

The Untenability of the Euclidean Brain

Fantastic Four

December 15th, 2018

Contents

Team Information	2
Note to the Reader	2
Abstract	2
Introduction	2
Description of Data Set	3
Descriptive Exploration of Data	5
Preprocessing	5
Voxelwise Analysis	5
ROI analysis	7
Visual Comparison of ROI Intensities	8
Exploration of Laterality	8
Region Intensity Across Stimuli	10
Planned Analysis	10
Construction of Correlation Matrices	10
K-Means Clustering	13
Alternate Classification Approaches	18
Random Forest	19
Support Vector Machines	19
Discussion	19
Appendix A.	20
Philosophical Assumptions	20
References	21

Team Information

Team Members: M.K. Finnegan¹, K. Anapindi; DH Hwang; R Dalmia

1. *Corresponding Author*

Note to the Reader

The code that generated these results can be found at [this github repository](#). Data is too large to submit with project. The data is available upon request to the corresponding author at mkfinne2@illinois.edu

Abstract

Mind wandering (a type of task unconstrained thought) is an important phenomenon that is presumed to occur during resting state imaging paradigms. Little is known about the temporal dynamics of these thoughts during the resting state; however the dynamics of brain connectivity which presumably represent such thoughts are beginning to be explored with dynamic connectivity in fMRI. Current techniques for classifying similarities of brain connectivity in this approach have not been evaluated for the meaningfulness of their classifications. This proposed study uses inferred cognitive states as the “true classification” of a given brain state allowing for testing of classification accuracy. Under the principled presumption that classification accuracy during task paradigms can provide an estimate of accuracy during rest paradigms we use two emotional processing tasks to evaluate these methods. Cognitive states are inferred by the stimuli being presented during two different block design tasks and metrics of psychopathology such as depression and anxiety. Modern clustering algorithms are tested as to whether they provide clusters that accurately overlap stimuli presentation during two block design tasks subject to potential covariance with measures of psychopathology symptoms.

Introduction

There is a wealth of evidence that when humans aren’t engaged in any externally oriented task, they are consistently immersed in a phenomenon called mind-wandering (1,2,3,4,5). This is a type of unconstrained thought that may facilitate creativity and problem solving (6), but in more maladaptive forms, may be responsible for the recurrence and maintenance of both depressive and anxious disorders (7,8,9,10).

Dynamic connectivity is comparatively recent development in functional neuroimaging analysis that shows immense promise towards understanding the role of mind-wandering in health and disease (11). For the interested reader, please refer to the original study proposal for a more comprehensive overview of the concepts involved in dynamic connectivity. Briefly, by calculating brain connectivity (*a.k.a* correlations of hemodynamic fluctuations across brain regions) on short time windows (on the order of seconds to minutes) rather than across an entire scan (which is typically on the order of 10 to 20 minutes), we can capture the temporal fluctuations of brain activity across a scan (12). These fluctuations presumably reflect, in part, the fluctuations of the internal contents of thought.

A current common practice to label similar connectivity matrices produced by this approach is k-means clustering using either the L1 (Manhattan distance) or L2 (Euclidean distance) norms (12,13). Typically, the k-means derived clusters are assumed to represent either unitary cognitive states or transitions between two unitary states. The differences of members across cluster are then either qualitatively or quantitatively analyzed (such as noting the degree of connectivity between executive and emotional systems) to make inferences about the nature of thought across time. In fact, this approach has been used to argue that short term meditation training decreases the mind’s propensity to wander (14) and that people who are more likely to be creative possess unique types of mindwandering (15).

K-means clustering however, will always produce clusters regardless of how well those clusters represent the underlying distribution of the data. Under certain distributions, k-means clusters have little, if any relation to the regularities inherent in the data. The algorithm underlying k-means seeks to minimize the within-cluster distance objective function. For the case of the commonly used Euclidean distance:

$$\sum_{i=0}^n \min_{\mu_j \in C} (\|x_i - \mu_j\|^2)$$

This makes the implicit assumption that the true clusters are both convex and isotropic (*i.e.* their boundaries can be defined by circles drawn around them). This means the k-means will not perform well when the data is non-convex, has anisotropic covariance, has spatially overlapping memberships, or contains outliers which artificially skew cluster means. There is no known evidence that that brain connectivity as represented in \mathbb{R}^n meets any of these assumptions.

