Response Letter

Dear Prof Zoltan Dienes,

Thank you for the opportunity to revise our study titled "*Can adults automatically process and translate between numerical representations*?". We sincerely appreciate all the constructive feedback, suggestions and comments. In this letter, our responses to each comment are in green. Changes in the manuscript have all been tracked. We have also uploaded a clean version of the revised manuscript to our OSF directory. We hope the revised manuscript meets with your approval and look forward to hearing from you at your earliest convenience.

Kind regards,
Iro

On behalf of all authors

Editorial comments:

Your article, entitled Can adults automatically process and translate between numerical representations?, has now been reviewed. The referees' comments and the recommender's decision are shown below. As you can see, the recommender suggests revisions.

We shall, in principle, be happy to recommend your article as soon as it has been revised in response to the points raised by the referees.

Thank you. Below, we address all the points raised by the referees. We are very grateful for the thorough review and feedback and acknowledge this now in our "Acknowledgements" section.

When revising your article, we remind you that:

1) Data must be available to readers, either in the text or through an open data repository such as Zenodo (free), Dryad (pay) or some other institutional repository. Data must be reusable, thus metadata or accompanying text must carefully describe the data;

Data and respective read-me guidance are provided in our OSF directory.

2) Details on quantitative analyses (e.g., data treatment and statistical scripts in R, bioinformatic pipeline scripts, etc.) and details concerning simulations (scripts, code) must be available to readers in the text, as appendices, or through an open data repository, such as Zenodo, Dryad or some other institutional repository. The scripts or code must be carefully described so that they can be reused;

R scripts and respective read-me guidance are provided in our OSF directory.

3) Details on experimental procedures must be available to readers in the text or as appendices;

All experimental procedures are outlined in detail in the manuscript. Accompanying material (e.g. stimuli, experimental scripts) are all provided in our OSF directory.

4) Authors must have no financial conflict of interest relating to the article. The article must contain a "Conflict of interest disclosure" paragraph before the reference section containing this sentence: "The authors of this article declare that they have no financial conflict of interest with the content of this article.";

Thank you for the detailed guidance. We have now added a "Conflict of interest disclosure" paragraph on p. 44 of our revised manuscript.

5) This disclosure has to be completed by a sentence indicating, if appropriate, that some of the authors are PCI recommenders: "X is a recommender at PCI Registered Reports.".

N/A

Revision round #1

Decision for round #1 : Revision needed
The role of working memory in translating between different number processing systems

The two reviewers are positive about the Stage 2, but make some important points. Reviewer 2 asks for some more clarity about the central executive; note that you cannot add the suggested extra sentences to the introduction, but there might be some way of addresing the point in the discussion itself. Reviewer 1 points out that the pre-registered protocol was not followed exactly with respect to the conditions you would use non-parametric tests. It is important to stick to the letter of the pre-registration.

Please see our responses in the respective points made by Reviewer 1 and Reviewer 2.

I checked your reported analyses against the Design Table and found other discerepancies:

p 28
"Finally, comparing the two interference conditions (PL and VSSP), we found no difference either for accuracy, t(80) = 0.58, p = .57, d = 0.06, or median RT, t (78) = 0.09, p = .93, d = .01."
and also

". Finally, performance in the two non-symbolic interference conditions (PL and VSSP) significantly differed, t(79) = 2.44, p = .02, d = 0.27, with participants performing better under PL than VSSP interference. Median RTs did not differ between these two conditions, t(78) = 1.48, p = .14, d = 0.17."
These are not in the Design Table; place in a separate exploratory section.

These analyses were preregistered in the detailed analysis plan under the sections for the symbolic, RQ1a in the detailed analysis plan, p. 56-57, and nonsymbolic condition, RQ1b in the detailed analysis plan, p.58-59. We have highlighted these in the detailed analysis plan for ease in the tracked-changes version of the manuscript. However, we understand how these could have been missed as the wording that we use in the results section in the Stage 2 manuscript is rather simpler compared to the original analysis plan for ease of clarity and understanding. To address this, on p.28 we have now also included in brackets how this is worded/referred to in the analysis plan for clarity. Specifically, it now reads:

"Finally, comparing the two interference conditions (i.e., symbolic primary task with PL dual-task vs symbolic primary task with VSSP dual-task)…"

Two things to bear in mind in drawing conclusions: Only conclude there was a difference between conditions if the difference was tested (and only draw such conclusions in the abstract or give them importance in the discussion if they were pre-registered); and also do not assert there was no effect because of a non-significant result unless you calculated power with a justified minimal effect for precisely that test.
p 34 "We also found that mechanisms for processing and translating numerical representations can differ for smaller (1 - 4) and larger (5 – 9) numerosities."
As the difference between numerosities was not tested this conclusion does not follow.

