Supplementary materials:

1/5 – preRNG further aspects

preRNG: further aspects

1 Immunity to hacking strategies

1.1 changing the analysis protocol in retrospect

Any changes made to the protocol folder after data acquisition will change the randomization scheme completely and will thus break the alignment of the data with the experimental protocol. Such misalignment will expose the author to the risk of being detected by readers who wish to validate the pre-registration.

1.2 multiple study protocols for the same data set

Since all random components in the experiment are determined by the same PRNG and using the same seed, only one protocol folder can be associated with a particular study.

1.3 reporting only successful repetitions

To date, no pre-registration scheme is immune to selective reporting of subjects or experimental repetitions. In theory, authors can repeat the same experiment over and over until observing the desired effect, and report only the last, successful, repetition. Similarly, in the case of multi-subject experiments, authors can run as many subjects as they wish, only to report a subset of these subjects whose data aligns with the prior hypothesis. Our scheme is not immune to such misconduct. However, this concern can be alleviated by introducing a dependency between subsequent subjects. One way to introduce such a dependency is to call the preRNG function with the experimental data of subject $n-1$ when determining random aspects of the experimental design of subject n. This way, the data acquired from subject n is dependent upon the data acquired from all the previous subjects and the original protocol folder, making it impossible to report a subset of the subjects without breaking the alignment between the data and the experimental randomization. This solution is not complete, as it does not prevent the researcher from exploring multiple candidates for subject n before deciding on the best one and moving on to the next subject. This can be mitigated by committing to a public identifier of subject $n+1$, where such an identifier is available, in addition to the data of subject $n-1$.

2 Studies involving multiple randomizations

In cases where multiple randomizations are needed for different repetitions of the same experiment, such as in the case of multiple subjects, the preRNG function can be called with an optional argument specifying the subject's serial number. The serial number will be appended to the protocol sum, and the SHA256 function will be applied to the resulting string. The new sum will be used to initialize the PRNG. This guarantees that (a) different repetitions will be initialized with completely different seeds and that (b) for all subjects, the randomization is fully dependent on the protocol folder. This option is supported by the accompanying implementations.

3 Studies with insufficient randomization

Some experimental designs do not include any random component, and in others the entropy of the experimental randomization is not sufficient to effectively time-lock the protocol folder. To use the preRNG scheme in such cases, one can add an additional randomized experimental phase only for the purpose of time-locking. For example, neuroimaging experiments can begin with a short block of events that give rise to robust sensory or motor activations, in random order and timing.

4 Constrained randomization

preRNG can be used even in the presence of constraints to the randomization scheme, as long as the min-entropy of the experimental randomization remains sufficient.

5 Comparing alternative randomization schemes

To verify the dependence of the data on the PRNG initialization, the verifier can generate a null distribution of results assuming different randomizations that were obtained using arbitrary seeds, while keeping the data constant. For example, the verifier can derive the result of a contrast between two experimental conditions assuming the random order of events that is dictated by the original protocol folder, and compare it to the result of the same contrast when assuming other possible orders generated by initializing the PRNG with alternative seeds. This is practically similar but conceptually different from the use of permutation testing for nonparametric inference, as here the effect of interest is assumed to be known, and inference is made on the true experimental randomization that was used to generate the data.

Importantly, in cases where (i) the entropy of the experimental randomization is high enough and (ii) the measured variable is noisy, it is probable that alternative randomization schemes will be more likely given the data than the one actually used. In such cases the similarity between the randomization schemes can support the validity of the pre-registration process. For example, in the case of an experiment consisting of 200 events of two conditions (100 of each) in random order, the chances that the maximum likelihood estimate of the randomization scheme will overlap by more than 180 events by chance alone is very small ($< 2^{-110}$). Therefore, evidence for that the protocol folder gives rise to randomization that is sufficiently similar to the one that maximizes the effect of interest (or the likelihood of the data) can be used to corroborate the validity of the registration process.

2/5: Alice's paper:

An example report of a neuroimaging study incorporating the preRNG scheme, as described in the main manuscript.

Cerebellum Involvement in Hand Movements: a Functional MRI Study

Alice

***Abstract*— To investigate the contribution of the cerebellum to motor behavior, I used a simple motor task inside an MRI scanner.**

I. Introduction

The cerebellum plays a critical role in behavior and motor control (Fox et al., 1985; Seitz et al., 1990; Sabatini et al., 1993; Ellerman et al., 1994). Here I provide further evidence for cerebellar contribution to simple hand movements (sequential fist opening and closing), using functional Magnetic Resonance Imaging (fMRI). I hypothesized that blood-oxygenation level dependent (BOLD) signal will increase for the cerebellar hemisphere ipsilateral to the active hand.