The purpose of this study is to examine the suitability of k-means in real data. In order to test the accuracy of the clustering results, two task-based paradigms are used in which the “true” classification of a brain state is assumed to be the stimuli presented at that time. See Appendix A (material originally introduced in the project proposal but for convenience to the reader, included here) for a detailed exposition of the philosophical assumptions justifying this approach. This allows us to test the derived classification against the “true” labels. Two tasks within the same subject cohort are employed as a means to assess the robustness of k-means clustering across paradigms while controlling for subject demographics. Each task was repeated across two scans within each subject (for a total of 4 scans per subject), which is leveraged to ensure algorithms are not fitting technological sources of variance such as magnetic gradient nonlinearities or subject head position induced distortions (16) which may vary across scans. As an extension of this work, alternate classification approaches are explored for potential improvements to current dynamic connectivity analysis practice.

Distribution of the raw data imaging data is restricted due to HIPAA considerations, thus is not included in this report. However, all derived data, including timecourses used to produce the correlation matrices can be found in the link to the team’s Box account submitted with this report.

Description of Data Set

Participants

The results reported here are based on a subset of women from a larger study examining the role of stress physiology in mother-infant interactions. They were primarily recruited from community agencies serving low-income women in a mid-size city of the Pacific Northwest. Twenty-five mothers with 3-month old infants agreed to participate in the additional neuroimaging reported here. Mean age for these mothers was $\mu = 26.4$ ($SD = 3.8$). Most identified as Caucasian (72%), with 12% reporting Latina ancestry, 8% Asian American, and 8% other. Eighty four percent reported some education beyond high school, with 24% having completed a college degree. No infants had serious health complications and the majority were born on time (4% before 37 weeks and 8% after 41 weeks gestation). This dataset is also described in 17. Ethical approval for the study was obtained from University of Oregon Institutional Review Board and all mothers provided written informed consent.

Scan Paradigm

Participant engaged in two separate tasks for this study, one involving infants expressing emotions and the other presenting emotionally expressive faces, however only the task observing infant emotional cues is reported here. The task was repeated across two scans. Scans were counter balanced across tasks to control for order effects and subjects were allowed to shift positions between scans in order to minimize discomfort related head movements. See Figure 1 for an overview of both paradigms.

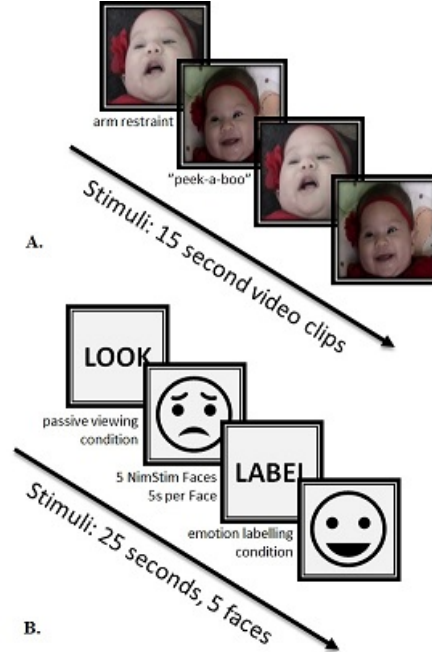


Figure 1: Schematic of the two tasks that produced the data. (A.) Mothers viewed either their own or others’ infants displaying either positive (induced by playing peek-a-boo) or negative (induced by arm restraint) emotions. (B.) Mothers were instructed to either passively view or label the emotional content of humans faces.

Infant Viewing Task

Mothers were instructed to “watch and respond as you would naturally” to videos of their or unknown others’ infants expressing either positively or negatively valenced emotional cues. Emotional expressions in the infants were evoked during a video recording session occurring in a home setting. Positive valence expressions occurred during a peek-a-boo game played with the mother. Negative valence was induced by arm-restraint, a task known to evoke frustration response in infants (18). Each run lasted for approximately 7.5 minutes and consisted of alternating blocks of the infant emotion videos edited to be 15s in length.

