In general, when outcome variables are correlated, information regarding one outcome overlaps with information yielded by the other outcome*." (p.577) (Van den Noortgate et al., 2013)

4 – sensitivity analysis was not performed:

- "*It is highly desirable to prove that the findings from a systematic review are not dependent on such arbitrary or unclear decisions by using sensitivity analysis (see MECIR Box 10.14.a). A sensitivity analysis is a repeat of the primary analysis or meta-analysis in which alternative decisions or ranges of values are substituted for decisions that were arbitrary or unclear*." (Deeks et al., 2023)

- *A sensitivity analysis is an important part of a meta-analysis as it aims to determine the robustness of the observed outcomes to the assumptions made in performing the analysis. Unfortunately this essential part of meta-analyses is often either not performed or reported*." (Bown & Sutton, 2010)

5 – Ammar et al. (Achraf Ammar et al., 2023) did not specify what are "sport practice" studies. As a result, they included or excluded studies that may or may not be relevant to their inclusion/exclusion criteria (a few examples are listed below).

3 - Ammar et al. assumed that constant practice equals blocked practice including Memmert study (Memmert, 2006). Unfortunately, it is not what CI effect describes since both blocked and random practice are variable practices (e.g., involving the practice of 3 skills, A, B, C). The only difference between random and blocked is the schedule (order). On the other hand, constant practice is not variable because it consists of practicing only one skill, e.g., A, B, or C. Therefore, including Memmert study in the Ammar et al. analysis is methodologically wrong.

4 – Ammar et al. (Achraf Ammar et al., 2023) included serial and alternating orders in their analysis, replacing purely random order if absent. Alternating practice is a practice schedule in which learners alternate between two or more tasks or variations of a task in a predictable pattern during a single practice session. This approach falls between blocked practice (where one skill is practiced repeatedly before moving to the next) and random practice (where skills are practiced in a completely randomized order). One can ask if such an approach is justified. Though some of the results on alternating order may be similar to random order, the mechanisms for both orders may not be the same. In an unpredictable random order, an individual cannot prepare a forthcoming Generalized Motor Program (GMP) prior to its execution because they do not know what will be performed. In contrast, in alternating order, an individual may prepare, as they learn the order of tasks performed. Without strong rationale and evidence that the mechanisms behind alternating and random practice are the same, in our opinion, equating them is unjustified.

5 – Ammar et al. included a study by Menayo et al. (Menayo et al., 2010), resulting in 16 outcomes (treated as independent), which utilizes simultaneous design. A simultaneous design provides a means of comparing two or more different treatments with an individual subject "*During the sessions, each tennis player executed a total of 192 repetitions of shots, distributed in blocked or random practice. The first 48 repetitions of each shot were executed in blocked practice, and the 144 remaining shots were executed in a random sequence*" (p. 665). Given each participant in this experiment practiced both: in random and blocked order or in random order only, the question is how Ammar et al. classified blocked group?

One could ask how many studies and outcomes could have left in Ammar's et al. meta-analysis, if the constant vs. blocked practice study, simultaneous design study, and even alternating/serial practice studies would have been excluded.

# COMMENT:

> Furthermore, PRISMA guidelines (2020) for conducting systematic reviews and MA encourage registration, but do not state that it is mandatory.

# RESPONSE:

Pre-registration, i.e., publishing details about a systematic review, though not obligatory (in some places it is), is strongly recommended (similarly to PRISMA, which is not obligatory either, though required by many) by authors and/or organizations specializing in meta-analyses. Well-recognized organizations specializing in systematic reviews require registration prior to the commencement of the study (they have their own registry). Examples include:

- Cochrane Institute "*Registering a systematic review involves publishing details related to the project prior to its commencement*." (*Early Career Professionals Network: Registering Protocols of Systematic Reviews | Cochrane Community*, n.d.)
- Campbell Collaboration "*A protocol (project plan) for the review is developed in advance and also undergoes peer review*." (*What is a systematic review? - The Campbell Collaboration*, n.d.)

On the other hand, registration has a few critical functions; specifically, it helps to avoid bias and replications:

- "*Registration is an important aspect of conducting a systematic review as it helps in reducing research waste, decrease bias and enhances the overall transparency of the work*" Cochrane Institute (*Early Career Professionals Network: Registering Protocols of Systematic Reviews | Cochrane Community*, n.d.)

