Can research findings be used in clinical neuropsychology? Analysis of randomized controlled trials of working memory intervention for children

Sacha Delphine Blause* (1), François Léonard (1), Sylvie Willems (1) & Ezio Tirelli(1)

(1) *Department of psychology, University of Liège, Liège, Belgium*

*Corresponding author at: Department of psychology, University of Liège, Place des Orateurs 2 B32, 4000 Liège, Belgium. Tel: 043662017; E-mail: sacha.blauseliege.be

Abstract

Objective: Working memory difficulties are prevalent in children with cognitive disorders, impacting their academic performance and quality of life. To offer optimal care, neuropsychologists rely on evidence-based practice (EBP) principles, particularly drawing from relevant research data. Randomized controlled trials (RCTs) are the gold standard for assessing intervention effectiveness, but for these studies to be reliable and clinically applicable, they must meet specific quality standards. This article aims to assess the quality of RCTs published on working memory interventions in children.

Methods: Two investigators independently evaluated the quality of a sample of 30 RCTs on the revalidation of working memory in children with the help of the CONSORT SPI grid (completeness of reporting) and the RoB2 tool (risk of bias).

Results: In our sample, only an average of 11.9 out of 45 CONSORT items were fully reported per article. Key elements like intervention descriptions, outcome definitions, and the blinding process were inadequately reported in over 80% of articles. According to RoB2 classification, 36 RCTs were deemed high risk for bias, 10 had some concerns, and one was rated as low risk.

Conclusions: The analyzed RCTs on cognitive rehabilitation in children reveal considerable transparency issues and high bias risks, limiting their clinical reliability. Researchers need to enhance the quality and clarity of their studies to better support clinical neuropsychologists.

Key words: Research quality; Scientific communication; Neuropsychology; Cognitive rehabilitation; Short term memory; working memory

Funding: This study was not supported by any funding

Conflict of interest: Non declared

Registration and protocol: This study was not previously registered and the protocol has not been published.

Availability of data: Some data are available in the supplementary material. For further information on the material used, the date or the statistical codes employed, please contact the first author.

Reporting: The PRISMA reporting form was used to draw up this document. As this study is not a systematic review, some points were not applicable (see Reporting guideline checklist).

Introduction

Children with poor memory skills are at high risk of academic failure; this observation applies to both long-term memory (Dehn, 2010) and working memory (Gathercole & Alloway, 2008). Memory disorders are common after acquired pathology (Dehn, 2010; Goh et al., 2021; Séguin et al., 2022), but also in developmental learning disorders, including specific language disorder, Down syndrom, autism spectrum disorder and attention deficit hyperactivity disorder (Alloway & Gathercole, 2006; Dehn, 2010). Targeted treatment is essential to improve long-term outcomes for the many children affected by memory disorders. In the clinic, multiple methods are used to improve children's memory skills and associated academic performance, and there is evidence to suggest that clinical groups of children can benefit in the short term from working memory training. However, these results are variable in different populations, and sometimes there is no evidence of a generalizable long-term effect (Bharadwaj et al., 2022; Villemonteix, 2018). Consequently, neuropsychologists cannot use this kind of method without an evidence-based approach. In this context, evidence-based practice (EBP) is an essential resource to help neuropsychologists see the situation more clearly.

The international "Evidence-Based" movement in neuropsychology and other healthcare fields aims to reduce ineffective healthcare interventions and promote clinical decisions based on the judicious use of the best available evidence (Sackett et al., 1996; Straus et al., 2019). In EBP, clinicians base their practice on four “pillars”: patient characteristics, context, scientific expertise and research (Hoffmann et al., 2017; Satterfield et al., 2009; Straus et al., 2019; Watson et al., 2018). This last pillar refers to the identification, selection and reading of high-quality evidence-based articles, to ensure that their results can serve as credible benchmarks for practice and guide neuropsychologists in their decision-making (Agoritsas et al., 2015). Among studies of interventions, the results of primary studies are often considered more reliable when obtained by means randomized controlled trials (RCTs), a study design seen as the "gold standard" for assessing the effects of interventions. RCTs allow researchers to draw causal inferences because of the use of an experimental group and a control group and the random allocation of participants to each of these groups (APA Presidential Task Force on Evidence-Based Practice, 2006; Hulley, 2013; Sessler & Imrey, 2015).

RCTs are designed to isolate the effects of an independent variable on a dependent variable by reducing the impact of environmental variables. In this way, they enable neuropsychologists to obtain information about the effectiveness of a treatment for a specific type of patient, by comparing it with another intervention or with no intervention. This approach facilitates the generalization of results and makes it possible to calculate the intervention’s effect size, essential information when choosing an intervention. Despite the methodological advantages of other types of designs such as Single-Case Experimental Designs (SCED), the RCT remains a preferred choice for evaluating interventions because of these characteristics. If the information provided by RCTs is to be useful for neuropsychologists, they must be carried out in a rigorous, reliable and solid manner. Meta-research in various areas of healthcare has revealed inappropriate research practices that adversely affect the quality, reliability and reproducibility of results (Anderson et al., 2013; Calin-Jageman & Cumming, 2019; Cuijpers et al., 2015; Ioannidis, 2005; Munafò et al., 2017; Wang et al., 2019; Westover et al., 2011).

Psychological research is no exception. Indeed, common research biases in the psychological literature include the excessive importance attributed to the *p*-value; lack of interpretation of the effect size (Faulkner et al., 2008; Gigerenzer, 2018); researchers' allegiance to a scientific trend (Falkenström et al.,

2013; Munder et al., 2012); the implementation of underpowered studies (Button et al., 2013); a lack of transparency in the writing of articles; and the use of questionable practices such as data dredging (carrying out numerous statistical tests on a set of data until statistically significant results are found), HARKing (Hypothesizing After the Results are Known) and *p*-hacking (use of various techniques to increase the chances of finding statistically significant results, even if the results are not really significant) (Scheel et al., 2021). These biases, defined as "the combination of various design, data, analysis, and presentation factors that tend to produce research findings when they should not be produced" (Ioannidis, 2005, p. 0697), encourage illusory conclusions (false positives, overestimation of effects, etc.). A significant amount of psychological research suffers from poor reproducibility, raising clear ethical concerns (Hardwicke et al., 2022; Ioannidis, 2005; Ioannidis et al., 2015). Broadly speaking, many findings from RCTs in clinical psychology may be of little practical use to clinicians (Faulkner et al., 2008). To date, no research has specifically assessed the quality of trials in clinical neuropsychology, particularly regarding the effectiveness of cognitive interventions. However, if methodological shortcomings exist, they could severely impede the adoption of EBP in clinical neuropsychology. A study among clinical neuropsychologists revealed that clinicians rarely use literature to guide their practice or answer clinical questions. While most neuropsychologists are aware of biases in trials, many feel they lack the expertise to assess their significance (Blause et al., 2023). Given this background, it is all the more essential for neuropsychologists to have at their disposal high-quality, transparent scientific articles on which clinicians can base their practice.

