Supplementary Material for: Socioeconomic status predicts hemispheric specialisation of the left inferior frontal gyrus in young children

Rajeev D. S. Raizada,*[1], Todd L. Richards[2], Andrew Meltzoff[1] and Patricia K. Kuhl[1]
[1] Inst. for Learning & Brain Sciences, Univ. of Washington, Box 357988, Seattle WA 98195
[2] Department of Radiology, Univ. of Washington Medical School, Seattle WA 98195

*To whom correspondence should be addressed; E-mail: raizada@u.washington.edu

Strengths and limitations of the FDR test

The FDR test shows that, considered with reference to the set of all the correlations between brain, behavioural and environmental variables, there is only a 5% chance of the SES/Broca's correlation being a false positive. As is the case for all statistical tests, the FDR test does not indicate whether or not that particular result *is* a false positive, nor does it show what the false-positive chances for that finding would be with a different reference set.

An FDR test is implemented by constructing a sorted ascending list of all the p-values in the set, and the FDR-threshold is the p-value at which this sorted list intersects with a line whose gradient is the desired corrected p-value threshold (usually 0.05). If there are a large number of other tests in the set that yield low p-values then it is easier for the sorted p-value list to "dip under" the 0.05-gradient line, so the presence of these low-p-value tests has the effect of making it easier for a different test in the set to be classed as significant. Conversely, a single very low p-value in the set may fail to pass an FDR test if it is not accompanied by a large number of other low-p-value tests, as without these other low p-values the isolated strong result will fail to "dip under" the 0.05-gradient line.

The data presented here actually exhibits both of these phenomena, depending on how it is subdivided. When all of the brain, behavioural and environmental correlations are considered together, as in the correlation matrix shown in Figure 1 of the main text, the existence of many strong brain-with-brain and behaviour-with-behaviour correlations does indeed make it easier for the environment-with-brain SES/IFG correlation to pass the FDR significance correction. However, when only the correlations between behavioural/environmental variables on the one hand and brain variables on the other hand are considered, the SES/IFG correlation stands apart as especially strong.

This is illustrated in the histogram of these correlation values shown in Supplementary Figure 6. Two correlations stand apart as being stronger than the others: the correlation between SES and all three parts of the IFG, and the correlation between SES and the opercular and triangular IFG-subparts together. The fact that these SES/IFG correlations distinguish themselves as being stronger by some margin than all the other correlations in this set supports the interpretation of them as meaningfully significant. However, as remarked above, a low p-value does not pass an FDR test by standing out from the crowd; on the contrary, it would need to be part of a crowd. This is illustrated in Supplementary Figure 7: the lowest p-value, which is that of the SES/IFG correlation, does not pass an FDR test of this reduced part of the dataset, precisely because FDR can capture the significance of individual tests only when they are accompanied by many other tests with equally low p-values. Thus, FDR is ill-suited for capturing the structure of this particular subpart of the dataset.

In summary, whether or not a given correlation passes an FDR test must be understood in the context of the reference set of other correlations that the test includes. Considering only the reference set of behavioural/environmental-with-brain correlations, the SES/IFG correlation stands out as especially strong, but not in a way that FDR can capture. Considering the global reference set of all the correlations between brain, behavioural and environmental variables, FDR does highlight the SES/IFG correlation as significant. The closer examination of the distribution of p-values provided here allows those results to be better understood.

The partial-correlation tests shown in Figure 3b further corroborate the fact that the SES/Broca's correlation stands apart as stronger than other correlations in the broader data set. To carry out these partial correlation analyses, we selected all the standardised test scores that correlated at $p<0.05$ (uncorrected) with either or both of the SES and Broca's measures. These were: WPPSI Verbal IQ, WPPSI General Language Score, PAT phoneme deletion, and the Peabody Picture Vocabulary Test. As is described in the Results section above, the SES/Broca's correlation remained significant at $p<0.05$ when any of these four measures were partialled out, and even when all four were partialled out simultaneously. Thus, the SES/Broca's correlation is strong not only considered in isolation, but also when compared against its closest brain-with-behaviour competitors.

What, then, can be concluded from about the statistical robustness of the SES/IFG correlation found here? This correlation between environment and neural specialisation is strong in itself, and is markedly stronger than the other correlations between behavioural/environmental and brain variables in our data set. Our findings draw attention to an empirical link, and it is our belief that they provide a motivation and rationale for further investigating that link in future work.

