Peer-Review comments and authors responses

"Caffeine and Cognitive Performance in Healthy Adults: A Systematic Review ofRandomized Controlled Trials"

Reviewer 1

Recommendation: Revisions Required
The manuscript entitled "Caffeine and Cognitive Performance in Healthy Adults: A Systematic Review of Randomized Controlled Trials" addresses an important and timely topic. Caffeine is one of the most widely consumed psychoactive substances worldwide, and its potential effects on cognition have significant implications for public health, daily functioning, and scientific understanding of neurocognitive processes. By focusing specifically on randomized controlled trials in healthy adult populations, this review adds rigor to the evidence base and helps clarify inconsistencies across individual studies. The study adheres to PRISMA guidelines and applies standard tools (Covidence, RoB 2), which is commendable.

Overall comments

- The introduction's section presents a good rationale and plausible justification for carrying out a systematic review.
- The aims (primary and secondary objectives) are well stated
- The study is methodologically sound and addresses a relevant research question. The main areas for improvement involve greater transparency of the search strategy, clearer justification of eligibility restrictions, and more detailed handling with data synthesis. With these revisions, the manuscript would make a strong contribution to the field.
- The results' section is comprehensive and well-structured, providing clear distinctions between primary and secondary outcomes. The inclusion of mood, arousal, psychomotor vigilance, and dual-task performance as secondary outcomes strengthens the validity by connecting cognition with real-world functioning.
- The inclusion of both narrative description and structured tables makes it easier for readers to follow the evidence.

Comment 1: Row 109, For the first time it refers to the "dose-dependent cognitive effects of caffeine (100–400 mg)". For clarity, you could insert "daily"

***Response**: We thank the reviewer for this helpful suggestion. The text has been revised for clarity.*

Methods

Comment 2: Row 147, "The search strategy was developed using Medical Subject Headings (MeSH) terms and refined based on our population, intervention, control, and outcomes (PICO) framework. We included articles published in English, with the following terms: [Caffeine OR Coffee OR Caffeinated OR caffeine low dose OR caffeine high dose] AND [Cognition OR Cognitive Function OR Cognitive Performance OR Executive Function OR Memory OR Attention OR Cognition] AND Placebo."
It seems that not all terms are MESH. Given the importance of reproducibility and the need for transparency, I strongly suggest including the full search strategies for each database as a supplementary file. The language restriction to English should be acknowledged as a limitation, as it introduces potential language bias.

Response: *The complete search strategies for each database (PubMed, Embase, Scopus, and Web of Science) have been added as Supplementary Table 1 (Table S1). Complete Search Strategies by Database. In addition, we have acknowledged the English-language restriction as a limitation in the "Implications for practice and future research" section.*

Results

> **Comment 3:** Figure 1 - In the text box "references removed," the "other reasons" are not showing any results

Response: *We have corrected the missing value in Figure 1.*

> **Comment 4:** Row 206: You stated that "several studies that initially appeared to meet the inclusion criteria were excluded after full-text review." Although the flowchart displays the reasons for exclusion (grouped according to PRISMA guidelines), the Cochrane Handbook recommends that systematic reviews provide a list of excluded studies at the full-text screening stage, along with the specific reasons for exclusion, as this demonstrates consistency in the decision-making process. Typically, this list is presented in an appendix or supplementary material, with each excluded study cited (author/year or full reference) and the corresponding exclusion reason.

Response: *Following the reviewer's suggestion, a full list of excluded studies and reasons for exclusion has been added as Supplementary Table 2 (Table S2). Full-text Articles Excluded with Reasons. This addition ensures transparency and compliance with PRISMA 2020 and Cochrane Handbook standards. A corresponding sentence was added in Methods section, Study selection process.*

> **Comment 5:** Row 232: "All included studies administered caffeine in doses ranging from 100 to 400 mg, consistent with the review's predefined thresholds for low (100–200 mg) and high (300–400 mg) exposure." Regarding the predefined thresholds terminology: Where do the thresholds used by the authors come from? Is there a reference, or were they chosen arbitrarily? Caffeine effects can vary significantly by body weight and habitual consumption. I was wondering if it wouldn't be important to inform whether they were standardized or justified.

