

# Revision round #1

Decision for round #1 : **Revision needed**

## Invitation to revise

---

Dear Dr. Karakashevska,

Thank you for submitting your Stage 2 manuscript to PCI-RR, we're looking forward to working with you again! The Stage 2 is well-prepared according to PCI-RR guidelines, however before we send the manuscript out for review, there are a few comments we'd like you to address. Looking forward to hearing from you soon.

Best,

Grace & Zoltan

The comments below are in-line with PCI-RR's review criteria of a Stage 2:

**2A. Whether the data are able to test the authors' proposed hypotheses (or answer the proposed research question) by passing the approved outcome-neutral criteria, such as absence of floor and ceiling effects or success of positive controls or other quality checks.** For the most part you have clearly fulfilled your plans for quality checks, however the edit on Page 20 referring to the behavioral performance needs to be reinstated as it was written in the IPA: The IPA stated that "Prior to analysis of the ERP data, we will check the behavioural performance on the luminance discrimination task. Any participants whose performance is below 80% on any block will be replaced. Given the results of Karakashevska et al. (forthcoming), it is likely that most participants will be over 95% correct on this task.". The Stage 2 currently states "Prior to analysis of the ERP data, the behavioural performance on the luminance discrimination task was checked to ensure participants were able to do the task without difficulty." It is not clear from this edit whether you upheld their quality check of the behavioral data outlined in the IPA. It is also not clear from the results section what the variance was within the behavioral data. Did any participant perform below 80% on any block? Furthermore, with the one participant whose behavioral data was not collected, was their data included in the EEG analysis? If their data is included, how did you confirm this participant was engaging with the task in the manner you registered in the IPA?

Thank you for noticing our mistake. There was some confusion between the authors here. Makin deleted the sentence about "Any participants whose performance is below 80% on any block will be replaced. Given the results of Karakashevska et al. (forthcoming), it is likely that most participants will be over 95% correct on this task."

I did not do the planned quality check to ensure that performance was always above 80% correct. This was an oversight on our part.

We have now combed through the behavioural data to avoid any inconsistencies. There were other minor oversights in the behavioural analysis.

We have now done the planned quality control checks and corrected this. Nearly all 120 participants met the pre-registered 80% inclusion criteria. However, one participant fell below in the baseline block (77.34%) two fell below in the monocular block (70.31% and 67.69%) and two fell below in the moving frame block (76.17% and 71.1%).

In total there were four participants who did not meet the 80% inclusion threshold in each block. The mentioned participant with missing behavioural data in the baseline block was almost certainly engaging during that block. This is likely given that their average performance was 95.83% correct in the other three blocks.

When double checking the behavioural data, we also identified another participant's behavioural data file was corrupted in one block (likely due to transfer of files). Their average performance was 85.83% correct in the remaining three blocks. There were also likely engaging during the missing block.

We apologise for this oversight and minor mistakes. We acknowledge that the submitted manuscript was not sufficiently transparent about the inclusion criteria of the behavioural data. We feel it would be justifiable, given resource constraints and unexpected results, to note this minor and inconsequential deviation from the original analysis plan in the revised manuscript.

We propose including the following in the manuscript:

*“The original plan was for participants whose performance fell below 80% correct on any block to be replaced. This was justified by the behavioural results of Karakashevska et al. (forthcoming). However, a minority of participants fell just short of the 80% threshold (77.34% for one participant in the baseline block, 70.31% and 67.97% for two participants in the monocular block, and 76.17% and 71.09% for two participants in the moving frame block). One participant had missing behavioural data in the baseline block, and one had missing behavioural data in the moving frame block. They were very likely above threshold, given that their average performance was >80% in the other three blocks. We believe it is justifiable to slightly relax the original inclusion threshold. Removing these 4 participants had almost no effect on SPN results, and it is very unlikely that replacing them would substantially alter the picture.”*

We believe this approach is adequate. However, we will happily replace the participants if the recommenders or editors feel it is necessary to uphold standards.

The results section with the 4 behavioural subthreshold participants removed would change like this (original statistics crossed out):

*“Hypothesis 1 was supported. Amplitude was more negative for symmetrical than asymmetrical patterns in the frontoparallel conditions across all 4 blocks ( ~~$F(1,119) = 226.91, p < .001, \eta p^2 = 0.656$~~ ) ( $F(1,115) = 219.119, p < 0.001, \eta p^2 = 0.651$ ).*

*Hypothesis 2 was also supported. The SPN was stronger (more negative) in the Baseline frontoparallel condition compared to the Baseline perspective condition ( ~~$t(119) = -4.39, p < .001, d_z = -0.401$~~ ) ( $t(115) = -3.995, p < 0.001, d_z = -0.371$ ).*