Thank you for pointing this out – we have now revised the sentence in p.35 as follows:
"Also, we observed different interference effects in processing and translating numerical representations for smaller (1 - 4) and larger (5 – 9) numerosities, which were examined separately."

With this sentence, we aim to summarise the findings regarding the effects of PL and VSSP interference on small and large numerosities. Indeed, some of these comparisons were preregistered but some were not – this distinction is highlighted in Figure 5, which provides the overview of all the comparisons.

p 35
"However, these speed-accuracy correlation coefficients did not differ significantly, which suggests that there was no significant change in participants' speed-accuracy trade-off across the conditions. Thus, a speed-accuracy trade-off on its own cannot explain the finding of improved accuracy in interference conditions."

A non-significant result without a properly justified power analysis for precisely that test does not justify asserting no effect.

Thank you for pointing this out – we have now revised the sentence on p.36: "*However, these speed-accuracy correlation coefficients did not differ significantly, which may suggest no change in participants' speed-accuracy trade-off across the conditions. Thus, a speed-accuracy trade-off on its own does not appear to explain the finding of improved accuracy in interference conditions. We should highlight here though that this is only a provisional assumption based on exploratory analyses for which we had not conducted power analysis, therefore respective assumptions should be considered tentative .*"

abstract
"albeit involving different components of WM, to a different extent."
Testing differences between dot and digit tasks, or between their respective dual task conditions, were not pre-registered; so claims about such differences should not appear in the abstract.

We agree and thank you for pointing this out. Indeed, we did not compare the symbolic and non-symbolic conditions. We had only preregistered and compared performance between the two interference conditions in each type of numerical representation. In the case of the non-symbolic condition, we found that performance was significantly worse in the non-symbolic task with VSSP interference vs the non-symbolic with PL interference. The way that the sentence in question was formulated could indeed be misleading, we, therefore, corrected it now as follows:

"We found that all three types of magnitude comparison necessitated WM resources, albeit in distinct ways"

One further point of clarity|:

"Surprisingly, in this task, accuracy improved"

This has been clarified now as follows: "*Surprisingly, in this task, accuracy improved under both WM interference conditions*"

Manuscript: https://osf.io/x76tz?view_only=bf8c069e022540aa9272452804f27db2 version: 1

Review by Hannah Dorothea Loenneker, 14 Oct 2024 20:04

The authors submitted a thorough and transparent stage 2 registered report. As the stage 1 version was already based on an in-depth consideration of theoretical implications and the current empirical evidence, stage 2 follows up with a structured performance of the respective experiments and a comprehensible discussion in the light of relevant theories. The current study adds to available literature as it clarifies which component of the working memory is related to

processing of and translating between different numerical modalities. This is obtained using a dual-task paradigm of (cross-modal) numerical comparison and working memory tasks. As the authors found working memory involvement to be crucial, they discussed that processing (non)symbolic magnitudes and translating between symbolic and non-symbolic numerical representations cannot be fully automatized.

I would like to stress that, on top of detailed reporting of results in the main text, the Supplementary Material provides interesting additional information. The authors also transparently reported deviations between stage 1 and stage 2.

Minor points:

The use of the past tense in the section "The present study" in the introduction is not entirely consistent so you might want to check whether you translated all present and future tense declinations from stage 1 to past tense for stage 2.

Thank you for the thorough and rigorous review - it is very much appreciated. This has now been corrected.

Formatting regarding paragraph indentation differs between sections – please unify

Corrected.

Page 18: "In the cross-modal comparison task, the side of presentation for the Arabic symbol was be counterbalanced." – there is a linguistic error in transforming the old future tense into the past tense, needs to be "was counterbalanced".

Corrected.

Also page 18: "("z" if the left quantity is larger, "m" if the right quantity is larger)" – should be in the past tense "if the left quantity was larger"

Corrected.

In your planned analyses section, you stated that you would use Wilcoxon tests for variables showing a skew > 3. However, in the results section, you only report t statistics despite several variables exceeding this threshold in RQ1a and 1b. However, for RQ2 you do report non-parametric results. Therefore, it remains unclear to the reader when you used parametric or non-parametric tests.