Response: *We included an explanation in the Methods section, eligibility criteria section (lines 118-125).*

Discussion

> **Comment 6:** Row 386 - The rationale provided for not conducting a meta-analysis, based on mixed and unclear RoB-2 assessments, is reasonable and highlights valid methodological concerns. The limitations section appropriately acknowledges important issues related to variability in cognitive outcomes and dosing regimens. However, I would encourage the authors to also acknowledge the considerable heterogeneity in study outcomes, cognitive tasks, and caffeine dosing protocols as additional limitations that restrict not only the generalizability but also the feasibility of pooling results. Explicitly stating both methodological quality issues and outcome heterogeneity would provide a more comprehensive justification for the narrative synthesis approach.

Response: *A paragraph acknowledging inter-study heterogeneity has been added to the Discussion section (lines 476-483).*

Reviewer 2

Recommendation: Revisions Required

Dear authors, thank you for the opportunity to review this valuable manuscript. It is well written and of good quality, requiring only some refinements.

Major comments

Methods

> **Comment 1:** Eligibility consistency (age range) and adherence. Methods specify adults 20–60 years yet included samples and other sections at times imply 18–60 or include participants outside the stated band. Please harmonize eligibility criteria everywhere (Abstract, Methods, Results, Tables) and, if any included study deviates, either justify the exception or adjust the criterion to match actual inclusion.

Response: We thank the reviewer for this suggestion. The right criteria are adults between 18-60 years, and it is now corrected in the reviewed draft.

Methods

> **Comment 2:** Reproducibility of the search (PRISMA-S).
> Report full database-specific search strings (including all Boolean operators/filters), the exact search dates, language limits, and the de-duplication process. If the review was not registered (e.g., PROSPERO), state this explicitly. Please also clarify the number of independent reviewers for screening/extraction and how disagreements were resolved.

Response: We have included this information in Supplementary Table 1 (Table S1). Complete Search Strategies by Database.

> **Comment 3:** Scope and outcome discipline. The Methods exclude athletic/performance-specific outcomes and sleep-related outcomes, but a few included trials touch on performance contexts or expectancy/time-of-day manipulations. Confirm that only cognitive endpoints meeting the inclusion criteria were extracted from such trials; if any exceptions were made, justify and acknowledge as a limitation.

Response: We confirm that only cognitive endpoints meeting the inclusion criteria were extracted from the articles chosen. We have added Supplementary Table 2, where the reviewer can also check the articles excluded and the reason for exclusion.

Results

> **Comment 4:** Link risk of bias to strength of conclusions. RoB-2 results are presented but not integrated into the interpretation. Please add a paragraph that summarizes the proportion of low / some concerns / high risk studies for each domain and indicates how this influences the certainty of evidence (e.g., "low-to-moderate confidence for attention; low for EF"), especially where funding or reporting issues were noted.

Response: We have included the interpretation of the RoB into our discussion.

> **Comment 5:** PRISMA completeness and exclusions. The flow diagram is reported; please add a concise table of exclusion reasons (e.g., non-RCT, wrong population, not acute, not cognitive endpoint) with counts, and ensure the numbers in the diagram and text match exactly.

Response: We have added Supplementary Table 2, where the reviewer can review the articles excluded and the reason for exclusion.

Conclusions

> **Comment 6:** Conclusions overreach. The manuscript currently claims that caffeine "supports cognitive enhancement in daily life." Given the mixed tasks and doses, possible expectancy effects,

and tolerance in habitual users, this is too strong. Please revise to state that benefits are context-dependent—for instance: benefits appear context-dependent, with the most consistent effects in attention/reaction time at 100–200 mg (as reflected in the included trials), while executive-function results are inconsistent.