*Hypothesis 3 was not supported. There was no difference between the perspective cost in Baseline and Monocular blocks ( ~~$t(119) = 0.05, p = .961, d_z = 0.004$~~ ) ( $t(119) = 0.24, p = 0.812, d_z = 0.022$ ); Baseline and Static frame blocks ( ~~$t(119) = -0.84, p = .403, d_z = -0.077$~~ ) ( $t(115) = -0.56, p = 0.578, d_z = -0.052$ ); Baseline and Moving frame blocks ( ~~$t(119) = -0.99, p = .322, d_z = -0.091$~~ ) ( $t(115) = -0.69, p = 0.490, d_z = -0.064$ ) or Static frame and Moving Frame blocks ( ~~$t(119) = -0.15, p = .878, d_z = -0.014$~~ ) ( $t(115) = -0.13, p = 0.898, d_z = -0.012$ ).*

*Hypothesis 4 was not supported. Perspective cost in the Moving frame block was -0.23 microvolts. This was significantly below 0 ( ~~$t(119) = -2.23, p = .014, d_z = -0.204$~~ ) ( $t(115) = 2.280, p = 0.012, d_z = -0.212$ ), and not significantly above -0.35 ( ~~$t(119) = 1.04, p = .151, d_z = 0.095$~~ ) ( $t(115) = 0.899, p = 0.185, d_z = 0.083$ ) *p* In other words, there was a small but significant perspective cost in the Moving frame block, contrary to our predictions.*

In the results section, we now report the range as well as mean behavioural performance in each block:

*“However, this effect was small in absolute terms. The mean percentage of correct responses were: Baseline (Mean = 95.90, SD = 3.71, min = 77.33, max = 100); Monocular (Mean = 93.91, SD = 5.23, min = 67.97, max = 100); Static frame (Mean = 96.43, SD = 3.35, min = 82.03, max = 100); Moving Frame (95.78, SD = 4.71, min = 71.09, max = 100).”*

Furthermore, we note that there was no correlation between behavioural performance and an SPN amplitude in any block in our new exploratory analysis section:

*“Finally, we examined correlations between individual SPN amplitudes, perspective cost, and behavioural performance (two participants with missing behavioural data were excluded from this analysis). The correlation matrix in Figure 12 shows that participants who had a large SPN in one condition tend to have a larger SPN in the other conditions (blue cells in top left). Unsurprisingly, participants had a larger frontoparallel SPN tend to have a larger perspective cost, while those with a larger perspective SPN tended to have a smaller perspective cost (alternating red and blue steps near diagonal). Those who performed well on the luminance task in one block tended to do so on other blocks (blue cells bottom right). However, there was no correlation between behavioural performance and*

ERP signals (maximum effect  $r = 0.174$ ,  $p = 0.06$ , uncorrected for multiple comparisons).”