It is crucial to overcome the lack of transparency and reduce biases, which are common and significant problems in the scientific literature (Faulkner et al., 2008; Gigerenzer, 2018; Munafò et al., 2017; Scheel et al., 2021; Thabane et al., 2013). Different tools such as CONSORT and RoB2 now exist to support researchers in the execution and reporting of RCTs. CONSORT is a statement proposed to enhance the reporting and transparency of RCTs (Moher et al., 2001). RCTs of psychological interventions must be reported transparently and in detail to minimize reporting bias and maximize the credibility and usefulness of research data (Cybulski et al., 2016; Ioannidis et al., 2014). Scientific journals are increasingly recommending the use of this tool (Shamseer et al., 2016). RoB2 is a critical appraisal tool designed to help evaluate of the risk of bias in RCTs but it can also be used in the methodological development of trials. Authors of meta-analyses regularly apply RoB2 to assess the quality of the studies included in their reviews.

Although these tools are valuable, they appear to be underutilized and often misapplied, which may explain the methodological flaws observed. Consequently, it is essential to evaluate the reliability of RCTs published in the field of cognitive remediation in neuropsychology and to identify their methodological limitations. This will contribute to improving practices and making scientific literature more accessible and valuable for clinical neuropsychologists. Our meta-research therefore focused on assessing the methodological quality of a sample of RCTs in cognitive remediation and identifying improvements to enhance the reliability, clarity, and reproducibility of published data. Exploratory analyses were also carried out to investigate potential links between quality and certain characteristics of these trials.

In this study, we hypothesize that, as in other areas of psychology, the methodological quality of RCTs in neuropsychology is not satisfactory and that an effort should be made to enhance reporting and rigor in the implementation of the methodological criteria necessary for producing data that are useful and relevant to clinical neuropsychologists. To this end, we created a sample of RCTs evaluating the effectiveness of cognitive interventions on children's working memory and assessed their quality.

Methods

Screening

A search of the literature was performed in December 2020 and in march 2025 to compile a sample of RCTs evaluating the effectiveness of interventions aimed at improving memory skills in children and adolescents published in English or in French and indexed in the online databases PsycInfo and PubMed. The electronic search keywords were developed by the coauthors (Table 1).

Two investigators (S.B. and S.W.) independently reviewed each title and abstract to exclude irrelevant articles and select only the studies that met the eligibility criteria (Table 2). A second selection phase was carried out on the basis of the full texts. After each phase, there was a joint discussion to resolve any potential conflicts and create an initial sample of articles.

Quality assessment

The quality of the sample was assessed in two stages.

First, CONSORT was used to assess the completeness of reporting of the sample, as it provides checklist of essential items that should be included in reports of RCTs (Montgomery et al., 2018). This tool is designed to help authors write their articles comprehensively and transparently. There are several versions of this reporting tool. CONSORT-SPI was designed for reporting on RCTs for social and psychological interventions. It is composed of seven main sections that structure the article and contains a total of 45 items that must be reported. Each individual CONSORT item receives a "yes" or "no" response depending on whether the item was or was not fulfilled referring to the description of each item on the CONSORT website (www.consort-statement.org, see the Supplementary material for more information).

Second, RoB 2 was used to assess some of the most important experimental biases (Sterne et al., 2019). This tool is composed of five areas in which bias may influence a study's results: bias arising from the randomization process, bias due to deviations from intended interventions, bias due to missing outcome data, bias in measurement of the outcome and, bias in selection of the reported result. Each area assessed by RoB 2 comprises a series of questions that can be answered "yes" or "probably yes", "no" or "probably no", "no information" and, in some cases, "not applicable". The tool uses algorithms to link all the answers to the various questions to an estimate of the risk of bias for each of the five methodological domains assessed. RoB 2 can also be used to assess the risk of bias for an article as a whole. The possible risk-of-bias judgements are Low risk of bias, Some concerns and High risk of bias.

These analyses were independently conducted by two researchers (S.B. and F.L.) and then reviewed in the presence of a third researcher (E.T.) to resolve any disagreements for more than 50% of the sample. Given the high agreement rate, the remainder of the sample was analyzed by a single researcher (S.B.), with the assistance of the other two researchers (F.L. and E.T.) in case of doubt. The data collected was encoded in an Excel table.

Data analysis

We used descriptive statistics to assess the general characteristics of articles and highlight the methodological rigor of RCTs by assessing their compliance with the CONSORT guidelines and identifying factors influencing the quality of RCTs with RoB 2. We will present the data using frequencies and percentages. To gain further insights, four exploratory analyses were conducted. We use four independent sample Student's t-tests with Jamovi (version 2.3.28) to compare the completeness of reporting across

different groups: articles that adhere to the CONSORT guidelines versus those that do not; articles published in high-ranking journals versus those in lower-ranking ones; articles with a large number of coauthors versus those with fewer coauthors and articles published before CONSORT SPI had been published (2018) and those published after. If the normality condition was not met (Shapiro-Wilk test), the non-parametric equivalent of this test (Mann-Whitney U) was used. The alpha threshold used is, by convention, 0.05.

Results

Search results

A total of 781 references were identified by searching the selected databases (Figure 1.). After selection on the basis of title and abstract, 101 articles were read in their entirety to obtain a first sample of 55 articles. It should be noted that we initially wanted to create a sample of articles investigating the effectiveness of cognitive interventions on all types of memory. However, given that the vast majority of articles dealt with the cognitive management of working memory, we elected to focus only on these articles in order to obtain a more homogeneous sample. After removing the few articles dealing with the rehabilitation of long-term memory, we obtained a final sample of 47 articles (see Table 3 for general characteristics and the Supplementary material for the references to the articles). As working memory is a widely studied area, we did not extend our sample to other executive functions (e.g. inhibition or shifting). Adding the term "executive function" to our search strategy would have resulted in too many irrelevant articles.

Completeness of reporting of the RCTs

Of the 45 items making up CONSORT, on average, 11.9 are reported completely by our sample of articles. The maximum number of items correctly reported by an article was 20, while the minimum was 3 (see Table 4 for all results and Supplementary material for results for each article).

The following CONSORT elements were fully reported by 50% or more of our sample: identification as a RCT in the title (68%), the scientific background and explanation of rationales of the study (100%), the information needed to ensure that the intervention was carried out as planned (57%), the description of the similarities in the different conditions of the study (57%), the participants' demographic data in the results (66%), the presentation of the trial's limitations (91%) and the generalizability of the results in the discussion section (55%); and the presence of potential conflicts of interest (85%).