Supplemental Figures

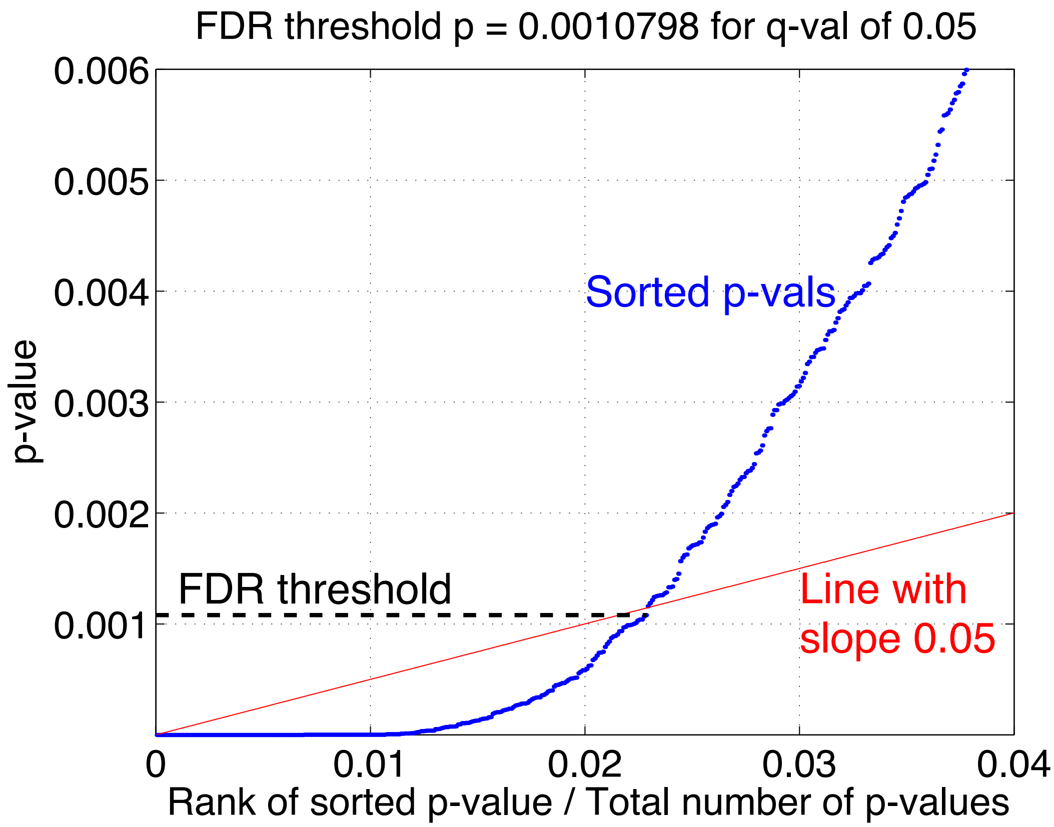

Figure 1: Implementation of the False Discovery Rate correction. All the p-values in the correlation matrix are sorted into an ascending list, and its point of intersection with a line with gradient of 0.05 is determined. The p-value at this intersection point provides the FDR-threshold.