Facial Viewing/Labelling Task

A second task involved a well studied paradigm of emotion regulation in which mothers viewed emotionally expressive faces from the NimStim database (19). Face images were race matched to the participants to take advantage of heightened neural responsiveness to in-group faces (20). Three second instructions appeared prior to each stimulus block with the words “Look” or “Label”, in which participants either were to passively view or label the emotional content of the faces. Blocks consisted of 5 faces of a given positive/negative valence (*e.g.* happy, sad, angry, etc.) each presented for 5 seconds each. Blocks were separated by 1-3s of jittered rest condition (alternately referred to as either “fixation” or “ISI”, short for interstimulus interval). Each scan lasted for approximately 7.6 minutes.

Scan Acquisition

Scans were acquired at the University of Oregon using a Siemens Allegra 3 Tesla magnet. Whole brain images were taken with a 32-channel phase array birdcage coil. Shimming routines were conducted to optimize signal-to-noise with a subsequent fast localizer scan (FISP) and Siemens autoalign routine at the outset of every session. The two functional scans were then conducted followed by a high resolution anatomical scan.

Function Scans

All tasks used a T2*-weighted gradient echo sequence, with TR = 2000 ms, TE = 30 ms, flip angle = 90°, 32 contiguous slices acquired in ascending interleaved order, with thickness = 4 mm, 64 × 64 voxel matrix; 226-228 volumes per run.

Anatomical Scan

Anatomical images used a T1-weighted 3D MP-RAGE sequence, with TI = 1100 ms, TR = 2500 ms, TE = 3.41 ms, flip angle = 7°, 176 sagittal slices 1.0 mm thick, 256 × 176 matrix FOV = 256 mm.

Descriptive Exploration of Data

Preprocessing

All preprocessing was carried out in the SPM12 software (21) running under MATLAB 2018a on Windows 10 using a custom written batch interface. Because of the extensive time needed to run computations serially, all preprocessing routines were parallized across subjects. Preprocessing included slice-timing correction which is use to account for the fact that, due to the interleaved slice acquisition order, adjacent data points will represent activity up to 1 second apart in time. Data points are sync interpolated either forward or backward in time to the reference slice, which was chosen to be the slice collected 1s into volume acquisition. Next, volumes within each scan underwent realignment and unwarping to correct for head motion artifacts across the scan.

Realignment consisted of a 2-pass procedure in which a mean image within a scan was calculated and volumes were put in alignment using a 6-parameter least squares approach to rigid body transformations. These estimated alignment parameters were then inspected to remove subjects with excessive head motion defined as more than 40% of volumes containing greater than 2mm of estimated headmotion. This is a very liberal motion limit, however in clinical or hard-to-recruit samples, it is very common to relax more stringent limits and the particular value was taken on recommendation of the PI who supervised data collection. With this criteria, one subject had to be discarded for excessive movement. Unwarping was also conducted to account for gradient field distortions that occur due to this movement. Although fieldmaps were acquired which provided an absolute measure of the degree of magnetic homogeneity throughout the field of view, they were not referenced during unwarping due to the fact that subjects moved substantially between scans making the field map not applicable to all runs. Instead deformations were estimated using SPM12 algorithms.

Functional images were then co-registered to the high resolution structural image for each subject. This involves similar computations to the realignment step, however alignment occurs across scans. The anatomical images were then used to calculate non-linear deformations to warp the brain into the Montreal Neurologic Institute (MNI) standard stereotaxic space in a process known as normalization. The detailed deformation fields were then applied to the low resolution functional images. This process allows for voxel-by-voxel comparisons across subjects and is necessary to apply the region-of-interest masks needed to extract data for k-means. Given that fMRI typically has a relatively poor signal-to-noise ratio, all functional images were smoothed with an 8mm FWHM Gaussian smoothing kernel in order to compensate for spurious fluctuations.

Voxelwise Analysis

In order to confirm there was task correlated signal that could potentially drive clustering, a voxelwise level analysis was conducted to identify the location of point sources of signal before collapsing into region-of-interest analyses. SPM12 was also used for this phase of data exploration. After preprocessing, a general linear model was calculated for each subject at each voxel using a design matrix consisting of several components. First the onset of predictors, but because fMRI measures changes in the ratios of oxygenated and de-oxygenated hemoglobin, there is a temporal delay between stimuli presentation and observed response. To