The specific objectives and precise hypotheses of the study (9%), the description of the intervention in sufficient details to allow replication (21%), the definition of the outcomes (13%), the blinding process (13%), the precise results for each outcome (30%) and the clinical interpretation of the results (21%) are examples of items that were under reported by the vast majority of the articles in our sample (see Table 4 for all the results).

The Mann-Whitney *U* was used for the exploratory analyses because the normality condition was violated (W= 0.976; p=.43). Regarding the comparison of the completeness of reporting between articles that claim to adhere to the CONSORT guidelines and those that do not, the test shows a significant difference between the means for the two groups (U= 104; p< .0001; r =0.623; large effect size). Articles claiming to use CONSORT are significantly more exhaustive (M= 14.1; SD = 4.54) than articles do not cite CONSORT (M = 9; SD = 3.08) (Figure 2). Articles that follow the CONSORT guidelines more often specify the study design in the title, provide clearer explanations of how the sample size was determined, and describe the method used to generate the random allocation sequence. They also report losses and exclusions more frequently, along with demographic variables. In addition, their results tend to be more thoroughly reported. Moreover, stakeholder involvement are mentioned more frequently in these articles (Table 4). However, the year of publication of the article has no impact on its transparency (U= 237; p= 813; r =0.044; small effect size). Articles published after the publication of the SPI CONSORT (2018 and after) did not score better (M= 11.8; SD = 4.81) than those published before (2017-2015) (M= 12; SD = 3.9) (Figure 2).

The ranking of the journal in which the article was published has impact on the transparency of the reporting. There was a significant difference (U= 122; p= .018; r =0.45; moderate effect size) between the

mean Consort scores of articles published in Q1-ranked journals (*M*= 12.8; *SD*= 4.25) and articles published in Q2-3-4-ranked journals (mean= 9.46; standard deviation= 4.25) (Figure 2). Although producing and drafting an RCT is complex and resource-intensive, the number of people involved in the process does not significantly influence the quality of the reporting. The CONSORT score did not differ significantly (*U*= 249; *p*= .578; *r* =0.0964; small effect size) between articles with 7 or more coauthors (*M*= 12.3; *SD*= 4.67) and those with fewer than 7 coauthors (*M*= 11.5; *SD*= 4.36) (Figure 2).

Risk of bias

Across our sample of 47 RCTs, according to the classification recommended by RoB 2, 36 RCTs were judged to be at high risk of bias, 10 were judged to raise some concerns and 1 was considered to be at low risk of bias.

If we consider the domains separately, we note that few or no articles are considered to be at high risk of bias for the randomization process (4/47), the management of missing data (0/47) and the reporting of results (5/47). According to RoB 2, more articles were classified as being at high risk of bias for deviations in the intervention (29/47) and the measurement of outcomes (17/47).

In Table 5, we summarize the RoB 2 results for the 47 RCTs (see Supplementary material for results for each article). A more in-depth examination of the results and the reasons for these biases is provided in the discussion.

Discussion

Memory difficulties in children are just one area in which clinicians need to turn to the literature to learn which intervention methods work for which clinical populations. When they do so, the evidence-based approach encourages them to consult review studies and, when these do not exist, RCTs, which many authors consider to provide a high level of evidence. However, RCTs credibility and reliability depend on the use of rigorous methodology, followed by equally rigorous reporting. Incomplete, non-transparent reporting prevents readers from assessing a study's strengths and weaknesses and the reliability of the data (Faulkner et al., 2008; Gigerenzer, 2018; Munafò et al., 2017; Scheel et al., 2021; Thabane et al., 2013). To this end, tools such as CONSORT have been created to help authors better report on RCTs. Other tools, such as RoB 2, have been designed for readers to help them assess the risk of bias in RCTs. Do researchers use these tools to help them design and present their randomised controlled trials? This study used CONSORT and RoB 2 to provide a global overview of the completeness of reporting and the risk of bias for RCTs that evaluate the effectiveness of cognitive intervention on children's working memory. Generally speaking, the results obtained point to a lack of transparency and a significant risk of bias in the majority of the articles evaluated.

With regard to the completeness of reporting, the scores obtained by our sample on CONSORT are unsatisfactory and indicate a lack of transparency, which is consistent with our hypotheses and with the results of studies carried out in other fields (Kilicoglu et al., 2023; Scheel et al., 2021). Indeed, on average, only 11.9 out of the 45 items in CONSORT are completely reported by our sample of articles. However, some of the underreported items provide information that is essential for understanding and implementing the research in neuropsychologists' the day-to-day clinical work.

For example, CONSORT recommends that the intervention tested in the trial should be described in great detail, with sufficient information to enable readers to replicate what was. Compliance with this item allows neuropsychologists to apply the intervention in their practice and offer it to suitable patients. Moreover, providing sufficient detail for replication increases the likelihood that other researchers will reproduce the study to verify or challenge the results (Royal Netherlands Academy of Arts and Sciences, 2018). Very few of the articles in our sample (21%) described their intervention in detail. Information on the number of sessions, the location of the intervention and the precise materials used was generally missing. Authors should more systematically include information in supplementary materials to solve this problem without making articles too long.

Information on blinding is also essential for the neuropsychologist so that they can assess the reliability of a study's results and critically evaluate the article. However, in our sample, only six articles provided all the necessary information on this subject. To be comprehensive, CONSORT recommends providing information on blinding for all stakeholders in the study, namely, the participants, the participants' supervisors (parent, physician, teacher, etc.), the experimenter administering the intervention, the experimenter evaluating the participants in pre- and post-tests, and the researcher analyzing the data. Most articles analyzed in this study report information on the participants or on the experimenter testing the children before and after the intervention. However, generally, no information is provided regarding the researcher analyzing the data. To facilitate the reporting of this information, a table summarizing the different stakeholders in the study with their level of awareness could be created.

Surprisingly, the number of articles in our sample that report their results exhaustively is low (31%). Indeed, although all analyzed articles presented their results, this was not done comprehensively. The CONSORT checklist recommends providing the *p*-value, the effect size, and the confidence interval for the effect size for each group. While most articles in our sample report the *p*-value relative to their outcomes,

the effect size is reported less frequently and is often only cited for outcomes considered statistically significant. The confidence interval for the effect size is also very rarely reported (31%). These data are, however, essential for the understanding and use of the literature by clinical neuropsychologists. Although the *p*-value is systematically reported, it only provides information on the probability of obtaining such a result if the null hypothesis is true. From a clinical standpoint, this is not enough information (Faulkner et al., 2008; Gigerenzer, 2018). Informing neuropsychologists of the effect size of the intervention they wish to use is more relevant (Faulkner et al., 2008). Furthermore, neuropsychologists are aware that all patients they encounter are different and that the effectiveness of an intervention may therefore vary. It is therefore essential to assist them in their decision-making by informing them about the variability of the effect (Gigerenzer, 2018). The confidence interval provides this information. In articles that correctly include this item, the results were presented in a comprehensive table, which seems to be the preferred way of presenting the results of each outcome exhaustively.