Response: We thank the reviewer for pointing out this claim in our conclusions. We have revised this and made the necessary changes. However, the reviewer can find in the Discussion section that we extensively discuss the considerations regarding our findings and how habitual high-dose consumers may experience tolerance and altered metabolic responses.

Minor comments

Abstract

Comment 7: Align age range with Methods; state the number of databases and date of last search; avoid mechanistic speculation not assessed by the review.

Response: We have corrected the age range throughout the manuscript. The databases used were indicated (line 134). The last date of search is already indicated in line 138.

Methods

Comment 8: Specify that both parallel and cross-over RCTs were eligible and note how carryover risks/washouts were handled in RoB-2. Define what effect metrics were extracted (means/SDs, change scores) and any conversions.

Response: *We have added the inclusion of parallel and crossover RCTs as eligible for our systematic review (line 113). In the case of washout periods in crossover studies, we included only articles in which the washout period was clearly defined. This helps ensure that the effects observed in the current intervention are not influenced by the previous treatment. We added this clarification in the text (line 175). We also included in the text which metrics were extracted from the articles (line 184-187).*

Results

Comment 9: Provide a brief sample characterization table (median trial n, mean age, % female, habitual caffeine distribution if available).

Response: *We have included a new table with the information requested above (Table 1).*

Comment 10: Tables/Figures: Ensure each study appears once per domain with consistent effect direction; standardize domain headings; check year/author concordance.

Response*: We thank the reviewer for pointing this out. We have completed Table 3 with the missing studies.*

Other

Comment 11: Funding/COI: Where applicable, explicitly note industry involvement among included trials and discuss as a potential bias.

Response: *The studies included in this systematic review were funded by governments, universities, or non-profit organizations/foundations. Some of them declared not receiving financial support and only 3 of them did not mention anything regarding funding.*

Comment 12: Language/style: Tighten repetitive sentences and standardize terminology ("executive functions," "working memory," "attention / processing speed").

Response: *We thank the reviewer for highlighting this point. We have corrected and standardized terminology throughout the text.*

Reviewer 3

Recommendation: Revisions Required

Dear Editor, I am thankful for this opportunity to peer review the manuscript titled: "Caffeine and Cognitive Performance in Healthy Adults: A Systematic Review of Randomized Controlled Trials."
I congratulate all the authors on their work and the extensive research needed to be put together for this manuscript. Below are my few observations recorded in my review.

Comment 1: Study design: Why weren't observational studies included in this review, where the confounders could have been strongly accounted for? (These effects could be observed with daily habitual caffeine consumers as well, right?) Were there no studies conducted? An observational study that strongly accounts for confounders also has the same validity as an RCT. The review could have an added value with observational and RCT being included, it is just my curiosity as well to know the reason.

Response: *We thank the reviewer for this comment. The decision to exclude observational studies from this review was primarily driven by the objective of this review: to assess causal relationships between caffeine consumption and cognitive enhancement. Randomized controlled trials (RCTs) are recognized as the gold standard for causal inference, as randomization minimizes both measured and unmeasured confounding.. In contrast, observational studies, although valuable, are inherently more susceptible to biases and confounding factors, which can complicate causal interpretation. Additionally, although well-designed observational studies can account for confounders employing rigorous adjustment methods (e.g., propensity scoring, multivariable regression), RCTs inherently minimize confounding by design. The exclusion of observational studies aimed to maintain a focus on the highest level of evidence regarding caffeine's causal impact. Future systematic reviews could expand on this work by integrating both RCTs and well-designed observational studies to enhance the generalizability of findings by examining broader populations with different caffeine consumption levels and demographic characteristics.*

Methods

Comment 2: Section: population Line: 220-221 "Most trials enrolled non-smoking participants without psychiatric, neurological, or metabolic disorders, and systematically excluded those with caffeine hypersensitivity, substance use, or the use of interfering medications" – I would like to ask the authors why these criteria weren't mentioned before in the exclusion criteria of the study?