- "*The goals of a registry of systematic reviews would be to avoid unnecessary duplication, allow better clarity around the conduct and analysis of reviews, avoid publication bias and selective reporting of outcome-related bias, promote collaboration and assist with prioritization*." (Straus & Moher, 2010)

- "*Registering the protocol reduces research bias, duplication of effort, resource waste, and provides greater transparency*" (Poklepović Peričić & Tanveer, 2019)

- "*The editors of Systematic Reviews believe that prospective registration of systematic reviews is an important development that will play a role in promoting transparency and avoiding bias that will ultimately serve to improve methodological standards*." (Stewart et al., 2012)

Had Ammar et al. (Achraf Ammar et al., 2023) registered their study, they would have noticed our protocol registered on 04th February 2021 (04th January 2021 first submission), the databases we included and search starting on 20 April 2020.

# COMMENT:

- Inclusion and exclusion criteria: Lack of equivalent baseline control

In the introduction, the authors mentioned, "…random practice condition hinders performance during acquisition, although it facilitates retention and transfer." This sentence suggests that a comprehensive assessment of CI effects requires focusing on both the first effect (suggested to be negative on acquisition) and the second effect (suggested to be positive on retention and transfer) of CI. The authors of the present MA elected to focus only on the second effect, particularly on pooling the results of the retention test in their analyses.

# RESPONSE:

The sentence cited explains (defines) the CI effect and does not indicate anything beyond that. As the title, abstract, methods, and discussion of our paper suggest, we focused exclusively on retention, similar to the approach taken by Brady (Brady, 2004), Lage et al. (Lage et al., 2021), or Sattelmayer (Sattelmayer et al., 2016).
This approach is also reflected in many experimental studies that focus solely on one of the aspects of the phenomenon, e.g., acquisition, retention, or transfer, or their combination.

Given the number of studies identified, we split our analysis into two publications: one dedicated to retention and the other solely to transfer. We did not focus on performance.

# COMMENT:

However, they did not control for the confounding variable of higher baseline performance in one group (e.g., random) compared to the other (e.g., blocked), which can partly explain the observed benefits of high CI in retention. Indeed, they did not exclude studies that failed to guarantee that the random and blocked groups were equivalent at baseline.

# RESPONSE:

Indeed, baseline imbalances may affect the pooled effect. Excluding the studies that do not report pre-test values is the most straightforward option, though not the most reliable. Ammar et al. should apply sensitivity analysis to estimate whether the excluded studies biased the pooled effect.

- "*For example, if the eligibility of some studies in the meta-analysis is dubious because they do not contain full details, sensitivity analysis may involve undertaking the meta-analysis twice: the first time including all studies and, second, including only those that are definitely known to be eligible*." (Deeks et al., 2023)
- "*Five general recommendations for dealing with missing data in Cochrane Reviews are as follows: 1) Whenever possible, contact the original investigators to request missing data. 2) Make explicit the assumptions of any methods used to address missing data: for example, that the data are assumed missing at random, or that missing values were assumed to have a particular value such as a poor outcome. 3) Follow the guidance in Chapter 8 to assess risk of bias due to missing outcome data in randomized trials. 4) Perform sensitivity analyses to assess how sensitive results are to reasonable changes in the assumptions that are made*" (Deeks et al., 2023).

In our study, baseline imbalances are considered in quality assessment. Analogically to Cochrane' guidelines, which, includes baseline imbalances in the risk of bias analysis "*The guidance for using the tool states that the presence of baseline imbalance could be assessed as part of the 'Other bias' domain. However, the guidance argues that imbalance arises primarily because of inadequate methods of randomisation (generation and/or concealment of the sequence) or through differential exclusions of participants (yielding incomplete outcome data) and so should be addressed via these standard items in the tool.*" (Corbett et al., 2014).

Given that "randomization" was a part of our quality assessment, we fulfilled the requirements for assessing baseline imbalance risk. Moreover, the quality of the papers on CI is generally low, as discussed in our paper in the Discussion Low Quality and Bias Problem section. Potential readers are aware that high heterogeneity and bias issues may affect the results. We addressed the heterogeneity problem by applying appropriate statistical models and performing sub-analyses.

