Results of noveltyVR

Joern Alexander Quent

21/06/2021

# Participants

A number of participants had to be replaced for the following reasons: One participant was replaced because of a coding error. Two participants did not complete the recognition task at all and another participant did not complete the recognition task in time (i.e. 40 hours after encoding). An additional participant’s Pr was at chance level as defined the bootstrapping procedure.

Furthermore for one participant, the recall data was lost and could not be retrieved. For another participant, the response on the VR questionnaire were not saved. Due to human error, we collected 2 people more than necessary.

The final sample size was therefore 72 participants (53 females, 18 males, 1 non-binary) with mean age M = 26.3 years (SD = 6.25 years). Participants in the novelty group completed the online recognition task M = 25.5 hours (SD = 2.38 hours) after encoding, while the control group completed the task after M = 25.1 hours (SD = 2.17 hours).

# Results

Note in the text we report original accuracy rates and probability estimates, while the statistical tests and effect size estimates we use arcsine transformed values.

## Encoding task

In the encoding task, two participants did not use the correct keys so that these had to be exclude from the analysis. In this task, responses were faster in the animacy condition, M = 1000 msec (SD = 227 msec), compared to the alphabetical condition, M = 1540 msec (SD = 351 msec), = 2.51e+21, *d* = 1.88. However, there was insufficient evidence whether accuracy in the animacy condition, M = 0.884 (SD = 0.0864), was different from accuracy in the alphabetical condition, M = 0.897 (SD = 0.101), = 2.51, *d* = 0.183.

## Memory tasks

Despite not explicitly registered in Stage 1, we also report the level of processing effect as it was the central manipulation with two-tailed tests. For recall, participants recalled more words in the animacy condition, M = 2.77 (SD = 2.85), than in the alphabetical condition, M = 15.4 (SD = 6.38), = 1.28e+22, *d* = 1.91. Similarly for the recognition, memory performance in the animacy condition, M = 0.107 (SD = 0.074), was better than in the alphabetical condition, M = 0.271 (SD = 0.124), = 2.49e+18, *d* = 1.6. We could therefore conclude with confidence that our level of processing manipulation was successful.

In the following, we report the results of registered statistical tests of the influence of novelty on memory. For Hypothesis 1.1, we found that participants in the novelty group, M = 16.8 (SD = 7.02) did not recall more words than participants in the control group, M = 19.4 (SD = 7.59) with strong evidence against the null hypothesis, = 9.2, *d* = 0.35. Based on this we stopped data collection. In a two-tailed test (not registered), there was next to no evidence that the control group had better memory than the novelty group, = 0.618. For Hypothesis 1.2, we found some evidence that the novelty group, M = 0.186 (SD = 0.088), did not have better memory than the control group, M = 0.192 (SD = 0.0838), = 5.22, *d* = 0.0818.

For recognition, I also analysed Pr-values.

For Hypothesis 1.2, we found some evidence that the novelty group, M = 0.196 (SD = 0.091), did not have better memory than the control group, M = 0.22 (SD = 0.106), = 7.72, *d* = 0.25.

For Hypothesis 2, we found the number of words recalled in the animacy condition minus the number of words recalled in the alphabetical condition does not reliably differ between the novelty group, M = 11.4 (SD = 7.71), and the control group, M = 13.7 (SD = 5.1), = 1.17, *d* = 0.356 (Hypothesis 2.1). The same is true for the same comparison between the novelty group,M = 0.147 (SD = 0.0896), and the control group, M = 0.182 (SD = 0.131), with averaged recognition estimates that also did not allow to draw any conclusion, = 0.826, *d* = 0.314 (Hypothesis 2.2).

For recognition, I also analysed Pr-values.

The same is true for the same comparison between the novelty group,M = 0.162 (SD = 0.108), and the control group, M = 0.205 (SD = 0.125), with Pr that also did not allow to draw any conclusion, = 1.29, *d* = 0.371 (Hypothesis 2.2).