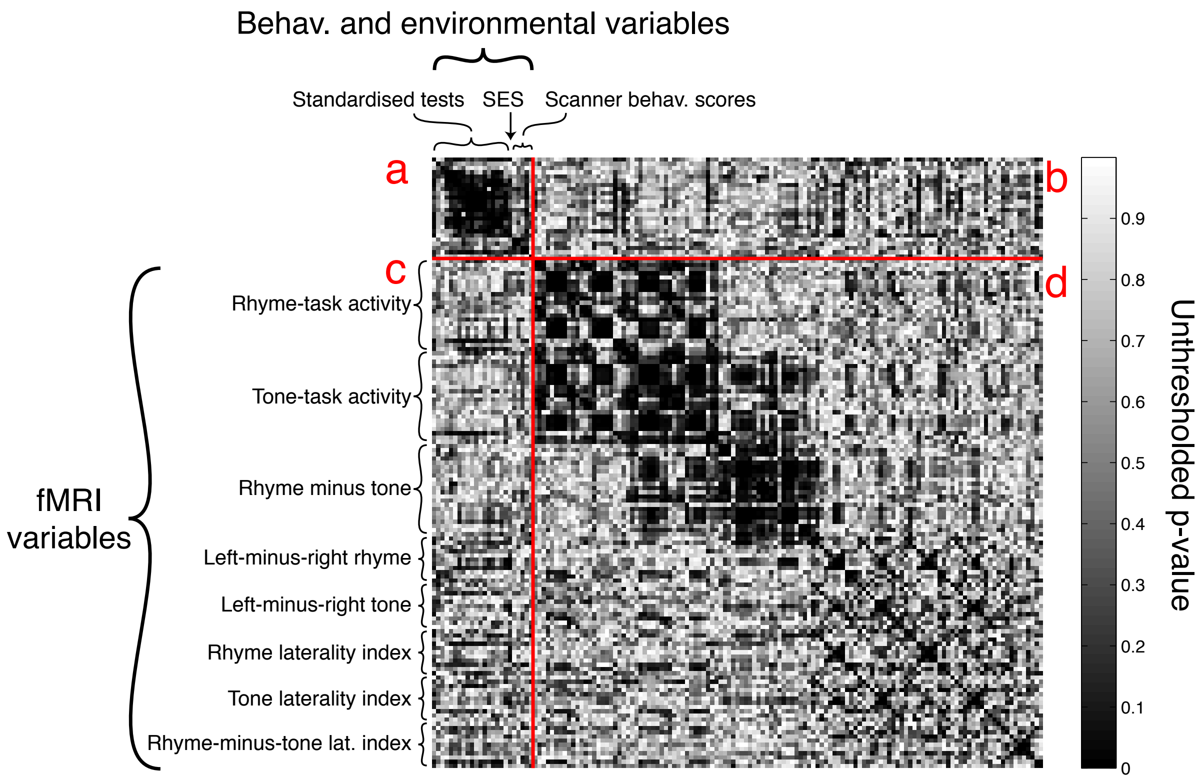

Figure 2: The same correlation matrix as in Figure 1 of the main text, but now showing the unthresholded p-values for the correlations, rather than the binary significant-vs.-nonsignificant shading in the main text's plot. The quadrants are labeled in the same way as main Figure 1: (a) contains correlations within the behavioural and environmental variables; quadrant (d) contains correlations within the fMRI variables; quadrants (b) and (c), which are symmetrically identical, contain the correlations between the behavioural/environmental variables and the fMRI data.