II. Materials and Methods

Unless otherwise stated, all materials and methods were determined prior to data collection, and were time-locked using the preRNG procedure (Mazor et al., 2017). The protocol folder is available on github.com/alicemdphd/cerebellum

A. Experiment

1) Participants: One subject (26 year old, right-handed female) participated in the experiment.

2) Procedure: The subject laid supine on the scanner bed, and viewed visual stimuli back-projected onto a screen through a mirror. Foam pads were used to minimize head motion. Stimulus presentation and timing of all stimuli were achieved using Python and PsychoPy (Peirce, 2008). The subject's eye movements were monitored using an EyeLink 1000 Plus eye-tracker. The experiment consisted of four 6 minutes experimental runs. The task was presented in a block design (8 seconds task blocks alternated with 10 seconds of rest, 20 blocks per run). During experimental blocks an arrow appeared on the screen, to which the subject was asked to react with sequential fist opening and closing of the appropriate hand (right or left, as indicated by the arrow direction), at her own pace. A fixation-cross appeared during rest periods. For each run, a right arrow was presented in 10 blocks, and a left arrow was presented in the other 10. The order of blocks was randomized within and between runs, using the preRNG procedure (Mazor et al., 2017) and based on the predetermined protocol folder (github.com/alicemdphd/cerebellum).

3) MRI Data Acquision: A Siemens 3-T Prisma scanner (located at the Edersheim-Levi Gitter Center for human brain imaging, Tel Aviv University, Israel) with a 64-channel Siemens Matrix head coil was used to collect all functional and anatomical scans. A single high-resolution structural scan was acquired using a magnetization-prepared rapid acquisition gradient echo (MP-RAGE) sequence (1 x 1 x 1 mm voxels). All functional runs were acquired parallel to the anterior-posterior commissure plane using the Center for Magnetic Resonance Research (CMRR) multiband accelerated gradient-echo EPI sequence (66 contiguous interleaved axial slices, 2 mm thickness, no gap; TR = 2000 msec; flip angle = 82; TE = 32.2 msec; in-plane resolution = 2 x 2 mm; matrix size = 96 x 96).

B. Data Analysis

1) Image Preprocessing and Statistical Analysis: The acquired data were analyzed using FEAT v6.00 (FMRI Expert Analysis Tool), part of FSL (FMRIB software library, version 5.0, www.fmrib.ox.ac.uk/fsl). Images were realigned to the central volume of each run to correct for head movements, and spatially smoothed using a 5 mm kernel. The data were then temporally filtered using both a high-pass filter with a cutoff of 50 seconds, and the FILM prewhitening tool. Functional images were registered to the brain-extracted T1 image, using boundary based registration. The anatomical image was registered to the standard MNI space (MNI152, 2mm) by first performing a linear registration with 12 degrees of freedom, and then using the FNIRT nonlinear registration tool with a warp resolution of 10 mm on the linearly registered image. First level analysis was executed using FILM. The model included 4 regressors: right-hand and left-hand blocks were modeled and convolved with a Double-Gamma HRF (Rh and Lh, accordingly). The temporal derivative of each of the resulting regressors was added to the design matrix as a second explanatory variable to account for minor temporal offsets. The design matrix then went through the same temporal filtering process as the empirical data, before beta values were extracted for each voxel in the brain by fitting the model to the voxel's time series. The four runs were modeled as a fixed effect. A GLM contrast between right-hand and left-hand regressors (Rh-Lh) was performed. I chose to restrict the analysis to the cerebellum, which was identified anatomically using the MNI atlas provided with FSL (MNI-maxprob-thr50-2mm). A small-volume false discovery rate (FDR) correction was applied to the voxels within this region, using the Benjamini-Hochberg procedure (Benjamini & Hochberg, 1995).

III. Results

In line with my prior hypothesis, I found stronger activation in the right cerebellar hemisphere for right-hand movements, and stronger activation in the left cerebellar hemisphere for left-hand movements (see Fig. 1). These

results survived FDR correction. Furthermore, exploration of the statistical parametric maps for this contrast revealed a positive linear modulation of the effect size in the left cerebellar hemisphere as a function of the run serial number. Results for this post hoc contrast survived FDR small volume correction, using the same anatomical mask that was used for the main contrast.