Although it is important for neuropsychologists to know the size and variability of an intervention's effect, it is even more important to understand what this implies concretely for their patients (Faulkner et al., 2008). The clinical interpretation of results is an essential phase in research and crucial for practitioners. In our sample of articles, only a small fraction takes this information into account (21%) even though it is a fundamental bridge between research and practice.

The focus on these different items emphasizes that the CONSORT checklist is very comprehensive but also complex. Several items require a significant amount of information, and it might even be possible to divide certain items into sub-items, as in the blinding section, for example. To make this study useful to researchers, we have created a table highlighting the information that is often missing but that we consider particularly relevant for readers and easy for researchers to include (see Supplementary material).Writing an article that provides exhaustive reporting requires a great deal of work by researchers and the use of reporting assistance tools should not be left until the final stages of article writing to verify that all necessary information is included. Rather, such tools should be introduced from the early stages of the writing process and could even be consulted at the beginning of study design with the help of the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT), for example (Chan et al., 2013). It is advisable to publish these protocols in peer reviews (Hardwicke & Wagenmakers, 2023). It would also be necessary for journals to mandate the use of CONSORT, as some already do (Shamseer et al., 2016). Indeed, as we have seen, explicitly citing CONSORT, even if it is not perfectly applied, improves the quality of reporting. Furthermore, the articles classified in Q1 seem to be more complete than the others - this could be the reason.

The risk of bias of the articles in our sample is high, and only one of these articles is considered to be at low risk of bias. This is in line with our hypotheses and previous studies (Leichsenring et al., 2017; Scheel et al., 2021). On closer examination, the analyses show quite clearly that the methodological biases of these studies are essentially in two of the five areas assessed by RoB 2: deviations from the intended intervention and measurement of the outcomes.

Concerning “deviations from the intended intervention”, it is a lack of information (about blinding, the potential deviation from intended intervention, the impact of these deviations on the results and the reasons for attrition) that puts most articles into the “high risk” class. The lack of transparency identified with CONSORT seems to be reflected here in the estimate of the risk of bias. Measurement of outcomes is at high risk of bias for many of the articles in our sample because the outcome assessor is not blinded. To avoid the presence of confirmation bias, it is recommended that the evaluator responsible for testing the outcomes not be informed of the condition to which the participants were allocated.

The RoB 2 tool is designed to assist readers in critically appraising articles. However, our analyses suggest that, if authors familiarized themselves with this tool before conducting their study, it could be beneficial for avoiding certain methodological errors and authorial shortcomings.

These results suggest that RCTs in neuropsychology often fail to meet the quality standards needed to provide clinicians with reliable data. The lack of transparency raises concerns about questionable methodologies. Can RCTs truly be considered the "gold standard" for assessing interventions in neuropsychology? Alternatively, could designs like SCEDs offer more reliable insights (Davison & Lazarus, 1995; Hawkins & Axelrod, 2008; Tate & Perdices, 2015)? While we may not have a definitive answer, it is clear that neuropsychologists need to critically assess RCTs before applying them in practice—a process that requires time, resources, and skills, which are often lacking (Blause et al., 2023).

Furthermore, the lack of transparency complicates replication efforts, reducing the practical utility of these trials for clinicians. To support clinicians in critically evaluating research, we recommend the use of appraisal tools such as RoB2 or the JBI's critical appraisal checklists (Barker et al., 2023). It is also essential to read full articles—or at the very least, thoroughly examine the methods section—in addition to the much more commonly consulted abstract, discussion, and conclusion (Blause et al., 2023).

We believe that critical reading skills should be emphasized in evidence-based practice education, both within university programs and in ongoing professional development for neuropsychologists. Moreover, understanding how research bias can affect clinical outcomes and ethics in scientific research should be a central focus in the training of future researchers.

This study has several limitations. First, while it provides insights into the quality of RCTs in neuropsychology, the sample includes only RCTs published between 2015 and 2025 in English or French, sourced from PubMed or PsycInfo, and focused on cognitive interventions for children's and adolescents' working memory. This limits the generalizability of the findings. Future research could expand the analysis to include other types of treatments, populations, or more recent studies from different databases. In addition, the CONSORT grid was used to assess report completeness, though it was not designed for this purpose, leading to potential biases such as assigning a score of 0 when elements were incomplete or interpreting specific criteria from the CONSORT website rather than the general statement. Despite a high agreement rate (90%) between authors, subjectivity may still affect scoring. Furthermore, we generated a total score for each item to reflect completeness, although not all items are equally relevant to neuropsychologists. For example, detailed descriptions of interventions are more useful than knowing who the study's stakeholders were. It might be beneficial to have a panel of experts weigh the CONSORT items to create a more accurate score. Lastly, it would have been interesting to dig deeper into the analyses linked to RoB 2, as we did for CONSORT. Unfortunately, the small number of "low risk" and "some concerns" articles in this sample meant that we were unable to go any further in that regard.

In conclusion, the RCTs published in the field of cognitive rehabilitation in children analyzed in this study lacked transparency and exhibited high risk of bias. These results could be a barrier to the utilization of scientific literature by clinical neuropsychologists. There is still work to be done by researchers to improve the quality of their research and their writing. This article offers avenues for improvement. Even small changes can contribute to improving the scientific literature overall, making it more accessible and narrowing the gap between research and clinical practice in neuropsychology.

References

Agoritsas, T., Vandvik, P., Neumann, I., Rochwerg, B., Jaeschke, R., Hayward, R., Guyatt, G., & McKibbon, K. (2015). Finding current best evidence. In *Users' guides to the medical literature : A manual for evidence-based clinical practice* (3rd ed). McGraw Hill.

Alloway, T. P., & Gathercole, S. E. (Éds.). (2006). *Working memory and neurodevelopmental disorders* (1st pub.). Psychology Press.