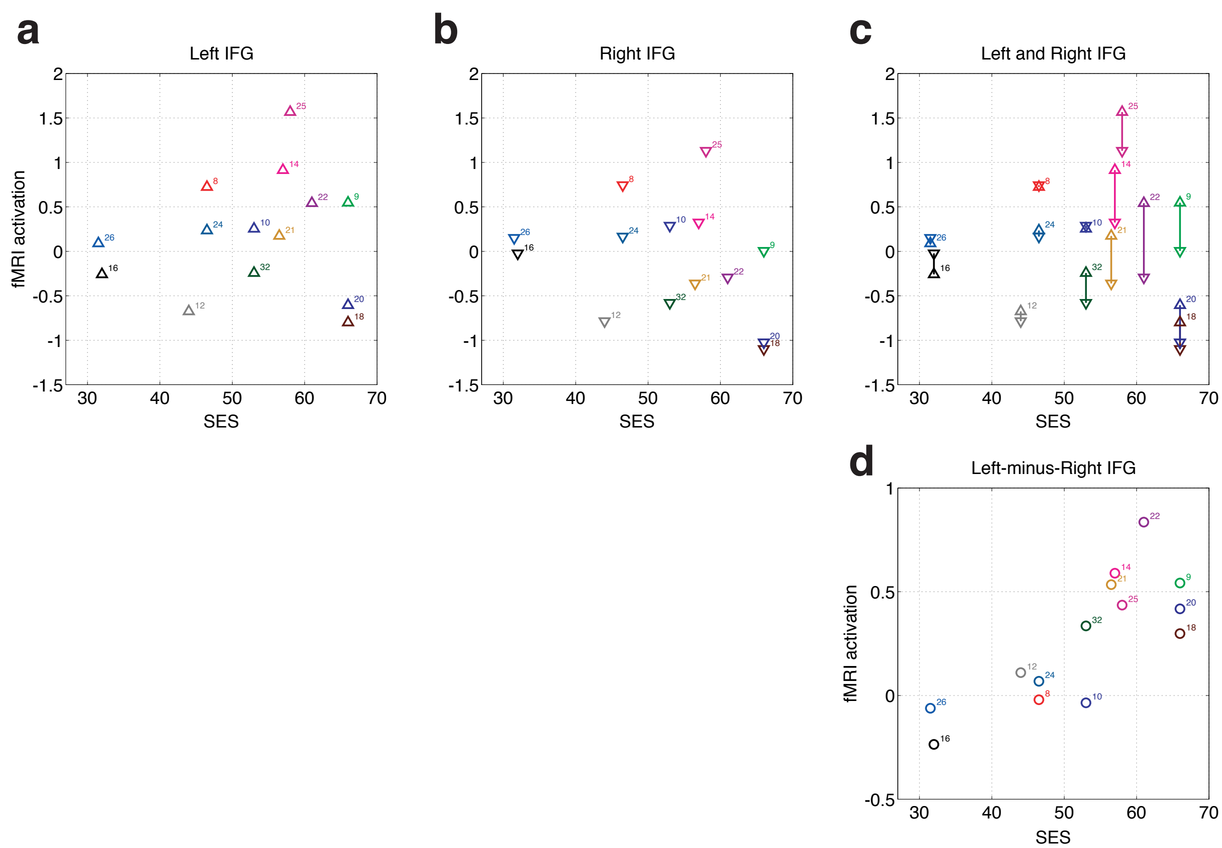

Figure 3: Rhyme-task activation in the left IFG and right IFG shown separately and together, all plotted in relation to SES. **(a):** Left IFG **(b):** Right IFG **(c):** Left and right IFG **(d):** Left-minus-right IFG. The numbers next to each point are the subject identity codes. It can be seen that neither the left nor right hemisphere activations considered individually correlates with SES, but that the left-minus-right activation difference is strongly correlated. A likely explanation for this result is that baseline fMRI activation values often show large variability across subjects. The average activation with a single ROI from a given subject will reflect this variability. However, by taking two measurements from within each subject and computing the difference between them, much of this inter-subject baseline-variability is canceled out. The left-minus-right activation is of course just such a within-subject subtraction. This left-minus-right measure should therefore suffer from much less baseline-variability noise than the single-hemisphere activations, thus allowing correlations with environmental variables such as SES to manifest themselves more cleanly.