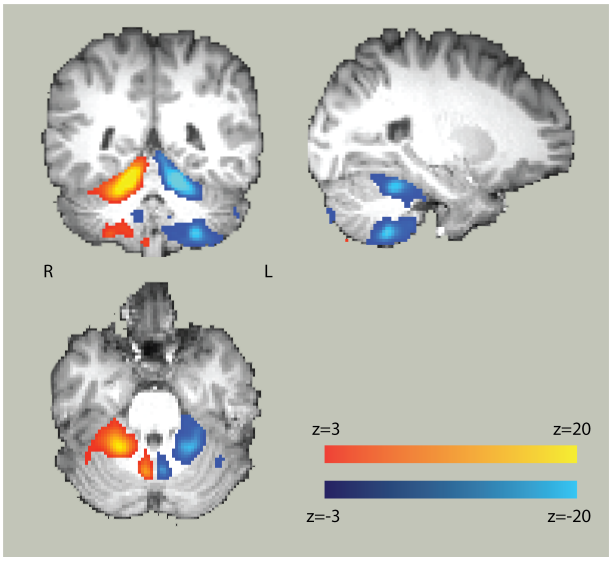

Fig. 1. Main contrast: Right-hand (red-yellow) - Left-hand (blue-light-blue). The contrast was restricted to the cerebellum only

IV. Discussion

The current study provides evidence for robust ipsilateral activation in the cerebellum for single-hand movements. These findings are in line with previous research reporting ipsilateral involvement of the cerebellum in motor processes (Fox et al., 1985; Sabatini et al., 1993; Ellerman et al., 1994).

References

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the royal statistical society. Series B (Methodological)*, 289–300.

Ellerman, J., Flament, D., Kim, S.-G., Fu, Q.-G., Merkle, H., Ebner, T., & Ugurbil, K. (1994). Spatial patterns of functional activation of the cerebellum investigated using high field (4 t) mri. *NMR in Biomedicine*, *7*(1-2), 63–68.

Fox, P. T., Raichle, M. E., & Thach, W. T. (1985). Functional mapping of the human cerebellum with positron emission tomography. *Proceedings of the National Academy of Sciences*, *82*(21), 7462–7466.

Mazor, M., Mazor, N., & Mukamel, R. (2017). Pre-registration: a cryptographic approach.

Peirce, J. W. (2008). Generating stimuli for neuroscience using psychopy. *Frontiers in neuroinformatics*, *2*.

Sabatini, U., Chollet, F., Rascol, O., Celsis, P., Rascol, A., Lenzi, G. L., & Marc-Vergnes, J.-P. (1993). Effect of side and rate of stimulation on cerebral blood flow changes in motor areas during finger movements in humans. *Journal of Cerebral Blood Flow & Metabolism*, *13*(4), 639–645.

Seitz, R. J., Roland, P. E., Bohm, C., Greitz, T., & Stone-Elander, S. (1990). Motor learning in man: a positron emission tomographic study. *Neuroreport*, *1*(1), 57–60.

3/5: Bob's verification:

Example verification steps of a neuroimaging study incorporating the preRNG scheme, as described in the main manuscript.

preRNG verification

Bob

Here I validated the pre-registration of "*Cerebellum Involvement in Hand Movements: a Functional MRI Study*" (Alice, 2017). Although Alice's paper is focused on the cerebellum, I decided to use primary cortical activations as a voucher for the pre-registration validity. I found these activations to be aligned with the randomization induced by the protocol folder, verifying the study's pre-registration.

1. I ran the function `createEVs` (that is found in the protocol folder) specifying the path to the compressed protocol folder as argument. This function calls the `preRNG.py` function, which initialized PRNG to the hashed protocol sum. It then generated a pseudorandom order of blocks for each of the four runs. This resulted in the following protocol sum and order of blocks:

```
>>> createEVs(os.path.join(home_dir,'protocolFolder.zip'))
'620d185c26237de2e54a0affaeb7a9d2b9ca5b0185ec1f1e90796fe6ff06152e'
```

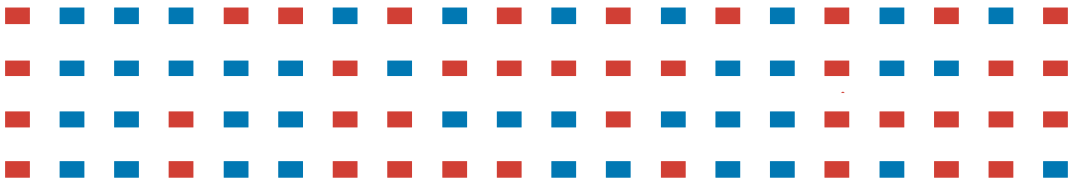

Figure 1: The resulting order of blocks in the four experimental runs. Blue rectangles represent blocks in which the subject was requested to move her right hand, and red rectangles represent blocks in which she was requested to move her left hand.