Anderson, B. L., Williams, S., & Schulkin, J. (2013). Statistical Literacy of Obstetrics-Gynecology Residents. *Journal of Graduate Medical Education*, *5*(2), 272-275. https://doi.org/10.4300/JGME-D-12-00161.1

APA Presidential Task Force on Evidence-Based Practice. (2006). Evidence-based practice in psychology. *The American Psychologist*, *61*(4), 271-285. https://doi.org/10.1037/0003-066X.61.4.271

Barker, T. H., Stone, J. C., Sears, K., Klugar, M., Tufanaru, C., Leonardi-Bee, J., Aromataris, E., & Munn, Z. (2023). The revised JBI critical appraisal tool for the assessment of risk of bias for randomized controlled trials. *JBI Evidence Synthesis*, *21*(3), 494-506. https://doi.org/10.11124/JBIES-22-00430

Bharadwaj, S. V., Yeatts, P., & Headley, J. (2022). Efficacy of cogmed working memory training program in improving working memory in school-age children with and without neurological insults or disorders : A meta-analysis. *Applied Neuropsychology: Child*, *11*(4), 891-903. https://doi.org/10.1080/21622965.2021.1920943

Blause, S., Tirelli, E., Wauquiez, G., Raffard, S., Didone, V., & Willems, S. (2023). What Information Do Neuropsychologists Use to Guide their Clinical Decisions? A Survey on Knowledge and Application of Evidence-Based Practice in a French-Speaking Population. *Archives of Clinical Neuropsychology*, acad057. https://doi.org/10.1093/arclin/acad057

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

Calin-Jageman, R. J., & Cumming, G. (2019). The New Statistics for Better Science : Ask How Much, How Uncertain, and What Else Is Known. *The American Statistician*, *73*(sup1), 271-280. https://doi.org/10.1080/00031305.2018.1518266

Chan, A.-W., Tetzlaff, J. M., Altman, D. G., Laupacis, A., Gøtzsche, P. C., Krleža-Jerić, K., Hróbjartsson, A., Mann, H., Dickersin, K., Berlin, J. A., Doré, C. J., Parulekar, W. R., Summerskill, W. S. M., Groves, T., Schulz, K. F., Sox, H. C., Rockhold, F. W., Rennie, D., & Moher, D. (2013). SPIRIT 2013 Statement : Defining Standard Protocol Items for Clinical Trials. *Annals of Internal Medicine*, *158*(3), 200. https://doi.org/10.7326/0003-4819-158-3-201302050-00583

Cuijpers, P., Karyotaki, E., Andersson, G., Li, J., Mergl, R., & Hegerl, U. (2015). The Effects of Blinding on the Outcomes of Psychotherapy and Pharmacotherapy for Adult Depression : A Meta-Analysis. *European Psychiatry*, *30*(6), 685-693. https://doi.org/10.1016/j.eurpsy.2015.06.005

Cybulski, L., Mayo-Wilson, E., & Grant, S. (2016). Improving transparency and reproducibility through registration : The status of intervention trials published in clinical psychology journals. *Journal of Consulting and Clinical Psychology*, *84*(9), 753-767. https://doi.org/10.1037/ccp0000115

Davison, G. C., & Lazarus, A. A. (1995). The dialectics of science and practice. In *Scientific standards of psychological practice : Issues and recommendations* (p. 95-120). Context Press/New Harbinger Publications.

Dehn, M. J. (2010). *Long-term memory problems in children and adolescents : Assessment, intervention, and effective instruction*. Wiley.

Falkenström, F., Markowitz, J. C., Jonker, H., Philips, B., & Holmqvist, R. (2013). Can Psychotherapists Function as Their Own Controls? : Meta-Analysis of the Crossed Therapist Design in Comparative Psychotherapy Trials. *The Journal of Clinical Psychiatry*, *74*(05), 482-491. https://doi.org/10.4088/JCP.12r07848

Faulkner, C., Fidler, F., & Cumming, G. (2008). The value of RCT evidence depends on the quality of statistical analysis. *Behaviour Research and Therapy*, *46*(2), 270-281. https://doi.org/10.1016/j.brat.2007.12.001

Gathercole, S. E., & Alloway, T. P. (2008). *Working memory and learning : A practical guide for teachers* (1. publ). SAGE Publ.

Gigerenzer, G. (2018). Statistical Rituals : The Replication Delusion and How We Got There. *Advances in Methods and Practices in Psychological Science*, *1*(2), 198-218. https://doi.org/10.1177/2515245918771329

Goh, M. S. L., Looi, D. S. H., Goh, J. L., Sultana, R., Goh, S. S. M., Lee, J. H., & Chong, S.-L. (2021). The Impact of Traumatic Brain Injury on Neurocognitive Outcomes in Children : A Systematic Review and Meta-Analysis. *Journal of Neurology, Neurosurgery & Psychiatry*, *92*(8), 847-853. https://doi.org/10.1136/jnnp-2020-325066

Hardwicke, T. E., Thibault, R. T., Kosie, J. E., Wallach, J. D., Kidwell, M. C., & Ioannidis, J. P. A. (2022). Estimating the Prevalence of Transparency and Reproducibility-Related Research Practices in Psychology (2014–2017). *Perspectives on Psychological Science*, *17*(1), 239-251. https://doi.org/10.1177/1745691620979806

Hardwicke, T. E., & Wagenmakers, E.-J. (2023). Reducing bias, increasing transparency and calibrating confidence with preregistration. *Nature Human Behaviour*, *7*(1), 15-26. https://doi.org/10.1038/s41562-022-01497-2

Hawkins, R. O., & Axelrod, M. I. (2008). Increasing the On-Task Homework Behavior of Youth With Behavior Disorders Using Functional Behavioral Assessment. *Behavior Modification*, *32*(6), 840-859. https://doi.org/10.1177/0145445508318846

Hoffmann, T., Bennett, S., & Del Mar, C. (2017). *Evidence-based practice across the health professions* (3rd ed). Elsevier.

Hulley, S. B. (Éd.). (2013). *Designing clinical research* (4th ed). Wolters Kluwer/Lippincott Williams & Wilkins.

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

Ioannidis, J. P. A., Fanelli, D., Dunne, D. D., & Goodman, S. N. (2015). Meta-research : Evaluation and Improvement of Research Methods and Practices. *PLOS Biology*, *13*(10), e1002264. https://doi.org/10.1371/journal.pbio.1002264

Ioannidis, J. P. A., Munafò, M. R., Fusar-Poli, P., Nosek, B. A., & David, S. P. (2014). Publication and other reporting biases in cognitive sciences : Detection, prevalence, and prevention. *Trends in Cognitive Sciences*, *18*(5), 235-241. https://doi.org/10.1016/j.tics.2014.02.010

Kilicoglu, H., Jiang, L., Hoang, L., Mayo-Wilson, E., Vinkers, C. H., & Otte, W. M. (2023). Methodology reporting improved over time in 176,469 randomized controlled trials. *Journal of Clinical Epidemiology*, *162*, 19-28. https://doi.org/10.1016/j.jclinepi.2023.08.004

Leichsenring, F., Abbass, A., Hilsenroth, M. J., Leweke, F., Luyten, P., Keefe, J. R., Midgley, N., Rabung, S., Salzer, S., & Steinert, C. (2017). Biases in research : Risk factors for non-replicability in psychotherapy and pharmacotherapy research. *Psychological Medicine*, *47*(6), 1000-1011. https://doi.org/10.1017/S003329171600324X

Moher, D., Schulz, K. F., & Altman, D. G. (2001). The CONSORT statement : Revised recommendations for improving the quality of reports of parallel-group randomised trials. *The Lancet*, *357*(9263), 1191-1194. https://doi.org/10.1016/S0140-6736(00)04337-3