Infant Viewing Task	Emotional Faces Viewing/Labeling Task
<u>InfOwnNeg > InfOwnPos</u> <u>InfOwn > InfOther</u> <u>InfOther > InfOwn</u> <u>InfPos > InfNeg</u> <u>InfNeg > InfPos</u> <u>InfPos > Fixation</u> <u>InfNeg > Fixation</u>	<u>EmoLabelPos > EmoLookPos</u> <u>EmoLabelNeg > EmoLookNeg</u> <u>EmoLabel > EmoLook</u> <u>EmoLookPos > EmoLookNeg</u>
<p><i>Key: Inf = infant viewing task; Own = mothers' own infant; Other = stranger's infant; Pos = peek-a-boo videos; Neg = arm restraint videos; Fixation = rest condition with fixation cross; Emo = emotional faces (NimStim) task; Look = passive viewing of faces; Label = labeling the emotion of faces; Pos = positive facial expressions; Neg = negative facial expressions</i></p>	

Figure 2: Voxelwise contrasts showing significant results

account for this, these predictors were convolved with what is known as the *hemodynamic response function* which estimates this delayed response for an instantaneous stimuli. Next, the estimated motion realignment parameters to remove any remaining variance associated with head motion. Then a 6 predictor cosine basis set provided by the SPM12 software and used to account for scanner drift, in which the magnetic fields slowly oscillate over time (which creates low-frequency oscillations in the voxel intensity values), this essentially acts as a high-pass temporal filter set *a priori* at 128Hz.

Due to the relatively high amounts of motion discovered in this data, additional regressors were added to deweight volumes exceeding 2mm of realignment. Data was pooled across scan types to reduce the number of individual models to estimate and dummy variables accounting for run were also used. Because of the large amounts of regressors, the design matrix was explored for collinearity using vector cosines, although there was some degree of collinearity, no regressors were flagged by SPM12 integrity checks for levels of dependence that would impeded estimation. Models were estimated using an implicit mask to restrict analysis only to voxels likely to be brain (determined by mean intensity across volumes), and thus reduce the problem of multiple comparisons in hypothesis testing. Only the regression coefficients associated with stimuli were explored for significant differences.

Contrasts were constructed for all pairwise contrast of regressors for every subject and t-values constructed using a fixed effects approach (the brain is the entire 'universe' of voxels). These statistics were used in a second group-level random effects analysis taken across subjects to explore signal across the entire dataset. Statistical correction for multiple comparisons was done in a hierarchical manner using the recommendations of (22) using a primary cluster-defining threshold of $p < .001$ uncorrected, and cluster selection with FWE-corrected cluster-level significance set at $p < .05$. The following table provides a summary of significant contrasts. The images that follow illustrate the extent of clusters discovered.

As can be seen in the table, both tasks provide significant differences in voxel intensity values across tasks. Within the infant task, there were significant differences in both directions across the Own/Other and Pos/Neg distinction. Within a mother's own infant, there did appear to be a consistent directional effect of valence across all voxels and similarly between both valence conditions and fixation. Although both tasks contain task-meaningful signal, the infant task produced less extensive differences as is illustrated in smaller clusters evident in the infant contrast images shown.

