

ADD TITLE HERE

Sebastian Speer¹, Antonio Schettino^{2,3}, & Ana Martinovici²

¹ ADD SEBASTIAN'S AFFILIATION

² Erasmus University Rotterdam

³ Institute for Globally Distributed Open Research and Education (IGDORE)

Author Note

The authors made the following contributions. Sebastian Speer: Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing; Antonio Schettino: Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing; Ana Martinovici: Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Correspondence concerning this article should be addressed to Ana Martinovici, Burgemeester Oudlaan 50, 3062 PA Rotterdam, Netherlands. E-mail: martinovici@rsm.nl

Abstract

One or two sentences providing a **basic introduction** to the field, comprehensible to a scientist in any discipline.

Two to three sentences of **more detailed background**, comprehensible to scientists in related disciplines.

One sentence clearly stating the **general problem** being addressed by this particular study.

One sentence summarizing the main result (with the words “**here we show**” or their equivalent).

Two or three sentences explaining what the **main result** reveals in direct comparison to what was thought to be the case previously, or how the main result adds to previous knowledge.

One or two sentences to put the results into a more **general context**.

Two or three sentences to provide a **broader perspective**, readily comprehensible to a scientist in any discipline.

Keywords: EEG Many Pipelines, scene categorization, vision, EEG, ERP, time-frequency analysis, Bayesian multilevel linear regression, TFCE

Word count: X

ADD TITLE HERE

1 Introduction

ADD INTRODUCTION HERE

2 Methods

2.1 Preprocessing

All preprocessing was done using MNE-Python software (Gramfort et al., 2013). EEG data were filtered with a low cutoff filter of 1 Hz to remove slow drifts, a high cutoff filter at 40 Hz to attenuate high frequency power. Subsequently, bad and noisy channels were detected using several different approaches as implemented in the PREP pipeline (Bigdely-Shamlo, Mullen, Kothe, Su, & Robbins, 2015). First, by means of correlation, we checked how well a given channel is correlated with all other channels (categorized as bad at $r > 0.4$); second, we checked by using the robust z -score deviation aggregates per channel (categorized as bad at $z > 5$); third, by using the robust z -score estimates of high-frequency noise per channel (categorized as bad at $z > 5$); and finally, we checked by using the random sample consensus (RANSAC) channel correlations, which is the correlation for each channel with itself across the original data versus the RANSAC predicted data (categorized as bad at $r > 0.75$) as implemented in the PREP pipeline (Bigdely-Shamlo et al., 2015). After detection, these channels were removed from the data and subsequently interpolated (i.e., estimated from surrounding channels). Interpolation was performed using the spherical spline method (Perrin, Pernier, Bertrand, & Echallier, 1989) as implemented in MNE-Python, which projects the sensor locations onto a unit sphere and interpolates the signal at the channels identified as bad on the signals for the good channels. Afterwards, the EEG data were re-referenced to the average signal across channels. Next, ocular artifacts were removed by performing an independent component

analysis on the data and then correlating the resulting components with the EOG channels to see which of the components represented the ocular artifacts. The component that correlated the highest with the EOG channels was then removed from the EEG data.

2.2 Epoching

The EEG data was then subsampled by a factor of four (i.e., 128 Hz; for more info, see <https://mne.tools/stable/overview/faq.html#resampling-and-decimating>) segmented into 800 ms epochs, with 300 ms before stimulus onset (onset of the scene) until 500 ms after stimulus onset. The epochs were baseline corrected using the 300 ms preceding the stimulus onset. The resulting epochs were then subjected to *Autoreject*, an automated artifact detection algorithm based on machine-learning classifiers, and cross-validation to estimate the optimal peak-to-peak threshold (Jas, Engemann, Bekhti, Raimondo, & Gramfort, 2017). This algorithm was implemented to remove artifacts not identified by previous preprocessing steps, and depending on the number of bad sensors for a given trial, either repairs the trial based on interpolation or excludes it from further analysis. The preprocessed data were then submitted to a Morlet wavelet analysis to transform the data into the time-frequency domain with 18 log-scaled frequency bins ranging from 4 to 40 Hz to have higher sensitivity in lower frequency ranges such as the theta band. To optimize both spectral and temporal resolution, the number of cycles to include in the sliding time window were defined by dividing each individual frequency by two. After transforming the data to the time-frequency domain, the data were decimated by a factor of two (sampling every second time point) to increase computational efficiency.