The baseline imbalances may also be addressed in separate meta-analyses. For example, see e.g. Wewege et al. (Wewege et al., 2022). A meta-analysis on baseline characteristics in CI could be an exciting option, but it was not within the scope of our paper and would extend much beyond what is accepted by journals in terms of page, word, and figure limits. However, we performed a quality assessment using a commonly recognized and accepted tool.

# COMMENT:

- Discussion Section: Comparing incomparable findings and poor explainability

In the discussion, the authors summarize the main findings of their MA, reporting a medium pooled effect size (SMD = 0.63 and 0.71 when applying the three-level mixed model and the random-effects model, respectively). However, they found an effect size slightly above negligible (SMD = 0.23 in the three-level mixed model and SMD = 0.28 in the random-effects model) when analyzing data from applied studies, which turned out to be statistically insignificant, as they reported. When comparing these findings with previous literature, they state: "These results are different from those reported by Ammar and colleagues. We found that the pooled effect size was medium 0.63 while Ammar et al. reported small." This is not accurate as Ammar et al mainly focused on sports settings, and thus should be logically compared with the negligible effect size found in the current MA when only data from applied studies were included. Comparing effect sizes calculated based on pooled data vs. those based on sport contexts (Ammar et al., 2023) doesn't make sense, especially since the authors of the current study have already computed separate effect sizes for applied setting studies.

# RESPONSE:

Ammar et el. (Achraf Ammar et al., 2023) does not define "sports settings". Therefore, the paper lacks a clear framework, and in one place, they refer to "sport setting", "sport practice", "sports-based contexts", "sport skills", in other in inclusion criteria, they mention "sport/gross motor skill". One could ask whether walking is a sport-setting task. Are any studies on cognitive processes (e.g., scheduling the feedback in blocked or serial order) in sport setting or not?

A few examples include:

1 – Ammar et al. included in their meta-analyses on retention the following:

- Novel tasks on throwing and kicking tasks performed while chair-sitting (Pollatou et al., 1997). The question is how a novel task performed on a specially construed chair, relates to “sport settings”?
- Dynamic balance training using the Nintendo Wii Fit system (Jeon et al., 2021). One could ask how Nintendo training relate to sport?
- Study in which feedback rather than practice was manipulated (Krause et al., 2014) Is feedback manipulation the same as practice manipulation?

2 - Ammar et al. omitted studies which can be easily defined as sport-related (they are not included in their meta-analysis nor in their Appendix S2 Excluded full texts after a careful review of the results section). E.g.:

- The study included "*college and high school baseball pitchers throwing straight balls and breaking balls*". (Tsutsui et al., 2013)
- The study included throwing and jumping (Jiménez-Díaz et al., 2018), i.e., gross motor skills, similar to Pollatou's study (Pollatou et al., 1997) which was included.
- The task was rifle sport shooting (Moretto et al., 2018). Rifle shooting can be classified as a gross motor skill “ *Rifle shooting can be considered a complex motor skill due to the challenge it presents to motor control. As proposed by Tuller, Turvey and Fitch [9], when trying to aim a gun the subject’s body is not completely motionless. The complexity of this task, then, revolves mostly around controlling many degrees of freedom of the arm segment, although other body parts are also involved in a way that a minimal change of position in any joint angle should be corrected by a compensatory movement of another joint(s) to achieve a successful shot. In other words, a shooter must constrain his muscles and joints to behave as a single coordinative structure*” (p. 100)(Moretto et al., 2018).

# COMMENT:

The authors explain these incomparable results by criticizing the previous literature with the same argument relating to database searching: "Probably the differences we found may be attributed to the search strategies, number of studies, and effect sizes included in both MA. Ammar et al. omitted the EBSCO database (including APA PsycArticles, APA PsychInfo, SPORTDiscus with Full Text, Medline, and Academic Search Complete), searching a publisher database instead (Taylor and Francis)." However the Ammar article reports inclusion of 37 studies, with 29 included in the retention phase analysis. This largely exceeds the number of studies included (only 19) in the applied setting analysis of the current MA (Figure 5). As around 10 studies in the sports context were included in the study of Ammar et al. (2023) and not in the present analysis of the applied context, the

authors of the present study are encouraged to argue why this important number of studies were not included, especially since they can easily identify and include them by screening the studies included in Ammar et al. (2023) and Ammar et al. (2024).