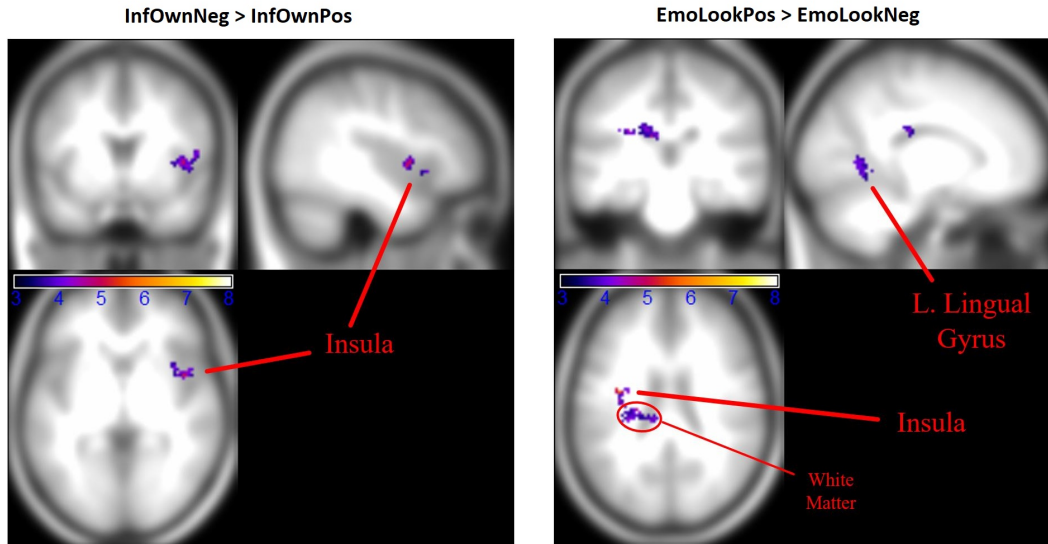


Figure 3: Result images of selected voxelwise contrasts

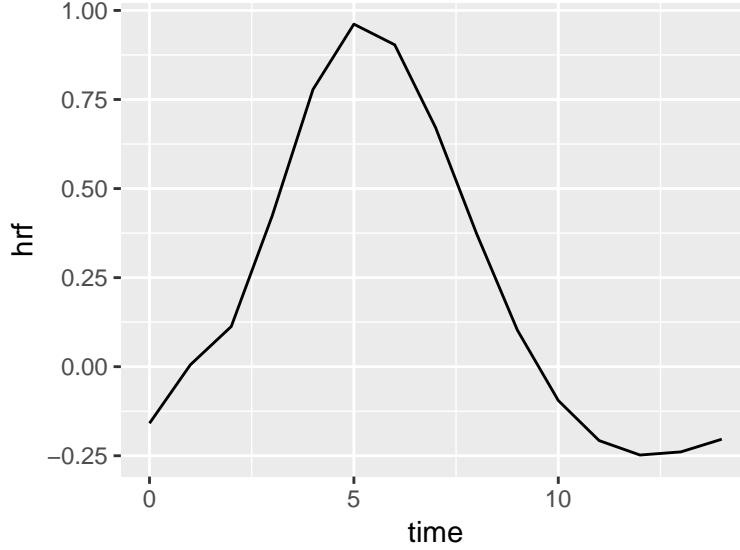
ROI analysis

Because the intended targets of clustering are not voxels themselves, but rather the correlations of anatomically defined voxel averages, we also explored how much of this task-dependent signal was preserved after voxel averaging. This is known as a region-of-interest (ROI) analysis. Regions of contiguous voxels were defined by reference to the Automated Anatomically Labelled Atlas (23) and the MarsBar toolbox addition for SPM (24). Mean intensity values were extracted for all defined regions, for all time points and for all subjects. All subsequent analyses reported here were conducted and visualized in R.

Linear Regression

Linear regression was again used to assess for assessing the strength of the relationship between task and signal, however several using a simplified regression model. Stimuli regressors convolved with the hemodynamic response function were still used, however we estimate this with the `hrfConvolve` function in the FAIR package. The HRF is implemented as a double gamma probability distribution parameters set for the commonly accepted model. A demonstration of this model on a finite impulse response shows the classic shape of the HRF that emerges.

```
# install.packages('FIAR') ## If needed
library('FIAR')
library(ggplot2)
fir = c(1,rep(0,14))
# Note that the data is assumed to be take every 1s.
# This was not listed in the documentation.
hrf=hrfConvolve(fir) # some example data
exmpl = data.frame(time=seq(0,14),hrf=hrf)
ggplot(data=exmpl, aes(x=time,y=hrf))+
  geom_line()
```



Unfortunately, R has not been extensively used for brain image analysis and there are very few resources for MRI-specific analysis. This is a poorly documented function and the package itself is still not fully tested, so results of this analysis should be seen as highly tentative. To simplify calculations, bulk head movement and volume deweighting was not included, but a 4 regressor cosine basis was included. All subjects were included in a single linear model with regressors of indicator variables for each ($n=24$) subject and the first or second run within subjects ($s=2$).

As can be seen in this representative linear model summary of the left caudate (Figure 4), several predictors and in particular fixation, InfOtherPos, and InfOwnNeg all significantly predicted ROI intensity values. This was consistent across most ROIs, although different combinations of predictors reached significance depending on the particular region. This left us with confidence that task correlated signal at the voxel level was indeed preserved through the smoothing process.