Montgomery, P., Grant, S., Mayo-Wilson, E., Macdonald, G., Michie, S., Hopewell, S., Moher, D., & CONSORT-SPI Group. (2018). Reporting randomised trials of social and psychological interventions : The CONSORT-SPI 2018 Extension. *Trials*, *19*(1), 407. https://doi.org/10.1186/s13063-018-2733-1

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie Du Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., & Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, *1*(1), 0021. https://doi.org/10.1038/s41562-016-0021

Munder, T., Flückiger, C., Gerger, H., Wampold, B. E., & Barth, J. (2012). Is the allegiance effect an epiphenomenon of true efficacy differences between treatments? A meta-analysis. *Journal of Counseling Psychology*, *59*(4), 631-637. https://doi.org/10.1037/a0029571

Royal Netherlands Academy of Arts and sciences. (2018). *Replication studies : Improing reproductibility in the empirical sciences*. Advisory Report.

Sackett, D. L., Rosenberg, W. M. C., Gray, J. A. M., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine : What it is and what it isn't. *BMJ*, *312*(7023), 71-72. https://doi.org/10.1136/bmj.312.7023.71

Satterfield, J. M., Spring, B., Brownson, R. C., Mullen, E. J., Newhouse, R. P., Walker, B. B., & Whitlock, E. P. (2009). Toward a transdisciplinary model of evidence-based practice. *The Milbank Quarterly*, *87*(2), 368-390. https://doi.org/10.1111/j.1468-0009.2009.00561.x

Scheel, A. M., Schijen, M. R. M. J., & Lakens, D. (2021). An Excess of Positive Results : Comparing the Standard Psychology Literature With Registered Reports. *Advances in Methods and Practices in Psychological Science*, *4*(2), 251524592110074. https://doi.org/10.1177/25152459211007467

Séguin, M., Gagner, C., Tuerk, C., Lacombe Barrios, J., MacKay, P., & Beauchamp, M. H. (2022). What About the Little Ones? Systematic Review of Cognitive and Behavioral Outcomes Following Early TBI. *Neuropsychology Review*, *32*(4), 906-936. https://doi.org/10.1007/s11065-021-09517-0

Sessler, D. I., & Imrey, P. B. (2015). Clinical Research Methodology 3 : Randomized Controlled Trials. *Anesthesia & Analgesia*, *121*(4), 1052-1064. https://doi.org/10.1213/ANE.0000000000000862

Shamseer, L., Hopewell, S., Altman, D. G., Moher, D., & Schulz, K. F. (2016). Update on the endorsement of CONSORT by high impact factor journals : A survey of journal "Instructions to Authors" in 2014. *Trials*, *17*(1), 301. https://doi.org/10.1186/s13063-016-1408-z

Sterne, J. A. C., Savović, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., Cates, C. J., Cheng, H.-Y., Corbett, M. S., Eldridge, S. M., Emberson, J. R., Hernán, M. A., Hopewell, S., Hróbjartsson, A., Junqueira, D. R., Jüni, P., Kirkham, J. J., Lasserson, T., Li, T., … Higgins, J. P.

T. (2019). RoB 2 : A revised tool for assessing risk of bias in randomised trials. *BMJ*, l4898. https://doi.org/10.1136/bmj.l4898

Straus, S. E., Glasziou, P., Richardson, W. S., & Haynes, R. B. (2019). *Evidence-based medicine : How to practice and teach EBM* (5th éd.). Elsevier Health Sciences.

Tate, R. L., & Perdices, M. (2015). N-of-1 Trials in the Behavioral Sciences. In J. Nikles & G. Mitchell (Éds.), *The Essential Guide to N-of-1 Trials in Health* (p. 19-41). Springer Netherlands. https://doi.org/10.1007/978-94-017-7200-6_3

Thabane, L., Samaan, Mbuagbaw, L., Kosa, S., Borg Debono, Dillenburg, Zhang, Fruci, V., Dennis, & Bawor. (2013). A systematic scoping review of adherence to reporting guidelines in health care literature. *Journal of Multidisciplinary Healthcare*, 169. https://doi.org/10.2147/JMDH.S43952

Villemonteix, T. (2018). L'entraînement de la mémoire de travail est-il bénéfique pour les enfants présentant un trouble déficit de l'attention/hyperactivité ? *Neuropsychiatrie de l'Enfance et de l'Adolescence*, *66*(1), 3-12. https://doi.org/10.1016/j.neurenf.2017.07.003

Wang, C., Xu, T., Geng, F., Hu, Y., Wang, Y., Liu, H., & Chen, F. (2019). Training on Abacus-Based Mental Calculation Enhances Visuospatial Working Memory in Children. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *39*(33), 6439-6448. https://doi.org/10.1523/JNEUROSCI.3195-18.2019

Watson, D. P., Adams, E. L., Shue, S., Coates, H., McGuire, A., Chesher, J., Jackson, J., & Omenka, O. I. (2018). Defining the external implementation context : An integrative systematic literature review. *BMC Health Services Research*, *18*(1), 209. https://doi.org/10.1186/s12913-018-3046-5

Westover, M. B., Westover, K. D., & Bianchi, M. T. (2011). Significance testing as perverse probabilistic reasoning. *BMC Medicine*, *9*(1), 20. https://doi.org/10.1186/1741-7015-9-20

Tables legends

Table 1: *Search keywords*

Table 2: *Eligibility criteria*

Table 3: *General characteristics of the sample*

Table 4: *Completeness of reporting of the sample (CONSORT SPI)*

Table 5: *Risk of bias of the sample (RoB2)*

Figure legends

Figure 1 *Flow chart:* n=number of articles

Figure 2 *Completeness of reporting – exploratory analyses:* Averages and standard deviations of CONSORT scores

Tables

Table 1
Search keywords

PsycInfo	
1	memory/
2	early memories/
3	episodic memory/
4	explicit memory/
5	implicit memory/
6	long term memory/
7	memory consolidation/
8	memory trace/
9	prospective memory/
10	reminiscence/
11	retrospective memory/
12	short term memory/
13	exp spatial memory/
14	exp verbal memory/
15	exp visual memory/
16	memory training/
17	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16
18	limit 17 to ("0300 clinical trial" and (childhood <birth to 12 years> or adolescence <13 to 17 years>) and yr="2015 -Current")
PubMed	
1	Memory [Mesh]
2	limit 1 to (“randomized controlled trial” and ‘child: birth to 18 years” and yr=”2015-Current”)