# RESPONSE:

Ammar et al. (Achraf Ammar et al., 2023) included 28 studies in the analysis on retention, unless they prove that "Wrisberg a 1991" is different than "Wrisberg 1991" (Fig 4, Ammar et al., 2023).

Unlike suggested, we found **ALL** studies Ammar et al. refer to, and we classified them differently (as the laboratory), excluded them during the full screening (see Appendix 2), or excluded them before full paper screening.

Moreover, our analysis of applied studies included 24 studies yielding 105 effect sizes for a three-level mixed model. Ammar et al. (2023) included 28 studies and 84 outcomes. The difference is four studies, not "around ten".

The studies we did not include in the forest plot (Fig. 5 and 6) in our meta-analysis on applied settings and are included in the meta-analysis on retention by Ammar et al. are the following:

1. (Boyce & Del Rey, 1990) - the paper was excluded as we could not retrieve the full version. It is worth mentioning that the Journal of Human Movement Studies was published until 2006/7. There are no electronic versions of the journal. We contacted the author, i.e. Boyce, via ResearchGate on 5 October 2020. There was no response. We tried to get access via a few university libraries and their networks (including libraries in South Africa, Poland and Czechia), though, unsuccessfully.

2. (Wrisberg & Liu, 1991) – Authors used alternating instead of random practice. The study is mentioned in the Discussion section, where we compare our results to Brady's (2004). Alternating practice is a practice schedule in which learners alternate between two or more tasks or variations of a task in a predictable pattern during a single practice session. This approach falls between blocked practice (where one skill is practiced repeatedly before moving to the next) and random practice (where skills are practiced in a completely randomized order). The reason for exclusion is listed in Appendix 2 (https://osf.io/r59zs/?view_only=61397e4508384d13960936a556890962).

3. (Pollatou et al., 1997) - Unlike Ammar et al. we included the study by Pollatou in laboratory settings since they stated "*For this study two apparatuses were invented and constructed to measure the selected motor skills. The tasks were novel to ensure similar initial learning for all subjects*" (p.489). One could wonder how novel throwing and kicking tasks performed while chair-sitting relate to the sport-setting.

4. (Smith et al., 2003) - Smith et al. used alternating practice instead of random. The reason for exclusion is listed in Appendix 2 (https://osf.io/r59zs/?view_only=61397e4508384d13960936a556890962).

5. (Keller et al., 2006) - Keller et al. did not include a random group using a serial order practice instead. The study is mentioned in the Discussion section, where we compare our results to Brady's (2004). The reason for exclusion is listed in Appendix 2 (https://osf.io/r59zs/?view_only=61397e4508384d13960936a556890962).

6. (Memmert, 2006) - We excluded this study while abstract screening – there was no blocked group. Memmert used a constant group instead "The constant practice group took 160 shots from the free throw line, while the variable practice group took 160 shots from different positions around the restricted area." (see his abstract). Unfortunately, it is not what CI effect describes since both blocked and random practice are variable practices (e.g., involving the practice of 3 skills, A, B, C). The only difference between random and blocked is the schedule (order). On the other hand, constant practice is not variable because it consists of practicing only one skill, e.g., A, B, or C. Therefore, including Memmert study in the Ammar et al. analysis is methodologically wrong.

7. (Jones & French, 2007) - Authors did not respond to our request to specify the number of participants allocated to each tested group. As the authors stated, "*Sixty-eight ninth-grade students were recruited from three high school physical education classes to participate. Students within each class were randomly assigned to one of three experimental groups (blocked, random, and random-blocked). Fifty-one students, 5 boys and 46 girls, completed the entire acquisition and testing trials. Seventeen participants missed two or more acquisition sessions so their data were dropped from analysis*" (p.885). One could guess how many participants were finally allocated into each of the three groups (given the dropouts is not clarified), however, it is not how the meta-analysis should be done. Knowing the number of participants is an essential part of statistical analysis. We contacted the corresponding author, i.e., L. Jones at laurajones2@boisestate.edu on 29 April 2021. There was no response. The reason for exclusion is listed in Appendix 2 (https://osf.io/r59zs/?view_only=61397e4508384d13960936a556890962).

8. (Vera et al., 2008) - As we noticed in the Appendix, the random practice group involved alternating practice. Vera et

al. (2008) study design was not compliant with our PICO. The reason for exclusion is listed in Appendix 2 (https://osf.io/r59zs/?view_only=61397e4508384d13960936a556890962).