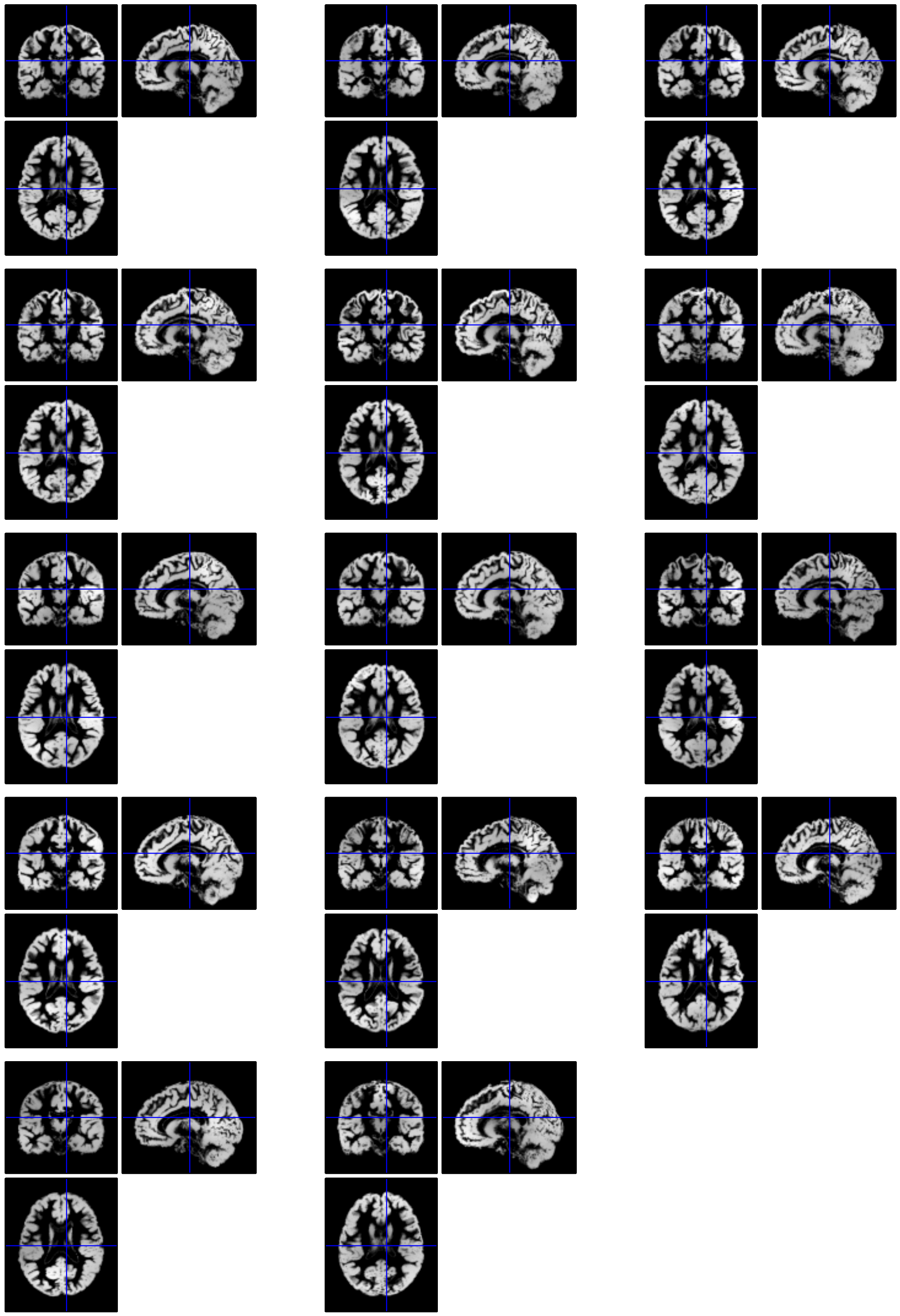

Figure 4: Grey matter segmentations for all fourteen of the children in the study, carried out using Christian Gaser's VBM Toolbox for SPM5. It can be seen that good quality segmentations were produced in all cases, despite the fact that some head-motion induced noise and artifacts were present in the original MRIs.