2. I performed a whole-brain contrast between right and left hand movements using the information I acquired about the block order in the previous step:

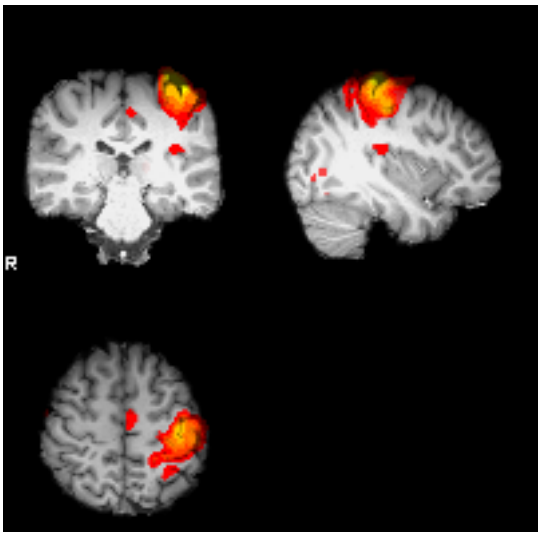

Figure 2: a contrast of right hand > left hand. Results are corrected for multiple comparisons, whole brain.

3. Introducing slight changes to the protocol folder (by adding a sequence of # symbols to the end of the analysis file found in the protocol folder) resulted in activations weaker by orders of magnitude:

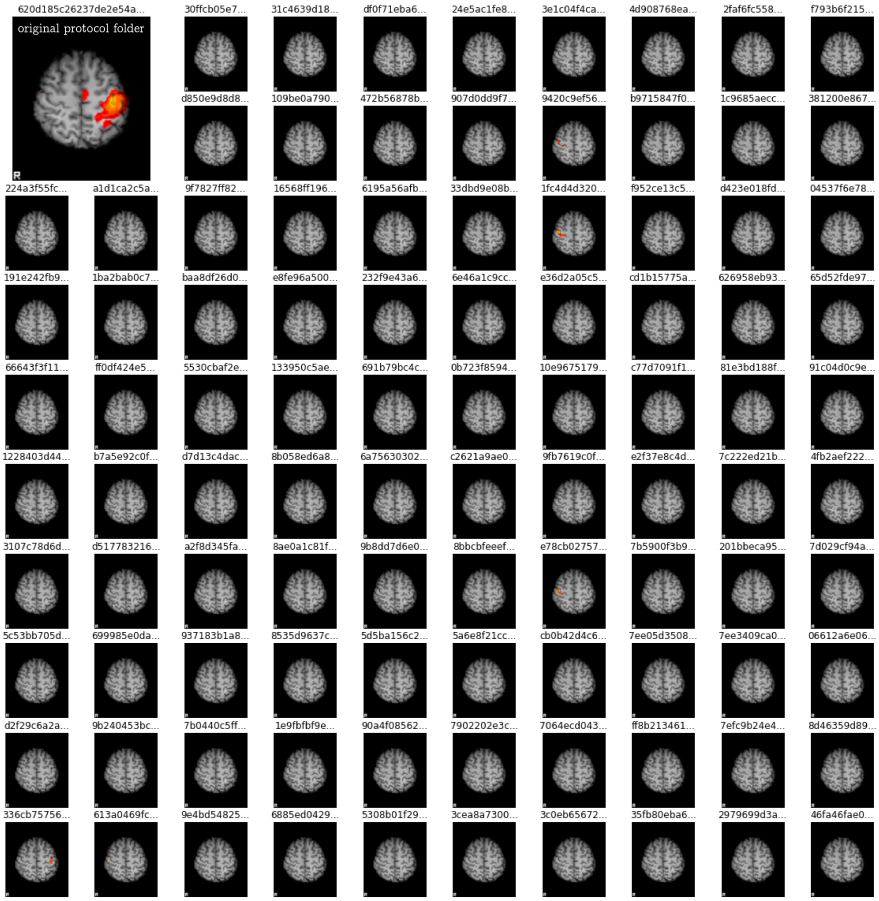

Figure 3: Appending #s to the file titled analyze.py and applying the preRNG function to the altered protocol folder dramatically changed the protocol-sum (first 10 letters appear above the image) and the presumed order of events. This slight change resulted in activation maps weaker by orders of magnitude compared to the one obtained from from the original protocol folder.