9. (Menayo et al., 2010) - As we indicated in the Appendix, we excluded Menayo et al. study since "During the acquisition phase, "A simultaneous treatment design with four repeated measures was used". (p.665) Simultaneous design provides a means of comparing two or more different treatments with an individual subject. Participants in this study practiced different tennis shots (two participants per group). Furthermore, "During the sessions, each tennis player executed a total of 192 repetitions of shots, distributed in blocked or random practice. The first 48 repetitions of each shot were executed in blocked practice, and the 144 remaining shots were executed in a random sequence" (p. 665). The reason for exclusion is listed in Appendix 2 (https://osf.io/r59zs/?view_only=61397e4508384d13960936a556890962).

10. (Rendell et al., 2010) – authors did not report mean values or SD. Instead, they presented results in the figures; however, they had such a small resolution that trying to assess the values based on the figures could be very misleading. We contacted one of the authors (Damian Farrow) via ResearchGate on 2/08/2021; however, there was no response. The reason for exclusion is listed in Appendix 2 (https://osf.io/r59zs/?view_only=61397e4508384d13960936a556890962).

11. (Sadri et al., 2013) - The Sadri et al. paper was excluded due to the very low quality (wrong numbers, doubled values, no means or SDs, no numbers of participants assigned to each group, no indication of the dependent variables – the scoring system was in Persian, etc.). The reason for exclusion is listed in Appendix 2 (https://osf.io/r59zs/?view_only=61397e4508384d13960936a556890962).

12. (Krause et al., 2014) - The study was excluded during screening since the authors manipulated feedback, providing it as blocked, serial-blocked, or serial schedule. Therefore, the study was not about the practice itself but about scheduling the feedback with no random group, and it was not compliant with our exclusion/inclusion criteria.

13. (Aiken & Genter, 2018) - The study is included in our systematic review but not in the meta-analysis since the authors reported only immediate retention testing (see inclusion and exclusion criteria in our manuscript). "*Following the acquisition phase the participants took a five minute break where they were prohibited from participating in any golf related task. They then performed two retention tests*." (p. 4).

14. (North et al., 2019) - As indicated in the Appendix, North et al. did not separately report means and SDs for retention tests. Their results represented the differences between the pre-test and the retention test. We have contacted both authors. N. Bezodis, contacted via Research Gate on 29 October 2021, and J. North via e-mail, on 2 November 2022. However, there was no response. The reason for exclusion is listed in Appendix 2 (https://osf.io/r59zs/?view_only=61397e4508384d13960936a556890962).

15. (Jeon et al., 2021) - Unlike Ammar et al. (2023) we classified and included this study in a meta-analysis of laboratory tasks. As Jeon et al. described, the purpose of their study was "*To compare the effectiveness of blocked and random practice schedules of balance training in dynamic balance abilities of older adults using* *(Nintendo*) Wii Fit balance game tasks*." (see their Abstract and subheading Wii Fit balance games for more details).

# COMMENT:

- Conclusion Overgeneralization

The authors conclude with: "The CI effect is a robust phenomenon in motor learning." Again, we do not agree with such an affirmative and generalized statement, which is not supported by the results of the current study, particularly those related to the applied setting. Additionally, the analysis conducted by the present authors only focused on the second CI effect, dealing with retention. The authors are therefore encouraged to avoid generalization and revise their conclusion to reflect only the current results. This recommendation is further supported by the potential publication bias and the fact that mainly low-quality studies were included in the present MA. Indeed, the authors stated in the "Low quality and bias problem" section of the discussion, "…we need to consider that the CI effect as described here may be biased. According to the Quality Assessment Tool for Quantitative Studies, only three articles out of 54 presented moderate or high quality." The authors are therefore encouraged to be more careful with the interpretation of their findings and to revise their generalized and affirmative statements on the beneficial effect of high CI throughout their manuscript.

# RESPONSE:

It is a good practice to cite a sentence referring to its context. The following sentence explains the rest. This is the full citation from our Conclusions:
"*The Ci effect is a robust phenomenon in motor learning. Our results evinced, however, that, similarly to Brady (2004), this claim is primarily based on laboratory studies in adults and older adults*."

We do not want to deal with cherry-picking comments.