3 Analysis

3.1 RQ2

To test whether there are effects of image novelty (RQ2; i.e., between images shown for the first time/new vs. repeated/old images) we conducted a multilevel analysis contrasting the EEG data from trials with old images against trials with new images. To test for differences in theta power at fronto-central channels we focused on the frequency range from 4-8 Hz and all frontocentral channels (FC1, FCz, FC2). At the first level (i.e., the participant level), we computed the averaged time-frequency maps for each of the two conditions. We then tested the resulting averaged maps at the second level for significant group effects, using a paired-sample t -test. We used cluster-based permutation testing as a stringent control for multiple comparisons (Maris & Oostenveld, 2007). Specifically, for every sample across the three channels, we quantified the experimental effect by a t value. Selection of samples for inclusion in a cluster was implemented using threshold-free cluster enhancement (TFCE) (Smith & Nichols, 2009). TFCE eliminates the free parameter initial threshold value that determines which points are included in clustering by approximating a continuous integration across possible threshold values with a standard Riemann sum. We subsequently clustered selected samples in connected sets based on temporal and spectral adjacency, and we computed cluster-level statistics by taking the sum of the t values within every cluster. Subsequently, we performed permutation testing using the Monte Carlo method to compute the posterior significance probability of our observed effect (Maris & Oostenveld, 2007). This analysis results in a cluster of adjacent data points across time, frequencies, and channels, which significantly differs in activity between old and new images. To test for differences in alpha power at posterior channels we focused on the frequency range from 8-13 Hz and all posterior channels (P7, P5, P3, P1, P2, P4, P6).

3.2 RQ3

The same analysis approach as described for RQ2 was implemented for RQ3. Specifically, to test whether there are effects of successful recognition of old images on spectral power, at any frequencies, at any channel, at any time, we contrasted time-frequency decomposed EEG data from trials containing old images that were correctly recognized as old with old images incorrectly recognized as new. Here we included all frequencies, timepoints, and channels in the analysis. The same thresholding procedure and permutation testing approach as above was used.

3.3 RQ4

To test whether there are effects of subsequent memory, we conducted exactly the same analysis as described in RQ3, with the only difference that here we contrasted trial containing images that will be successfully remembered vs. forgotten on a subsequent repetition.

4 Results

No significant clusters were identified for any of the research questions (at $\alpha = 0.05$).

4.1 Software

We used R (Version 4.1.3; R Core Team, 2022) and the R-packages *citr* (Version 0.3.2; Aust, 2019), *papaja* (Version 0.1.0.9999; Aust & Barth, 2022), *tinylabels* (Version 0.2.3; Barth, 2022), and *tinytex* (Version 0.38; Xie, 2019) for all our analyses.

5 Discussion

ADD DISCUSSION HERE

6 References

- Aust, F. (2019). *Citr: 'RStudio' add-in to insert markdown citations*. Retrieved from <https://github.com/crsh/citr>
- Aust, F., & Barth, M. (2022). *papaja: Prepare reproducible APA journal articles with R Markdown*. Retrieved from <https://github.com/crsh/papaja>
- Barth, M. (2022). *tinylabls: Lightweight variable labels*. Retrieved from <https://cran.r-project.org/package=tinylabls>
- Bigdely-Shamlo, N., Mullen, T., Kothe, C., Su, K.-M., & Robbins, K. A. (2015). The PREP pipeline: Standardized preprocessing for large-scale EEG analysis. *Frontiers in Neuroinformatics*, 9, 16. <https://doi.org/10.3389/fninf.2015.00016>
- Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., ... Hämäläinen, M. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, 7, 267. <https://doi.org/10.3389/fnins.2013.00267>
- Jas, M., Engemann, D. A., Bekhti, Y., Raimondo, F., & Gramfort, A. (2017). Autoreject: Automated artifact rejection for MEG and EEG data. *NeuroImage*, 159, 417–429. <https://doi.org/10.1016/j.neuroimage.2017.06.030>
- Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, 164(1), 177–190. <https://doi.org/10.1016/j.jneumeth.2007.03.024>
- Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. F. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalography and Clinical Neurophysiology*, 72(2), 184–187. [https://doi.org/10.1016/0013-4694\(89\)90180-6](https://doi.org/10.1016/0013-4694(89)90180-6)
- R Core Team. (2022). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Smith, S. M., & Nichols, T. E. (2009). Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference.

- 154 *NeuroImage*, 44(1), 83–98. <https://doi.org/10.1016/j.neuroimage.2008.03.061>
- 155 Xie, Y. (2019). TinyTeX: A lightweight, cross-platform, and easy-to-maintain LaTeX
- 156 distribution based on TeX live. *TUGboat*, (1), 30–32. Retrieved from
- 157 <https://tug.org/TUGboat/Contents/contents40-1.html>