Table 2

Eligibility criteria

Inclusion criteria
Study concerned children or adolescents under 18 years of age with or without a pathology
Trial investigated the effect of cognitive rehabilitation
Aim of the intervention was to improve the subjects' memory capacity
RCTs
Published between 2015 and 2025
In English or French
Exclusion criteria
Quasi-experimental studies or non-random allocation
Trials testing the effectiveness of non-cognitive interventions (sport, diet, emotional management, etc.)
Letter, thesis, comment, abstract, chapter, erratum, dissertation or editorial

Table 3

General characteristics

Characteristics	*Category*	*N*	*%*	*Mean*	*Median*	*Max*	*Min*
Articles							
Mention of the use of CONSORT		24	51				
Open access		27	57				
Registration of protocol		15	32				
Authors							
Number of authors	1-2 authors	6	13				
	3-4 authors	11	23				
	5-6 authors	8	17				
	7 authors or more	22	47				
H-index of first author				15.98	11	90	1
Journal							
Cite score (Scopus 2023)				836	7.55	41.2	1.8
Journal ranking	Q1	35	74				
	Q2	10	21				
	Q3	1	2				
	Q4	1	2				

N: numbers of articles concerned; %: percentage of articles concerned; Mean: mean results; Median: median results; Max: maximum value obtained; Min: minimum value obtained.

Table 4

Completeness of Reporting – CONSORT-SPI

Section	Item		Item present		CONSORT (n = 24)		Not CONSORT (n = 23)	
Title and abstract			N	%	N	%	N	%
Title	1	Identification as a randomized trial in the title	32	68	22	92	10	43
Abstract	2	Structured summary of trial design, methods, results, and conclusions (see CONSORT SPI for Abstracts)	0	0	0	0	0	0
Introduction								
Background and objectives	3	Scientific background and explanation of rationale	47	100	24	100	23	100
	4	Specific objectives or hypotheses	4	9	2	8	2	9
	5	If pre-specified, how the intervention was hypothesized to work	3	6	1	4	2	9
Methods								
Trial design	6	Description of trial design (e.g., parallel, factorial), including allocation ratio	0	0	0	0	0	0
	7	Important changes to methods after trial started (e.g., eligibility criteria), with reasons	8	17	7	29	1	4
Participants	8	Eligibility criteria for participants	21	45	12	50	9	39
	9	Settings and locations of intervention delivery and where the data were collected	2	4	1	4	1	4
Interventions	10	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	10	21	4	17	6	26
	11	Extent to which interventions were actually delivered by providers and taken up by participants as planned	27	57	16	67	11	48
Outcomes	12	Completely defined pre-specified outcomes, including how and when they were assessed	6	13	2	8	4	17
	13	Any changes to trial outcomes after the trial commenced, with reasons	4	9	4	17	0	0
Sample size	14	How sample size was determined	19	40	13	54	6	26
	15	When applicable, explanation of any interim analyses and stopping guidelines	0	0	0	0	0	0
Randomization								
Sequence generation	16	Method used to generate the random allocation sequence	25	53	19	79	6	26
	17	Type of randomization; detail of any restriction (e.g., blocking and block size)	8	17	6	25	2	9
Allocation concealment mechanism	18	Mechanism used to implement the random allocation sequence, describing any steps taken to conceal the sequence until interventions were assigned	8	17	5	21	3	13
Implementation	19	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	6	13	4	17	2	9
Awareness of assignment	20	Who was aware of intervention assignment after allocation (e.g., participants, providers, those assessing outcomes), and how any masking was done	6	13	4	17	2	9
	21	If relevant, description of the similarity of interventions	27	57	13	54	14	61
Analytical methods	22	Statistical methods used to compare group outcomes	1	2	1	4	0	0
	23	How missing data were handled, with details of any allocation method	9	19	4	17	5	22
	24	Methods for additional analyses, such as subgroup analyses, adjusted analyses, and process evaluations	3	6	2	8	1	4
Results								

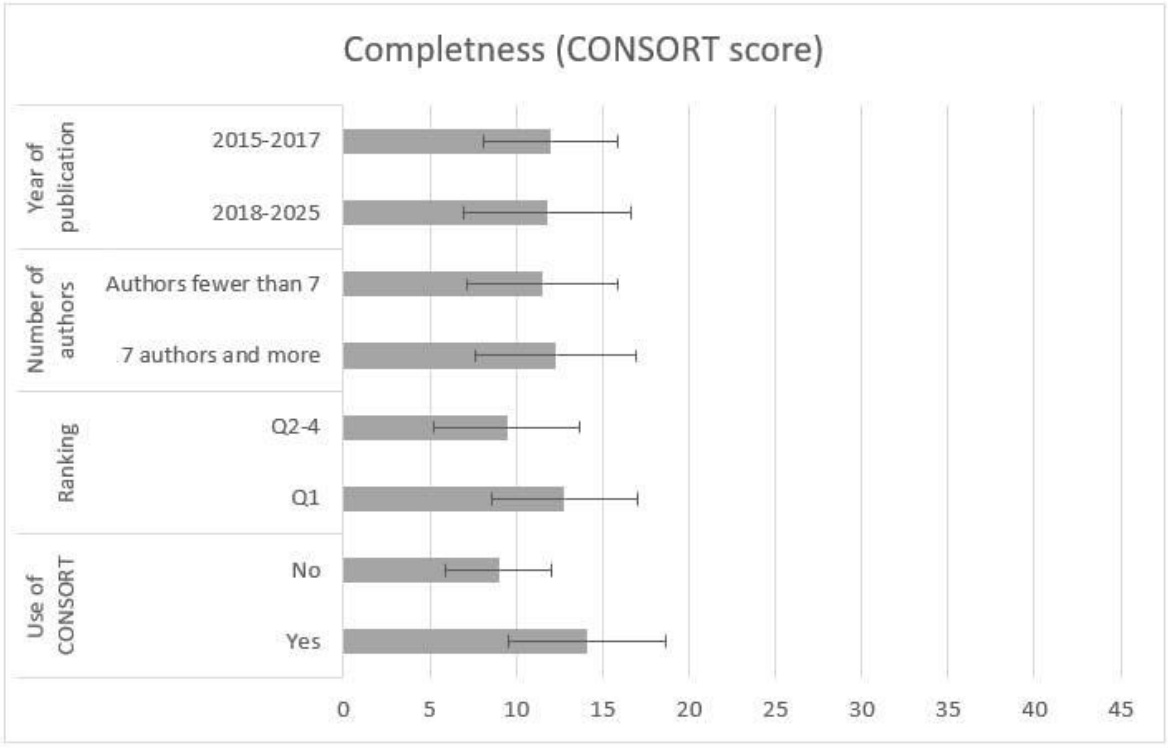
Participant flow (a diagram is strongly recommended)	25	Where possible, the number approached, screened, and eligible prior to random assignment, with reasons for non-enrolment	8	17	7	29	1	4
	26	For each group, losses and exclusions after randomization, together with reasons	12	26	9	38	3	13
Recruitment	27	Dates defining the recruitment and follow-up periods	2	4	1	0	1	4
	28	Why the trial ended or was stopped	0	0	0	0	0	0
Baseline data	29	Include socioeconomic variables where applicable	31	66	22	92	9	39
Numbers analyzed	30	For each group, number included in each analysis and whether the analysis was by original assigned groups	12	26	8	33	4	17
Outcomes and estimation	31	For each outcome, results for each group, and the estimated effect size and its precision (e.g., 95% confidence interval)	15	31	12	50	3	13
	32	Indicate availability of trial data	10	21	5	21	5	22
	33	For binary outcomes, the presentation of both absolute and relative effect sizes is recommended	0	0	0	0	0	0
Ancillary analyses	34	Results of any other analyses performed, including subgroup analyses, adjusted analyses, and process evaluations, distinguishing pre-specified from exploratory	3	6	2	8	1	4
Harms	35	All important harms or unintended effects in each group	3	6	2	8	1	4
Discussion								
Limitations	36	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	43	91	23	96	20	87
Generalizability	37	Generalizability (external validity, applicability) of the trial findings	26	55	14	58	12	52
Interpretation	38	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	10	21	7	29	3	13
Important information								
Registration	39	Registration number and name of trial registry	3	6	2	8	1	4
Protocol	40	Where the full trial protocol can be accessed, if available	15	32	12	50	3	13
Declaration of Interests	41	Sources of funding and other support; role of funders	16	34	7	29	9	39
	42	Declaration of any other potential interests	40	85	20	83	20	87
Stakeholder investments	43	Any involvement of the intervention developer in the design, conduct, analysis, or reporting of the trial	11	23	8	33	3	13
	44	Other stakeholder involvement in trial design, conduct, or analyses	3	6	2	8	1	4
	45	Incentives offered as part of the trial	20	43	11	47	9	39