Regarding Hypothesis 3 (two-tailed), we found that the difference between recognition parameters (r - f) between the novelty group, M = -0.0509 (SD = 0.152) , and the control group, M = -0.024 (SD = 0.164), was not conclusive = 0.385, *d* = 0.244.

To test for the interaction of Hypothesis 4, we first calculated the difference between the two encoding conditions (animacy & alphabetical) for each group and for each parameter (r & f). We then calculated the difference of the condition difference (animacy & alphabetical) by substracting f - r and subjected these difference scores for the novelty group,M = -0.0162 (SD = 0.195), and the control group, M = -0.0715 (SD = 0.186), to statistical test, = 0.486, *d* = 0.3.

Table X: A summary of the results of the four primary hypothesis

|  |  |  |
| --- | --- | --- |
| Hypothesis | BF10 | BF01 |
| 1.1 | 0.109 | 9.200 |
| 1.2 | 0.130 | 7.720 |
| 2.1 | 1.170 | 0.855 |
| 2.2 | 1.290 | 0.774 |
| 3 | 0.385 | 2.600 |
| 4 | 0.486 | 2.060 |

All in all, we have been unable to provide any evidence in favour of our hypotheses.

## Post VR questionnaire

Table X: Summary of the results comparing the ratings for each post VR question with two-tailed tests.

|  |  |  |
| --- | --- | --- |
| Question | BF10 | BF01 |
| In the computer generated world I had a sense of “being there”. | 0.277 | 3.61 |
| Somehow I felt that the virtual world surrounded me. | 0.260 | 3.85 |
| I felt like I was just perceiving pictures. | 0.246 | 4.07 |
| I did not feel present in the virtual space. | 0.245 | 4.08 |
| I had a sense of acting in the virtual space, rather than operating something from outside. | 0.251 | 3.98 |
| I felt present in the virtual space. | 0.275 | 3.64 |
| How aware were you of the real world surrounding while navigating in the virtual world? (i.e. sounds, room temperature, other people, etc.)? | 0.329 | 3.04 |
| I was not aware of my real environment. | 0.259 | 3.86 |
| I still paid attention to the real environment. | 0.275 | 3.64 |
| I was completely captivated by the virtual world. | 0.246 | 4.07 |
| How real did the virtual world seem to you? | 0.458 | 2.18 |
| How much did your experience in the virtual environment seem consistent with your real world experience? | 0.358 | 2.79 |
| How real did the virtual world seem to you? | 0.306 | 3.27 |
| The virtual world seemed more realistic than the real world. | 0.389 | 2.57 |
| This experience was novel. | 0.260 | 3.85 |
| This experience was exciting. | 0.709 | 1.41 |
| This experiences was uncomfortable. | 0.251 | 3.98 |

As can be seen in Table X, none of the questions showed a reliable group difference. The fact that participants in the novelty group,M = 76 (SD = 31.9) , did not rate their experience as more novel as the participants in the control group, M = 72.6 (SD = 44.7) , = 0.26, *d* = 0.0884 is noteworthy here.

As an additional non-registered analysis, we examined the data from post VR questionnaire that we collected. For this, the data was rescaled to vary from 0 to 6 like the original scale (see Schubert et al., 2001) and items that were inversely scaled reversed. The IPQ score was then calculated by summing across all items. This analysis showed that the novelty group, M = 51.8 (SD = 10.9) , did not differ from the control group, M = 52 (SD = 10.1), *d* = 0.0238, BF10 = 0.246. In addition, we asked participants to rate the questions: “This experience was novel”, “This experience was exciting” and “This experience was uncomfortable”, however group differences did not arise for any of these statements. Especially for statement 1, the novelty group rated their experience, M = 5.28 (SD = 0.957), and the control group rated theirs, M = 5.18 (SD = 1.34), *d* = 0.049, BF10 = 0.249.