Visual Comparison of ROI Intensities

Given the discrepant findings between the relevance of task at the voxel and the ROI level, we plotted the average intensity vs. the anatomical brain region of interest for visual inspection. This stacked area graph represents all the five different stimuli, with ISI being the baseline. The other 4 stimuli (EmoLabelPos, EmoLabelNeg, EmoLookPos, EmoLookNeg) are stacked on top of it. From this graph, we can visualize whether the brain activity in a particular region changes w.r.t. the stimuli or if the activity remains constant irrespective of the stimuli. From this graph, we did observe abrupt changes in the mean voxel intensity for certain regions (depicted by the step like pattern towards the top right) w.r.t to the control ISI intensity. This tells us that the voxel intensities do capture the brain changes in intensity w.r.t the stimuli and although this may not be apparent within individual ROIs, is still reasonable to expect to be represented in the correlation matrices that reflect between ROI coordination.

Exploration of Laterality

This cohort was relatively high in depression symptoms with 9 out of 24 mothers meeting criteria for past or current depressive episodes or diagnosis. Depression is a very common postpartum complication (25) with devastating consequences for both mother and infant (26,27,28) and is 2nd leading cause of death (via suicide) for mothers in the first year of the infant's life (29). It also has been known in neuroscience for a very long time that the brain exhibits a laterality in depression (30) typified by persistent left frontal and

Infant 2
Region: Caudate_R

```
Call:
lm(formula = roi.comb[, i] ~ design.comb)

Residuals:
    Min       1Q   Median       3Q      Max
-44.997  -0.740   0.021   0.781   6.263

Coefficients: (5 not defined because of singularities)
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    395.262379    0.102644  3850.799 < 2e-16 ***
design.combfixation    0.035497    0.017542   2.024 0.043063 *
design.combInfotherNeg -0.006379    0.017574  -0.363 0.716628
design.combInfotherPos  0.034734    0.017625   1.971 0.048810 *
design.combInfownNeg    0.057155    0.017232   3.317 0.000917 ***
design.combInfownPos      NA         NA         NA      NA
design.combt1    -0.071678    0.029404  -2.438 0.014812 *
design.combt2      NA         NA         NA      NA
design.combt3      NA         NA         NA      NA
design.combt4      NA         NA         NA      NA
design.comb      49.534699    0.138762  356.977 < 2e-16 ***
design.comb     -32.533872    0.138607 -234.721 < 2e-16 ***
design.comb     -46.618546    0.138608 -336.335 < 2e-16 ***
design.comb     -28.695415    0.138605 -207.029 < 2e-16 ***
design.comb      0.345853    0.138762   2.492 0.012717 *
design.comb     -8.256510    0.138608  -59.567 < 2e-16 ***
design.comb      9.595926    0.137999   69.536 < 2e-16 ***
design.comb      3.867909    0.137698   28.090 < 2e-16 ***
design.comb     10.368539    0.137996   75.137 < 2e-16 ***
design.comb     25.228370    0.137996  182.820 < 2e-16 ***
design.comb     109.293036    0.138148  791.129 < 2e-16 ***
design.comb      34.367227    0.137843  249.322 < 2e-16 ***
design.comb     10.818193    0.137999   78.393 < 2e-16 ***
design.comb     21.045814    0.137996  152.511 < 2e-16 ***
design.comb     63.682048    0.138145  460.979 < 2e-16 ***
design.comb     95.764307    0.138300  692.439 < 2e-16 ***
design.comb     52.101537    0.138146  377.147 < 2e-16 ***
design.comb     25.314772    0.137994  183.449 < 2e-16 ***
design.comb     48.424333    0.138148  350.524 < 2e-16 ***
design.comb     -15.079633    0.138299 -109.037 < 2e-16 ***
design.comb     -16.309612    0.137845 -118.319 < 2e-16 ***
design.comb     -91.843689    0.137844 -666.287 < 2e-16 ***
design.comb      16.698848    0.137998  121.008 < 2e-16 ***
design.comb      NA         NA         NA      NA
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.477 on 5438 degrees of freedom
Multiple R-squared:  0.9988,    Adjusted R-squared:  0.9988
F-statistic: 1.661e+05 on 28 and 5438 DF,  p-value: < 2.2e-16
```