Again, we are strongly convinced that this statement is justified. According to both analyses we applied, high CI has a medium beneficial effect (statistically significant) on the entire population. Our title and method section refer to retention. We mentioned the possible bias due to the low quality of the included studies, although we discussed it as a potential risk. At this point, we must advance hypothesis H0, which assumes the lack of evidence of bias as long as it is not proven (alternative hypothesis). This is how the scientific approach works, and it is reflected in statistical inference (H0 – there are no differences, it does not exist; HA – there is a difference, it exists). The claim that high CI does not improve retention in motor learning resembles *argumentum ad ignorantiam*, i.e., when one claims that a proposition is true because it has not been proven false or that a proposition is false because it has not been proven true. As long as it is not proven that the bias affects the pooled effects, we should assume there is no evidence that it does and that high CI improves retention in motor learning.

To be honest, we originally wanted to refer to:
"*Although the CI effect is one of the most robust and replicable effects in motor learning, the exact nature of "interference" or precisely why it is beneficial for long-term learning remains unclear*" (p. 389) (Taylor et al., 2022)

## References:


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

Ammar, A., & Schöllhorn, W. I. (2024). *Comment on Czyż et al. (2024) on Contextual Interference in Motor Learning*. https://doi.org/https://doi.org/10.51224/SRXIV.435

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

Bown, M. J., & Sutton, A. J. (2010). Quality Control in Systematic Reviews and Meta-analyses. *European Journal of Vascular and Endovascular Surgery*, *40*(5), 669–677. https://doi.org/10.1016/J.EJVS.2010.07.011

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*(4), 189–200.

Brady, F. (2004). Contextual Interference: A Meta-Analytic Study. *Perceptual and Motor Skills*, *99*(1), 116–126. https://doi.org/10.2466/pms.99.1.116-126

Corbett, M. S., Higgins, J. P. T., & Woolacott, N. F. (2014). Assessing baseline imbalance in randomised trials: implications for the Cochrane risk of bias tool. *Research Synthesis Methods*, *5*(1), 79–85. https://doi.org/10.1002/JRSM.1090

Czyż, S. H., Wójcik, A. M., Solarská, P., & Kiper, P. (2024). High contextual interference improves retention in motor learning: systematic review and meta-analysis. *Scientific Reports 2024 14:1*, *14*(1), 1–45. https://doi.org/10.1038/s41598-024-65753-3

Deeks, J., Higgins, J., & Altman, D. (2023). Chapter 10: Analysing data and undertaking meta-analyses. In J. Higgins, J. Thomas, J. Chandler, C. M, T. Li, M. Page, & V. Welch (Eds.), *Cochrane Handbook for Systematic Reviews of Interventions version 6.4 (updated August 2023)*. Cochrane. www.training.cochrane.org/handbook.

Dundar, Y., & Fleeman, N. (2017). Developing My Search Strategy. In A. Boland, M. G. Cherry, & R. Dickson (Eds.), *Doing a Systematic Review. A student's guide.* (2nd ed., pp. 61–78). Sage.

*Early Career Professionals Network: Registering Protocols of Systematic Reviews | Cochrane Community*. (n.d.). Retrieved August 1, 2024, from https://community.cochrane.org/news/early-career-professionals-network-registering-protocols-systematic-reviews

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

Harari, M. B., Parola, H. R., Hartwell, C. J., & Riegelman, A. (2020). Literature searches in systematic reviews and meta-analyses: A review, evaluation, and recommendations. *Journal of Vocational Behavior*, *118*, 103377. https://doi.org/10.1016/J.JVB.2020.103377