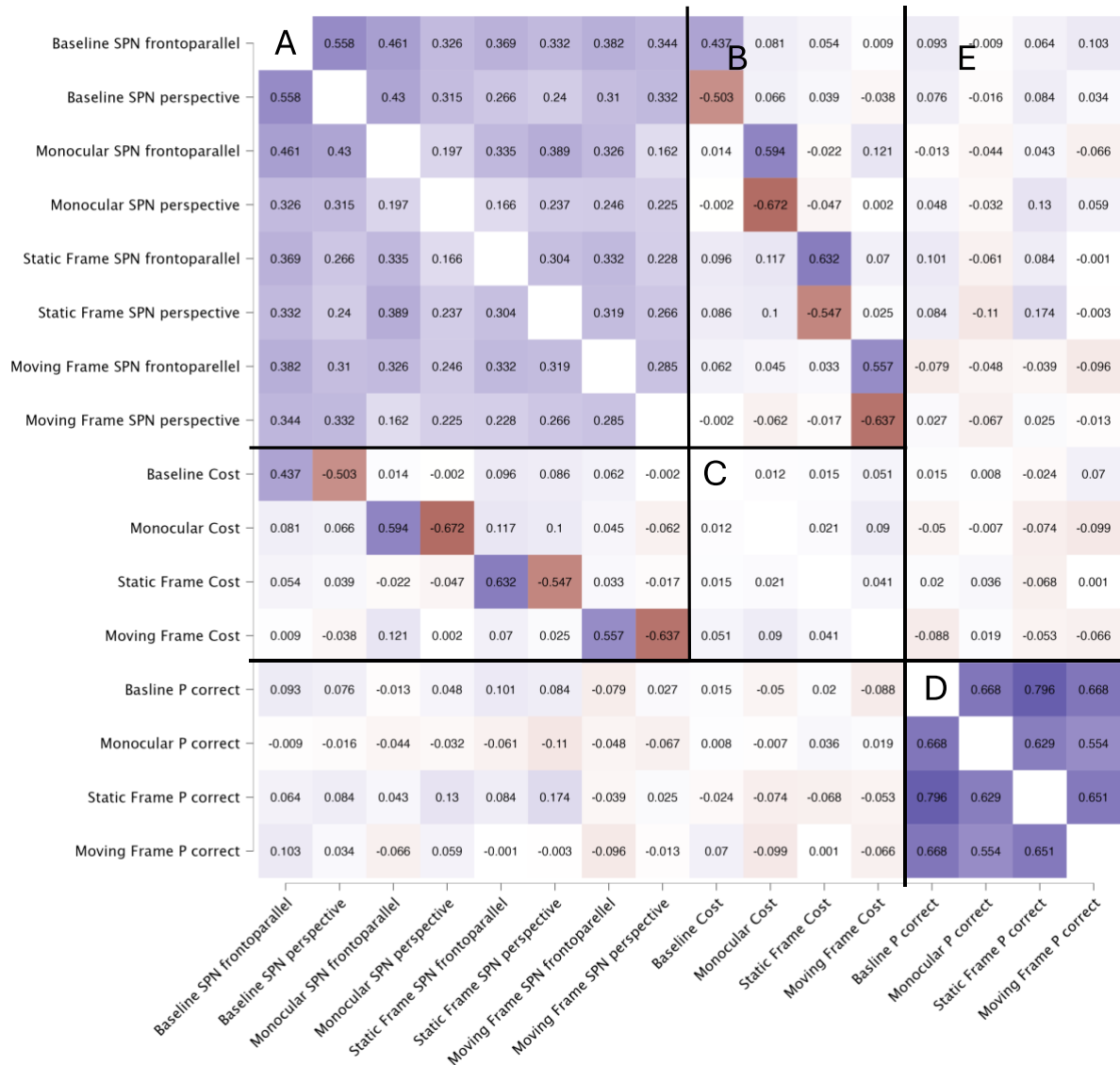


Figure 12. Correlation matrix of ERP and behavioural data. Blue cells indicate positive relationships, red indicate negative. Zones A, B, C, D and E are interpreted in text.

This analysis is also informative in other ways, because it shows the stability of SPN signals across conditions.

**2B. Whether the introduction, rationale and stated hypotheses (where applicable) are the same as the approved Stage 1 submission.** Aside from changes in tense (e.g. future tense to past tense), correction of typographic and grammatical errors, and correction of clear factual errors, the introduction, rationale and hypotheses of the Stage 2 submission seem to have remained identical to those in the approved Stage 1 manuscript. The one change we noted which would justify a footnote is the change in terminology when stating how the SPN will be calculated on page 7 under Study Aims and Hypotheses (and later in the methods). You previously stated the SPN would be calculated as the difference between the symmetry and random waves in the IPA Stage 1, however you now state the SPN will be calculated as the difference between thy symmetry and asymmetry waves in the Stage 2. It seems that the stimuli have not been changed from the Stage 1 to the Stage 2, therefore this is likely a terminology change. Please could you add a clarifying footnote accordingly.

We have added a clarifying change as footnote 1.

**2C. Whether the authors adhered precisely to the registered study procedures. This criterion assesses compliance with protocol.** Alongside the two issues mentioned in the above comments, the authors should also add a footnote to explain the edit on Page 13 under Apparatus which highlights the non-dominant eye will be covered for the monocular condition in the Stage 2. In the IPA, it was originally written that the left eye will be covered in the Stage 1. As you state under Participants on Page 9 of the IPA manuscript that the preferred sighted eye will be determined for the monocular viewing condition, this edit seems to be the correction of a mistake in the IPA. Please clarify this discrepancy in a footnote.

It was an error at this part of the Stage 1 manuscript as my non-dominant eye is the left one. We planned to test for eye preference and cover non-dominant eye rather than the left eye every time. This was already explained on page 10 in Stage 1. We have added this clarifying change as footnote 2.

**2D. Where applicable, whether any unregistered exploratory analyses are justified, methodologically sound, and informative.** In our opinion, the exploratory analyses seem justified and address questions relevant to the unexpected findings. However, the addition of the extra behavioral experiment to assess the perceptual experience of the participants is beyond the scope of exploratory analyses. At most we suggest that you move this experiment to an appendix, and ensure you only reference it briefly in the discussion with the main conclusion following from the approved registered study.

While the perceptual experiences check is something more than exploratory analysis, it is something less than an extra behavioural experiment. It was merely two questions

administered to 8 participants after a very short block. Furthermore, it facilitates interpreting the results of the approved pre-registered study. We would like to keep this in the manuscript; however, we have changed the language, so it is not referred to as a 'behavioural experiment'. However, if this is against the guidelines we are happy to move it to supplementary material.

**2E. Whether the authors' conclusions are justified given the evidence.** Please see the comment above.