Figure 4: Summary of regression modelling in the Infant Task

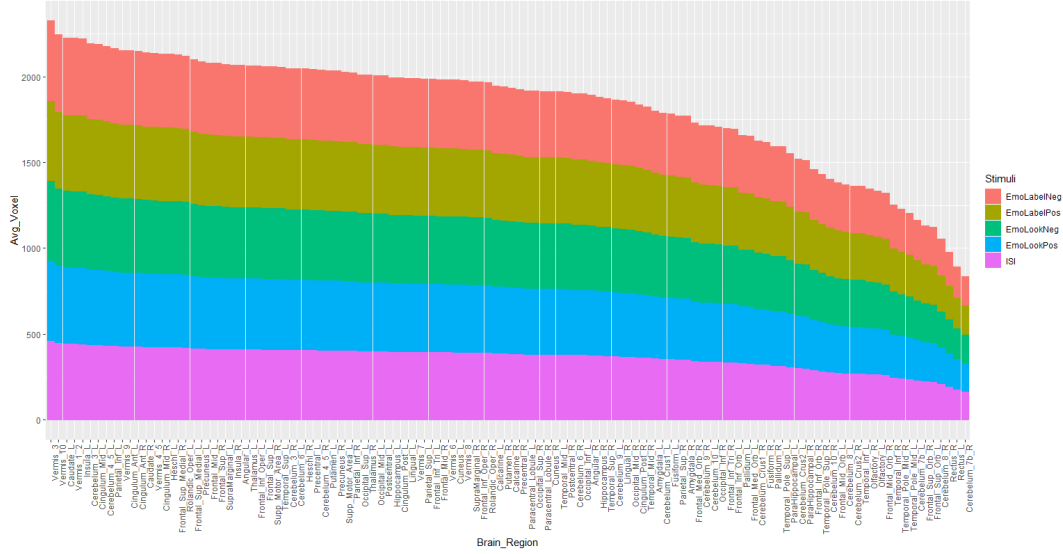


Figure 5: Stacked graph of ROI intensity across regions

right parietal hypoactivation (31). To detect the presense of laterality between left-right homologous ROIs, we performed an paired, 2 -tail t-test for all the regions of a given stimulus between the left and right brain regions. Before performing the t-test, we tested for the distribution of the data (See Figure 6.) to ensure that all the necessary conditions to perform the hypothesis test are met.

The results that we got from the t-test were quite interesting. After accounting for the multiple testing correction (Benjamini Hochberg), it turned out that almost every single region of interest actually had a significantly different ($p < 0.05$) mean voxel intensity between the left and the right regions. It is unclear how much of this laterality difference may be driving the results presented below, but should be noted that this raises the ponential of implicit markers of psychopathology skewing results.

Region Intensity Across Stimuli

A final aspect of the data that we wished to characterize was the inter-region variability vs. the inter-stimulus variability of the voxel intensity. A box-whisker plot (Figure 7) of representative regions was created to visually inspect this relationship. Regions from the boundaries of the field-of-view were excluded from these plots as they tend to have artificially low intensity values due to signal dropout near the edges of the shimming region. Based this plot, it is evident that the inter-region variability is much higher compared to the inter-stimulus variability.

This result discouraged us from pursuing a clustering based on regional intensity values (as opposed to the correlations across regions), because there did not seem to be sufficient separation across tasks within any regions to warrent the computational expense. This is however, left as an avenue of future exploration.

Planned Analysis

Construction of Correlation Matrices

The impetus for this project is to assess the accuracy of k-means clustering when (under the philosophical assumptions in Appendix A) the “true” label of the brain states are known. To be reflective of the broader literature, brain states were conceptualized as connectivity and thus correlation matrices of ROI timecourses