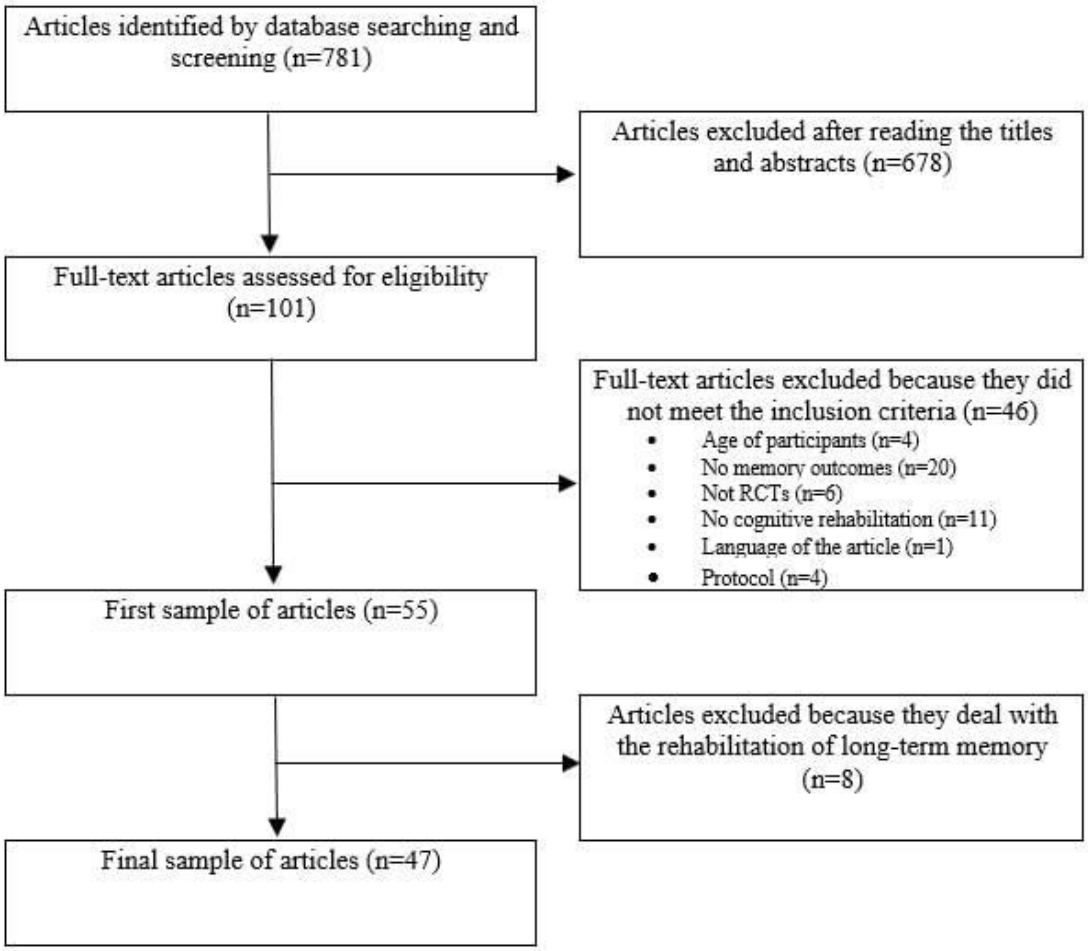
N = number of articles; CONSORT = articles mention the use of CONSORT; Not CONSORT = articles do not mention the use of CONSORT.

Table 5. *Risk of bias – RoB 2*

Domain	Risk of bias	N	%
Risk of bias arising from the randomization process	Low risk	15	32
	High risk	4	9
	Some concerns	28	60
Risk of bias due to deviations from the intended interventions	Low risk	8	17
	High risk	29	62
	Some concerns	10	21
Missing outcome data	Low risk	20	43
	High risk	0	0
	Some concerns	27	57
Risk of bias in measurement of the outcomes	Low risk	8	17
	High risk	17	36
	Some concerns	22	47
Risk of bias in selection of the reported result	Low risk	37	79
	High risk	5	11
	Some concerns	5	11
Overall risk of bias	Low risk	1	2
	High risk	36	77
	Some concerns	10	36

N= number of articles


Completness (CONSORT score)
Year of publication
2015-2017
2018-2025
Number of authors
Authors fewer than 7
7 authors and more
Ranking
Q2-4
Q1
Use of CONSORT
No
Yes
0
5
10
15
20
25
30
35
40
45


Articles identified by database searching and screening (n=781)
Articles excluded after reading the titles and abstracts (n=678)
Full-text articles assessed for eligibility (n=101)
Full-text articles excluded because they did not meet the inclusion criteria (n=46)
Age of participants (n=4)
No memory outcomes (n=20)
Not RCTs (n=6)
No cognitive rehabilitation (n=11)
Language of the article (n=1)
Protocol (n=4)
First sample of articles (n=55)
Articles excluded because they deal with the rehabilitation of long-term memory (n=8)
Final sample of articles (n=47)