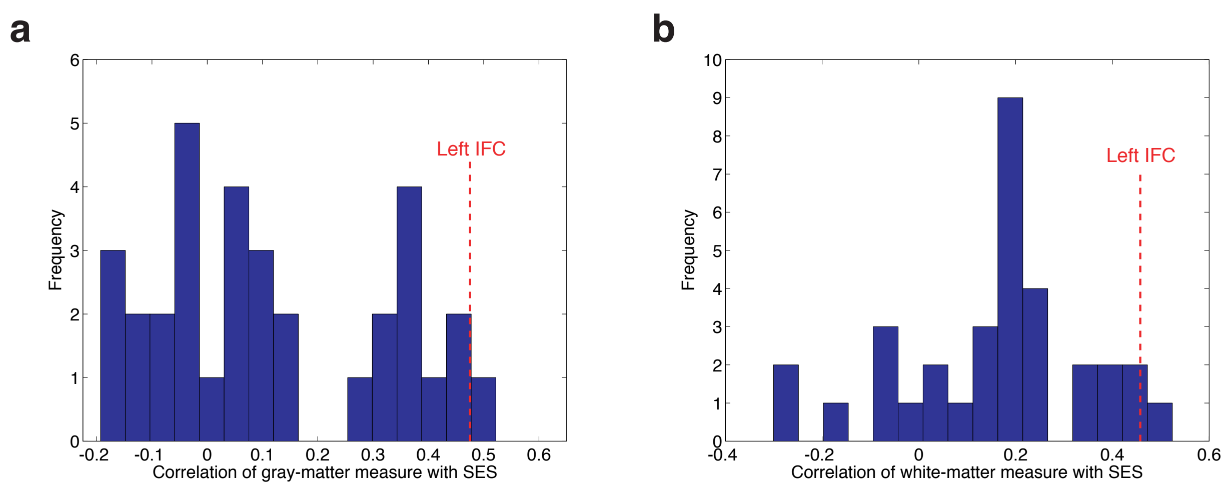

Figure 5: No ROI in either hemisphere outside of the left IFG had a stronger correlation with SES than did the left IFG itself, either for grey matter volume, white matter volume, or for left-minus-right volume for either gray or white matter. The only stronger correlations that emerged were from subparts of the left IFG itself. **(a):** Histogram of correlations with SES for all the ROIs for all of the above measures. The left IFG grey matter correlation ranks second, slightly behind the grey matter correlation for the left IFG orbital subpart. **(b):** Histogram of white matter volume correlations with SES. The left IFG white matter correlation again ranks second, this time slightly behind the grey matter correlation for the left IFG opercular-triangular compound subpart.

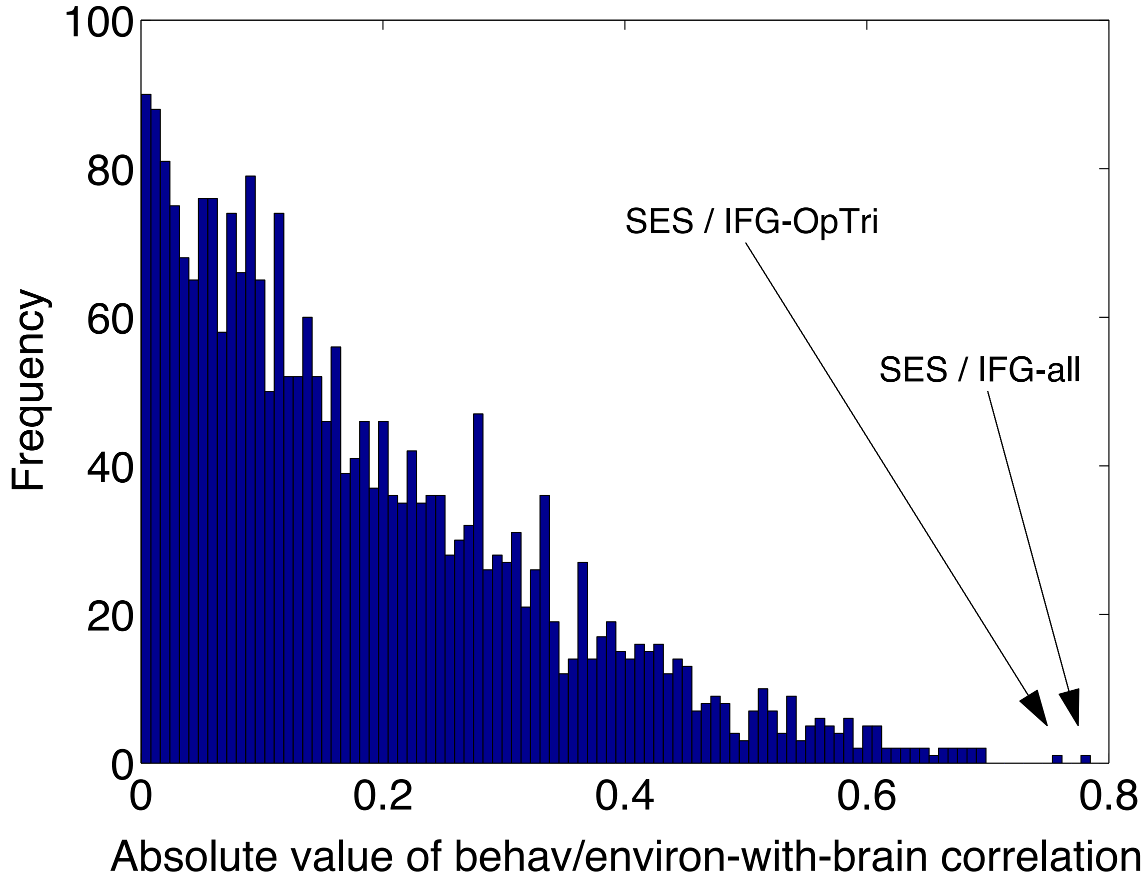

Figure 6: Histogram of the correlations between behavioural/environmental and brain variables. It can be seen that two of the correlations stand out as being stronger than the others. The strongest of these is the correlation between SES and the left-minus-right rhyme-task asymmetry in all three parts of the IFG taken together. The second-strongest is the correlation between SES and the ROI consisting jointly of the opercular and triangular parts of the IFG.