Response:
Thank you for pointing this out. The descriptive data provided in Table 1 in the Stage 2 manuscript included the extreme outliers, which, as preregistered, were then excluded from the analyses. Thus, the skewness data reported in that Table were not the ones that had informed the choice of parametric vs non-parametric analysis. Apologies for the confusion. In the revised manuscript, this is clarified in p.23 and the data in Table 1

now depict the descriptive statistics for all primary and secondary task performance variables after having excluded the extreme outliers for each variable. This table now also includes a column indicating the sample size (n) for each variable after the removal of the outliers. It should be noted that decisions regarding the choice of a parametric or non-parametric test were based on the normality checks of the data included in a particular comparison. For example, the variable for symbolic comparison accuracy with VSSP interference (n = 81) violates the kurtosis criterion (>4) even after removing extreme outliers. However, when comparing this with symbolic comparison accuracy without interference, an additional two participants were excluded (who were extreme outliers on the symbolic comparison accuracy without interference variable). This left 79 participants, for whom the kurtosis criterion was not exceeded on either variable and therefore a t-test was performed. To clarify this we added the following text on p. 28 of the revised manuscript:

"Note that for some bivariate comparisons the data did not exceed kurtosis limits once all relevant exclusions had been performed and therefore a t-test was performed even if the individual variables exceeded this limit."

Could you add descriptive statistics for the comparison small vs. large quantities as well in the main text?

Response: Thank you, this table (Table 2) has now been included in the main text right after Table 1.

The graphical summary of all results was very helpful. The figure would be easier to read if you added a note explaining the arrows a little bit more, i.e., stating that an arrow facing upwards indicates better accuracy/ slower reaction time for dual and an arrow facing downwards indicates worse accuracy/ faster reaction time for dual.

Response: Thank you, this has now been clarified in the caption.

Page 32: "Non-symbolic comparison necessitated VSSP WM but also PL albeit to a lesser extent." – on which result is this "to a lesser extent" based on? Did you compare effect sizes?

Response: Thank you for pointing this out. This conclusion is derived from several results: 1) Performance in the two non-symbolic interference conditions (PL and VSSP) significantly differed, t(79) = 2.44, p = .02, d = 0.27, with participants performing better under PL than VSSP interference (p. 29). In other words, VSSP interference affected performance more than PL interference, 2) The observation of the overall interference effects that we tested: Under VSSP interference, performance both in the primary and secondary task was affected, whereas under PL interference only primary task performance. This is now clarified in the manuscript on page 35.

"Under VSSP interference, performance was reduced in both the primary and the secondary task, whereas under PL interference only in the primary and significantly less than the VSSP."

Page 34: "On the contrary, we observed widespread interference effects in this primary task (see Figure 5)" this sentence lacks a full-stop & "Thus, it appears that attentional mechanisms can mitigate WM interference under certain conditions." – text does not need to be underlined

Response: Corrected. Thank you for the thorough review, which is highly appreciated.

Review by Xinru Yao, 17 Nov 2024 14:15

This is a fascinating study that explores a common yet often overlooked phenomenon: whether symbolic and non-symbolic representations are processed and translated automatically or require the involvement of working memory. It also exmines how the processing of numerical information differs between small and large quantities. The study addresses an important research gap and contributes valuable insights to our understanding of numerical cognition. I like the introduction of this report, which provides us a clear overview of numerical representation and theoretical framework of the current study. Also, the method section is well-detailed with clear operational definitions as well as power analysis.

One minor point i noticed while reading is that in the introduction part, you briefly mentioned central executive (CE) of working memory once in Page 11, but the relationship between the CE and the other two components (VSSP and PL) is not explained. However, in discussion part, you discussed some about CE with VSSP and PL. For me, the mention of CE in the discussion feels a bit abrupt. It might be helpful to add one or more sentences in Page 11 to clarify the relationship between CE and VSSP or PL. This would provide better context and make the discussion more cohesive.

Response: We completely agree; this was an omission in the introduction, but we are conscious of the fact that we cannot make changes to the introduction at this stage. Instead, as helpfully suggested by the Editor, we have included two sentences at the beginning of the Discussion section which we hope clarify the role of the CE in our design (pp 34-35):

"Interference was implemented by asking participants to perform the numerical comparison tasks while also undertaking a WM task where they had to retain either letters that they heard (PL) or visuospatial patterns that they saw (VSSP) and then recall them backwards. As these secondary tasks required the manipulation of the stored elements, they also involved the CE component of WM, i.e. the limited-capacity attentional system overseeing and coordinating the activities of the PL and VSSP (Baddeley, 2001; Hitch, Allen, & Baddeley, 2024; Repovš & Baddeley 2006;)."

For limitations and future directions: Beyond its methodological strengths and contribution to understanding the causal mechanisms underlying basic number processing skills, does this study offer any guidance for practical applications?

Response: Thank you for the suggestion. We have added a paragraph regarding this in the "Limitations and Future Directions section" on p.42-43.

A small suggestion for Appendix C: Since the table spans multiple pages, consider repeating the header row with explanations at the top of each new page. This will make it easier for readers to navigate and understand the content without needing to refer back to the first page.

Response: Thank you for your thoughtful suggestion regarding Appendix C. We have now adjusted the table accordingly and hope it is now clearer and easier to navigate.