

Study Registration

For the Koestler Parapsychology Unit Study Registry

The registration information for the study is given below. Each section can be expanded as needed.

1. The title or name of the experiment (for listing the experiment in the registry).

Does characteristic alpha EEG activity predict upcoming sensory events?

2. The name, affiliation, and email address for the lead experimenter(s) for the study.

Julia Mossbridge, Northwestern University, j-mossbridge@northwestern.edu

3. A short description or abstract of the purpose and design of the experiment.

The first experiment was already performed before I had learned of the existence of this registry. However, I have yet to analyze the data from the second (replication) experiment, so I am registering this upcoming analysis. The purpose of the experiment is to determine whether the phase of characteristic alpha EEG activity predicts whether an upcoming sensory event will have a given motor response. The sensory event dictates the motor response, in that participants are told that they should press the left mouse key as quickly as possible if they see a number "1" on the screen or hear a low tone, and the right mouse key as quickly as possible if they see a number "2" on the screen or hear a high tone. Thus it is not clear whether it is the motor response or the sensory event that is being predicted; that will have to wait for future experiments, as the error rate is too small to determine the answer with this paradigm.

All responses were made with the right hand. The analysis of the first experiment revealed a significant phase difference in characteristic alpha activity preceding the presentation of the upcoming stimulus, about 550 ms pre-stimulus (similar to Ben Libet's 1983 result). Pattern classification showed that alpha phase at three left parietal electrodes significantly predicted the upcoming left- or right-button press (and thus the upcoming stimulus type).

4. A statement or list of the specific hypothesis or hypotheses being tested, and whether each hypothesis is confirmatory or exploratory.

I am registering the analysis of a replication data set. Participants performed exactly the same experiment, and I will run the same analysis to determine if I get the same result as the initial, exploratory experiment.

The hypothesis is that the phases of the peak EEG alpha frequency (relative to 1900 ms pre-stimulus) at left parietal electrodes differentiate upcoming stimuli into their behaviorally relevant groups. Note that here the dependent variable is the upcoming stimulus and the independent variables are the alpha phases at each of 64 electrodes. This registration is for the analysis of a second, confirmatory experiment.

5. The planned number of participants and the number of trials per participant.

I've already run the replication, I just need to analyze the data. There are the same number of participants (20) as in the original study. Each participant performed 20 fewer trials than in the original study (120 trials in the original study, 100 trials in the replication) because it appeared that the effect was strong enough not to require the additional trials.

6. A statement that the registration is submitted prior to testing the first participant, or indicating the number of participants tested when the registration (or revision to the registration) was submitted.

All participants have been tested. The second replication has yet to be analyzed, so I am registering prior to performing that analysis.

In addition to the minimum content above, further information is highly recommended—particularly for well-planned confirmatory experiments. The additional information includes some or all of the following:

7. The specific statistical test method that is planned for each hypothesis, including which statistical test will be used, whether the unit of analysis is the participant or the individual random event, what p value (or confidence interval level) is significant, whether the statistical test (or confidence interval) is one or two-tailed, and any adjustment for multiple analyses. For example, “to analyze overall ψ , a z-score binomial test with continuity correction will evaluate whether the overall rate of direct hits for all trials in the experiment is greater than 25%, with significance set at $p \leq .05$ one-tailed,” or “the difference between the two conditions will be analyzed with a two-sample t -test with the number of hits for

each participant as the unit of analysis and significance set at $p \leq .05$ two-tailed.” (This information can be included in section 4 above.)

The average phases of the peak EEG alpha frequency (relative to 1900 ms pre-stimulus) at each of 64 electrodes are processed with a random forest pattern classification algorithm to develop classification criteria for the type of upcoming response (right vs. left). Each participant's mean phase data at each electrode are entered as one instance for the classifier; thus for 20 people there are 20 instances representing pre-right phases and 20 instances representing pre-left phases. Random forest classification algorithms have a built-in generalization test (out-of-bag error estimation), so only the error rate based on attempts at classifying the upcoming response using the instances that are not used during the training phase are reported as the final error rate. This method is used to determine the critical electrodes for classifying upcoming responses (based on generalization error).

There are several ways to test these data for significance. We use two methods, reporting the results of both:

- 1) Finding two distributions of classification error rate across 1000 attempts at classification for each of two data sets, and performing a distribution tail-test between the distributions of error rates. One distribution consists of error rates based on the original (real) data. The other distribution consists of error rates based on a scrambled version of the original (real) data. There are many methods of scrambling these data, and each gives different results. We are trying to determine which is the best compromise between these methods (not too conservative, not too liberal). Regardless, we will use the same scrambling method on both data sets. We think we have settled on a scrambling method that places 50% of one trial type in the other category, but without replacement. Here, the critical alpha for the two-tailed distribution test is 0.05.
- 2) Performing a circular t-test on the phase data (within-participant means for upcoming-left vs. upcoming-right stimuli) at the most critical electrode for classification. The adjustment here for multiple analyses puts the critical alpha value for the two-tailed t-test at 0.00078 ($.05/64$ electrodes). Note that here, the phase data are the DVs and the upcoming stimuli are the IVs.

8. The power analysis or other justification for the number of participants and trials.

To perform a power analysis, we'd need to come to an agreement on how to calculate effect size for the classification analyses. It's not straightforward, because the effect size depends on the difference between the error rate on the pattern classified original data and the error rate on scrambled data -- and there are multiple scrambling methods which create different effect sizes. However, using what we believe to be a middling method (in terms of being conservative -- not too conservative, not too liberal), the effect size (Cohen's d) is 3.58, making 20 participants probably overkill. However, we note that we reserve the right to change the scrambling method if our simulations discover an error in our scrambling method that produces such big effect sizes (obviously we're a bit

concerned about that). If we do change the scrambling method in the final paper, the effect size will be different than the one given here.

9. The methods for randomization in the experiment.

We used a pseudo-random number generator, the one built into the Presentation software package. Because it wasn't a hardware RNG, we are also examining possible transition statistics issues that could potentially influence the results.

10. A detailed description of the experimental procedure.

Briefly, 20 undergraduates were fitted with 64 active EEG electrodes. Participants used their right hand to press the left mouse button as soon as they saw a "1" or heard a low tone, and the right button if they saw a "2" or heard a high tone. All stimuli were randomized using a pseudo random number generator. Raw EEG was current-source density transformed and artifacts were removed before analysis. The dependent variables were the phases of the mean (across-trial) peak alpha (7.5-12 Hz) frequency for each person at each electrode relative to 1900 ms preceding the onset of the upcoming stimulus presentation (left vs. right button press). A random forest pattern classification algorithm was used to classify the dependent variables according to the nature of the upcoming stimulus presentation. We have successfully used this algorithm in the past and have found it to be capable of processing noisy data. To be statistically conservative, the classifier was executed using the actual data versus a randomized version of the same data. Follow-up circular t-tests on the phase angles at the 64 electrodes confirmed the classification.