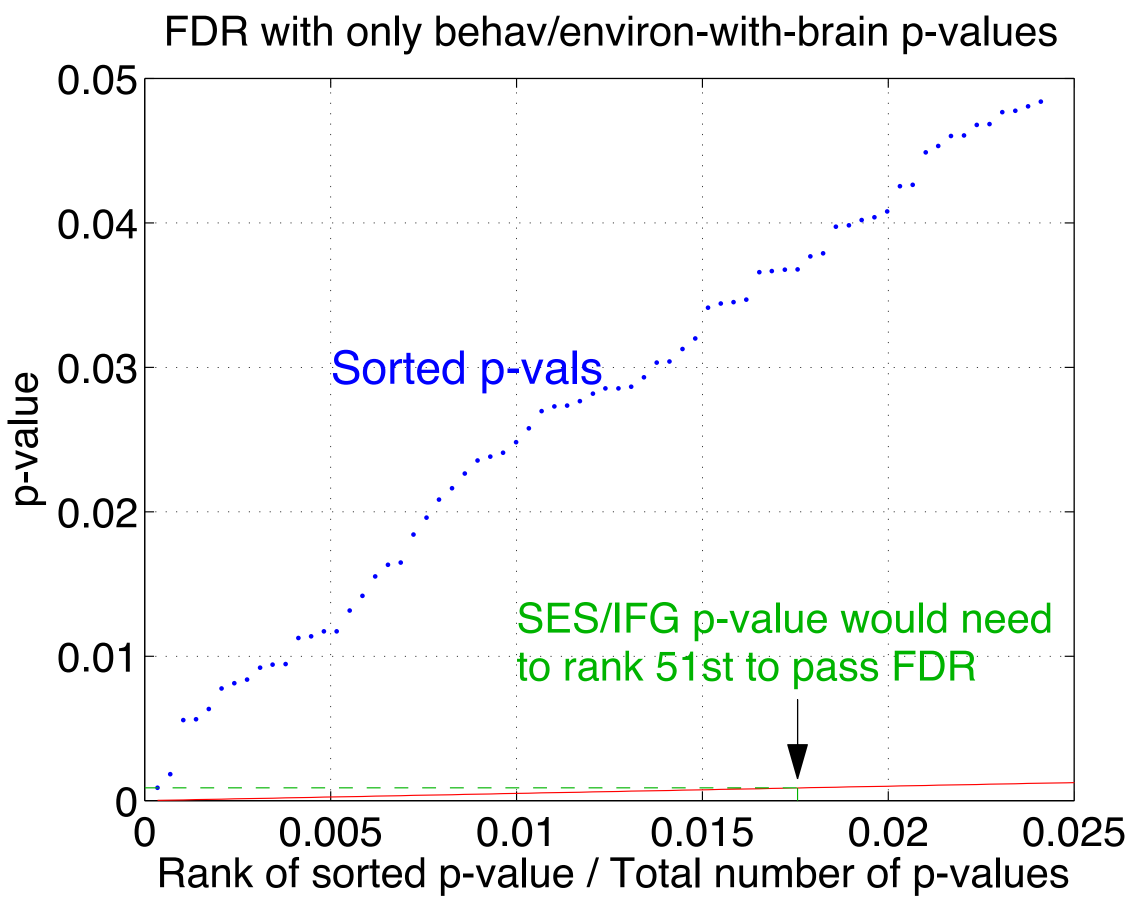

Figure 7: When only the correlations between behavioural/environmental variables on the one hand and brain variables on the other hand are considered, the lowest p-value (which corresponds to the SES/IFG correlation), does not fall underneath the line with gradient 0.05, and therefore would not pass an FDR test. As the histogram in Supp.Fig.6 also illustrates in a different form, two correlations stand apart as being stronger than the others: the correlation between SES and all three parts of the IFG, and the correlation between SES and the opercular and triangular IFG-subparts together. FDR can capture the significance of individual tests only when they are accompanied by many other tests with equally low p-values. Thus, FDR is ill-suited for capturing the structure of this particular dataset.