Response: *The exclusion criteria that we chose are clearly stated in the Methods section. What the reviewer mentioned above is the observation that we made when analyzing the results of the articles that we reviewed for this manuscript. We acknowledge that these patients' characteristics mentioned could have been included in our exclusion criteria, but at the time of defining this, we did not consider them.*

Comment 3: Individuals with prior cognitive impairment of any kind – even a short term / any neuropsychiatric conditions could impair the results, just the age groups below 18 & > 60 were mentioned as an exclusion, however participants even in this population group between 18-60 can have significant memory / cognitive impairment, which ought to be in exclusion criteria reaffirming the elimination of bias or confounders even though the studies have enrolled healthy population.

Response: *In fact, any neuropsychiatric condition could be a confounding factor that could impair the result. Even though we did not add it explicitly as an exclusion criterion (by healthy individuals, we meant to exclude any individual with neuropsychiatric conditions), we should have stated more clearly in our exclusion criteria. However, the studies included in this systematic review excluded those patients from their study, so the results are not affected by this potential confounder.*

Comment 4: Since most of the studies included smaller sample size / participants, did the studies also consider conditions like GERD / esophagitis which could be an underlying benign not very troublesome condition but could be exacerbated by even low doses of caffeine? Since the population is adults between 20-60 most middle aged adults usually have stress related gastritis as well, was this taken into consideration ? if the results – which shows improvement in cognition and memory , shouldn't on the other hand be worsening something else due to their lifestyle or an underlying mild disorder?

Response: *None of the included RCTs specifically assessed GERD or related gastrointestinal conditions; most enrolled overtly healthy adults, often university students. While caffeine can transiently lower the lower esophageal sphincter pressure and may acutely trigger reflux symptoms in sensitive individuals, consistent evidence of clinically significant GERD worsening at typical doses (100–400 mg/day) is lacking. As the included studies were short-term and focused on cognitive outcomes, gastrointestinal effects were beyond their scope.*

Comment 5: Of the entire population, only about 5-6 studies roughly < 300 participants are aged 40 -60 years. I am asking since this is the age group which usually is prey to all kinds of lifestyle disorders with increase in stress and work related burn out, most women approaching menopause as well, and this would be the group which usually takes pain killers/antacids regularly and it is available over the counter – non prescribed, I do not see a mention of these meds not being taken, as well as don't you think the concomitant meds un prescribed like these, we should be more cautious about while prescribing caffeine? Also arguably this improvement in cognition and memory is much required in this age group as well, and this age group is very underrepresented in the review.

Response: *Indeed, adults aged 40–60 years were underrepresented across the included RCTs, which primarily enrolled younger, healthy participants such as university students. Most trials did not report or control for concomitant medication use, including over-the-counter analgesics or antacids. We agree that midlife adults—who often experience higher stress levels, perimenopausal changes, and polypharmacy—represent a population of particular clinical relevance. Future studies should specifically include this group and account for common medication use to better assess both cognitive effects and potential interactions with caffeine.*

<u>Results</u>

Comment 6: (TABLES): though the results of the dosages are explained in the section, however as per the study having varied doses of caffeine, not consistent in all the studies, the results if explained dose wise as in the group of studies which received low dose 100-200 mg, the ones which received moderate dose up to 300 mg, the ones which received high dose > 300-400 mg – the dose wise results in a tabular form would make the results more clearer, as to what was the population receiving the dose , how many participants and what was the effect of it.