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
Jiménez-Díaz, J., Morera-Castro, M., & Salazar, W. (2018). The contextual Interference Effect on the Performance of Fundamental Motor Skills in Adults. *Human Movement*, *19*(1), 20–25. https://doi.org/10.5114/hm.2018.73608
Jones, L. L., & French, K. E. (2007). Effects of Contextual Interference on Acquisition and Retention of Three Volleyball Skills. *Http://Dx.Doi.Org/10.2466/Pms.105.3.883-890*, *105*(3), 883–890. https://doi.org/10.2466/PMS.105.3.883-890
Keller, G. J., Li, Y., Weiss, L. W., & Relyea, G. E. (2006). Contextual Interference Effect on Acquisition and Retention of Pistol-Shooting Skills. *Http://Dx.Doi.Org/10.2466/Pms.103.1.241-252*, *103*(1), 241–252. https://doi.org/10.2466/PMS.103.1.241-252
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. *Http://Dx.Doi.Org/10.2466/23.30.PMS.118k17w3*, *118*(2), 384–399. https://doi.org/10.2466/23.30.PMS.118K17W3
Lage, G. M., Faria, L. O., Ambrósio, N. F. A., Borges, A. M. P., & Apolinário-Souza, T. (2021). 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
Memmert, D. (2006). Long-Term Effects of Type of Practice on the Learning and Transfer of a Complex Motor Skill. *Http://Dx.Doi.Org/10.2466/Pms.103.3.912-916*, *103*(3), 912–916. https://doi.org/10.2466/PMS.103.3.912-916
Menayo, R., Sabido, R., Fuentes, J. P., Moreno, F. J., & García, J. A. (2010). Simultaneous Treatment Effects in Learning Four Tennis Shots in Contextual Interference Conditions. *Http://Dx.Doi.Org/10.2466/Pms.110.2.661-673*, *110*(2), 661–673. https://doi.org/10.2466/PMS.110.2.661-673
Moretto, N. A., Marcori, A. J., & Okazaki, V. H. A. (2018). Contextual interference effects on motor skill acquisition, retention and transfer in sport riffle schooting. *Human Movement*, *19*(2), 99–104. https://doi.org/10.5114/hm.2018.74065
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
Park, S., & Beretvas, S. N. (2019). Synthesizing effects for multiple outcomes per study using robust variance estimation versus the three-level model. *Behavior Research Methods*, *51*(1), 152–171. https://doi.org/10.3758/S13428-018-1156-Y/TABLES/6
Poklepović Peričić, T., & Tanveer, S. (2019). *Why systematic reviews matter*. https://www.elsevier.com/connect/why-systematic-reviews-matter
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
Rendell, M. A., Masters, R. S. W., Farrow, D., & Morris, T. (2010). 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
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. http://aassjournal.com/article-1-69-en.html
Sattelmayer, M., Elsig, S., Hilfiker, R., & Baer, G. (2016). A systematic review and meta-analysis of selected motor learning principles in physiotherapy and medical education. In *BMC Medical Education* (Vol. 16, Issue 1). BioMed Central Ltd. https://doi.org/10.1186/s12909-016-0538-z
Smith, P. J. K., Gregory, S. K., & Davies, M. (2003). Alternating versus blocked practice in learning a cartwheel. *Perceptual and Motor Skills*, *96*(3 (Suppl.)), 1255–1264. https://doi.org/10.2466/pms.2003.96.3c.1255
Stewart, L., Moher, D., & Shekelle, P. (2012). Why prospective registration of systematic reviews makes sense. *Systematic Reviews*, *1*(1), 1–4. https://doi.org/10.1186/2046-4053-1-7/METRICS
Straus, S., & Moher, D. (2010). Registering systematic reviews. *CMAJ*, *182*(1), 13–14. https://doi.org/10.1503/CMAJ.081849
Taylor, S., Fawver, B., Thomas, J. L., Williams, A. M., & Lohse, K. R. (2022). Practice Schedules Affect How Learners Correct Their Errors: Secondary Analysis From a Contextual Interference Study. *Journal of Motor Learning and Development*, *10*(3), 388–411. https://doi.org/10.1123/JMLD.2022-0021
Tsutsui, S., Satoh, M., & Yamamoto, K. (2013). Contextual Interference Modulated by Pitcher Skill Level. *International Journal of Sport and Health Science*.
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/TABLES/5
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. *Http://Dx.Doi.Org/10.2466/Pms.106.2.447-460*, *106*(2), 447–460. https://doi.org/10.2466/PMS.106.2.447-460
Wewege, M. A., Hansford, H. J., Shah, B., Gilanyi, Y. L., Douglas, S. R. G., Parmenter, B. J., McAuley, J. H., & Jones, M. D. (2022). Baseline imbalance and heterogeneity are present in meta-analyses of randomized clinical trials examining the effects of exercise and medicines for blood pressure management. *Hypertension Research 2022 45:10*, *45*(10), 1643–1652. https://doi.org/10.1038/s41440-022-00984-3
*What is a systematic review? - The Campbell Collaboration*. (n.d.). Retrieved August 1, 2024, from https://www.campbellcollaboration.org/what-is-a-systematic-review.html
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