4. I extracted activation timecourses from two voxels in the right and left central sulci. These timecourses aligned well with the order of events as obtained from applying the preRNG function to Alice's original protocol folder:

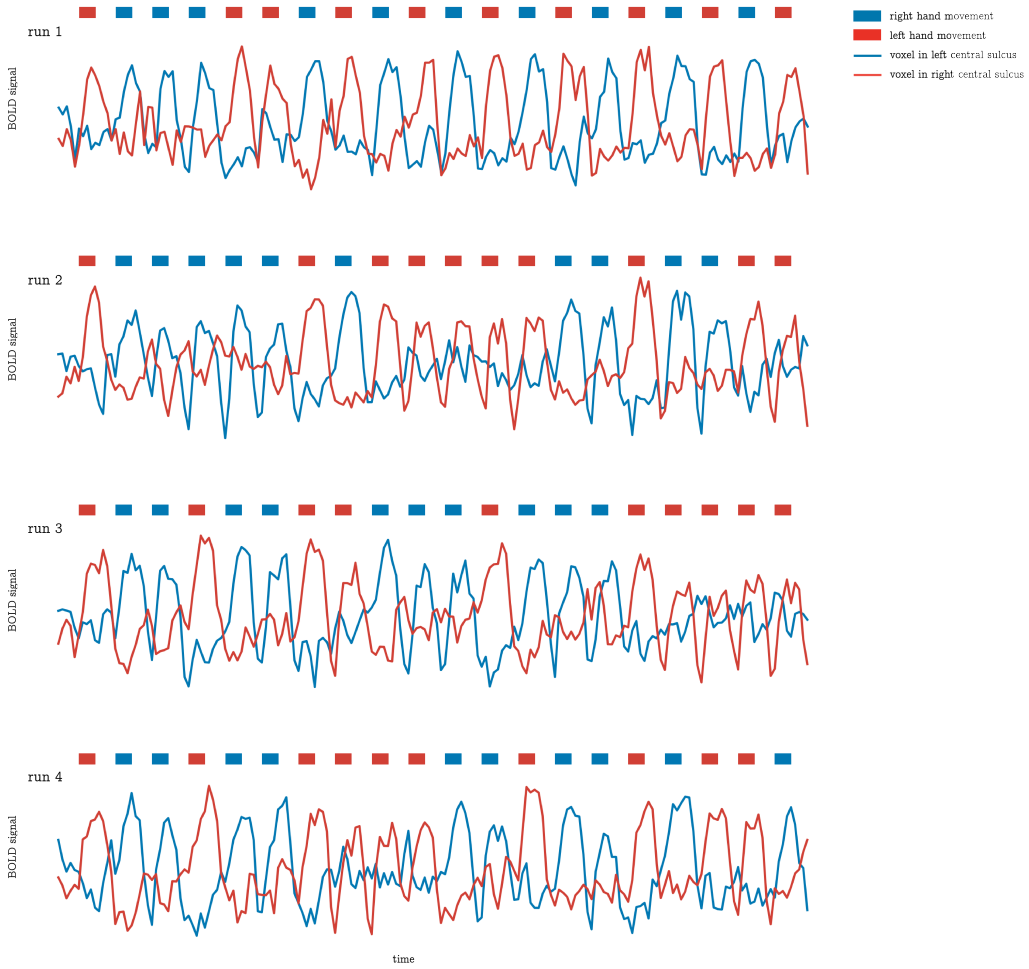

Figure 4: Activation timecourses for the four experimental runs, as extracted from the bilateral hand knobs in the central sulci. blue rectangles represent blocks in which the subject was requested to move her right hand, and red rectangles represent blocks in which she was requested to move her left hand. Blue and red lines correspond to signal from the left and right central sulci accordingly.

5. To further explore the alignment of the primary cortical activations with the presumed order of events, I computed the ratio between the likelihood of the voxels' timecourses given neighboring experimental designs (designs that are identical to the original, except for one swap of right and left experimental blocks; $Design'$) and the original experimental design ($Design$). I computed this Bayes factor ($\frac{p(S1activations|Design'}{p(S1activations|Design}$) for each neighboring experimental design, using a naive linear regression model (not accounting for autocorrelations). None of the Bayes factors exceeded 1, suggesting that the original experimental design maximized the likelihood of these data, at least locally.