Response: *We thank the reviewer for this comment. We did not separate results regarding caffeine dosage given that in some articles, the amount of caffeine was measured in mg/kg and not in total mg. In those cases, it was not possible to calculate the exact amount of caffeine consumed by the participants, therefore we decided not to include the results based on the amount of caffeine consumed.*

Comment 7: Section: Risk of bias (TABLES): Out of the 20 RCT selected you have only 3 studies with low risk, 4 with high risk and rest 13 with some concern. How do you explain the robustness of internal and external validity with very low participant inclusion of just 951 participants?, agreed it is multi ethnic in inclusion, diverse population but is the population number enough to suggest it to larger subset of the people considering the concerns with bias? , all confounders not noted or accounted for? all population age groups like ex: 40-60 having a lesser number representation?

Response: *We acknowledge that most included trials presented some concerns regarding risk of bias and that the overall sample size (n=951) limits the strength of generalizability. Nevertheless, internal validity was supported by the randomized, double-blind designs and standardized cognitive assessments used across studies. The heterogeneity in populations, settings, and caffeine administration enhances external validity but also introduces variability. We agree that future large-scale, well-powered trials—particularly including middle-aged adults—are needed to confirm these findings and address residual confounding.*

> **Comment 8:** Again, your bias in DOMAIN 4: measurement of outcome is 90% which you have mentioned in your ROB explanation as well, so how do you account for the robustness of the results in this case? Do you consider the results with low representation of population still valid enough in this review?

Response: *We appreciate the reviewer's thoughtful concern. Most studies employed validated cognitive tests under standardized, blinded conditions, supporting reasonable internal validity. We acknowledge that the limited sample size and moderate risk of bias warrant cautious interpretation, and our conclusions emphasize the need for larger, high-quality trials to confirm these findings.*

<u>Others</u>

> **Comment 9:** MINOR CORRECTIONS WORD LIMIT

1. Your introduction is too long comes to more than 800 words kindly cut short the introduction
2. Your entire manuscript comes up to 5182 words , as per the PPCR journal guidelines , the full text of the article should be maximum 4000 words , you are crossing the limit by 1182 words, the whole limit including references should be 5000 words – kindly cut short the length of the article as per the journals word limit guidelines

Response: *According to the suggestions, the revised manuscript has a shorter introduction and meets the word limit required by the journal.*

Reviewer 4

Recommendation: Revisions Required

Dear authors, it was a pleasure for me to review this manuscript, as I consider myself an avid coffee enthusiast. This systematic review aims to synthesize the existing evidence on the effects of caffeine on cognitive functions.

Strength-wise, it is very detailed and follows the PRISMA guidelines accordingly. The way you described each outcome measure used in the RCTs adds a degree of scientific writing that is worthy of praise for you as students of PPCR. Also, I believe it shows your learning throughout your assignments for this review. I congratulate you on an excellent writing effort.

However, the review has flaws in its writing, consistency, and methodology. Adding a section for the design of the review itself might clarify any methodological doubt. Aside from this, proofreading also reveals that certain sections were written without regard for past phrases or comments, making some sentences feel redundant/repetitive. Please address this by carefully reading through each section, specifically the introduction, results, and discussion. Given these changes, I endorse the publication of this review once the authors have adequately addressed all comments/suggestions.

Response: *We sincerely thank Reviewer 4 for the careful and insightful review of our manuscript. We greatly appreciate the constructive feedback, which has helped us to strengthen the methodological transparency and clarity of the paper. All comments were carefully considered, and corresponding revisions have been implemented throughout the manuscript. Below, we provide a point-by-point response, detailing the changes made and indicating where each revision can be found.*

> **Comment 1:** The review has flaws in its writing, consistency, and methodology. Adding a section for the design of the review itself might clarify any methodological doubt.

Response: *The section for the design of the review has been added at the beginning of the Method section, under the name of "Review process".*

> **Comment 2:** Certain sections were written without regard for past phrases or comments, making some sentences feel redundant/repetitive. Please address this by carefully reading through each section, specifically the introduction, results, and discussion.

Response: *We revised the introduction, main results, and discussion sections with regard to past phrases or comments, making sentences less redundant or repetitive.*