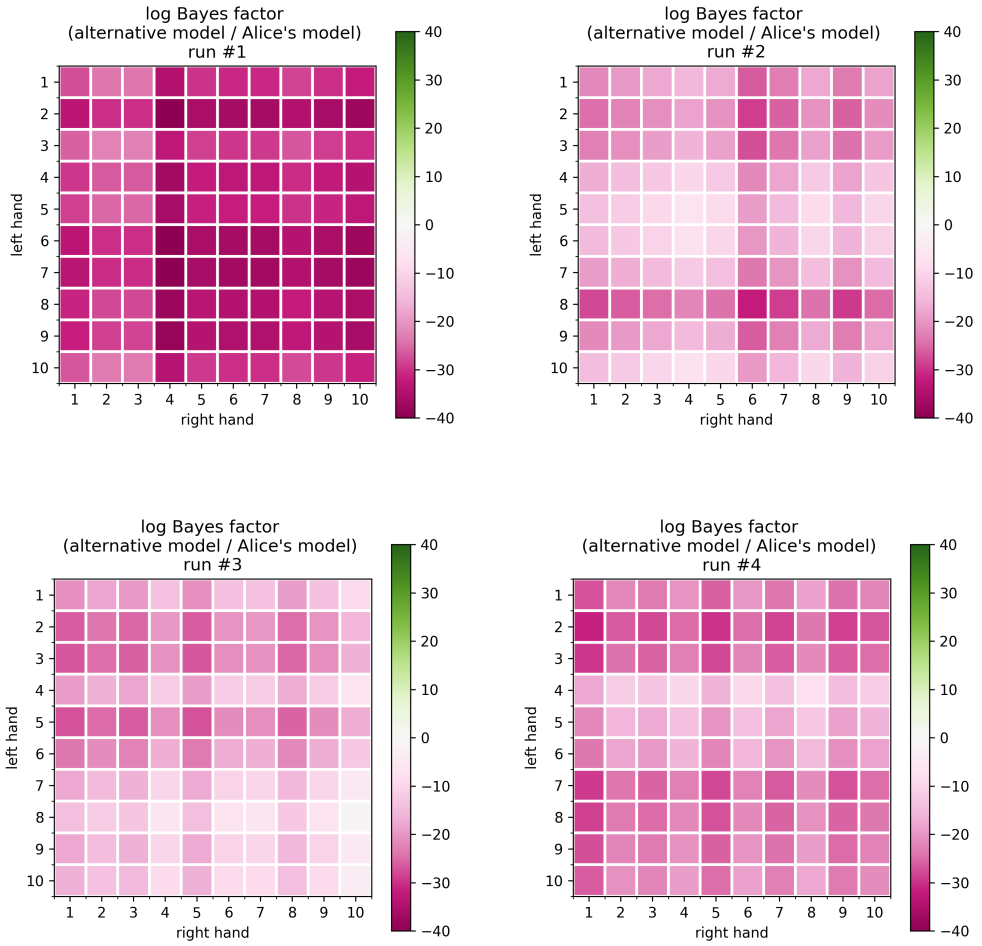

Figure 5: Matrix entries represent the log likelihood ratio between the neighboring design and the original order of events. For example, entry 9,4 in the first matrix compares the likelihood of the data given a design in which the ninth left hand event has been replaced with the fourth right hand event, to the likelihood of the data given the original experimental design. The log-likelihood is negative, suggesting that the alternative model is less appropriate. Note that none of the swaps across all four experimental runs increases the likelihood of the data.

6. Finally, to examine whether other event orders maximize the likelihood of the chosen voxels' activations, I performed gradient descent optimization starting from an arbitrary order of events. I then performed the following steps:

   (a) Generate the log Bayes factor matrix A by computing a Bayes factor for each possible swap of left and right events, as in step 5

   (b) Extract the maximum log Bayes factor $A_{i,j}$ from the matrix

   (c) if $A_{i,j} > 0$, swap the i^{th} left hand event with the j^{th} right hand event, and repeat steps a-c. otherwise, stop.

   For all four runs, the algorithm converged to the same order of events that was obtained by using the preRNG function with Alice's protocol folder, validating the alignment of this randomization with the raw experimental data.

Pre-RNG implementations: implementations of the preRNG scheme for Python, Matlab and R.

https://github.com/matanmazor/preRNG

Bob's verification code: The Python code used for the verification process as described in Bob's report.

https://github.com/matanmazor/AliceAndBob