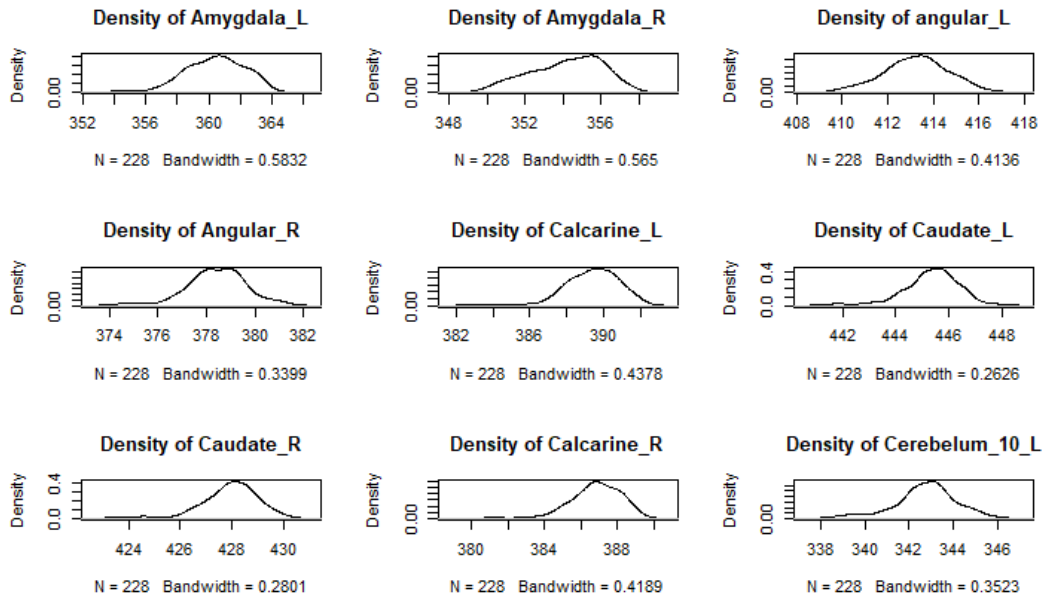


Figure 6: Density distributions of select ROIs

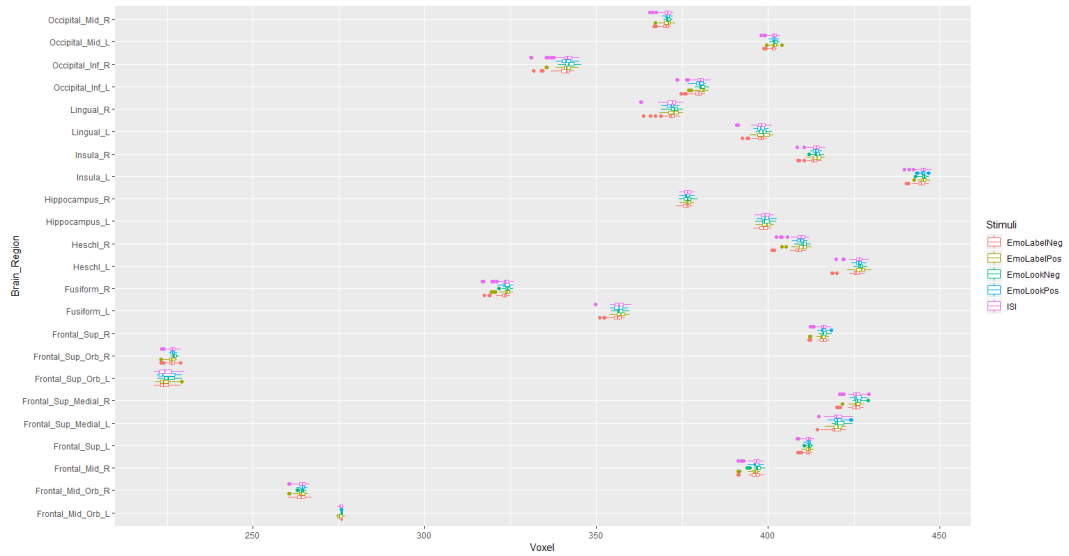


Figure 7: Box-whisker plots of select ROIs across stimulus type

were constructed. The planned method was originally to use a sliding window approach to extract correlation matrices representative of the stimuli being presented, but it was later discovered that the timing of the stimuli did not provide wide enough windows for this approach to be feasible. A modified version was implemented in which stimuli labels were shifted forward in time to be representative of when they were contributing the most to the hemodynamic response. This was accomplished by creating indicator variables of the time points in which stimuli were presented, $I(s, t) = 1, 0 \text{ o.w.}$. These were then convolved with the hemodynamic response function to produce the design matrices used in the exploratory linear regression. The code used to generate these shifts and matrices is outlined below.

```
ShiftVarNames <- function(stimuli){
  # Creates the design matrix, shifts everything by the HRF, then assigns new labels
  # Build design matrix
  design = data.frame(matrix(0, nrow = length(stimuli), ncol = length(levels(stimuli))))
  colnames(design) <- levels(stimuli)
  for(s in levels(stimuli)){
    design[which(stimuli == s),s] <- 1
  }

  # Due to an undocumented feature of the hrfConvolve function, when raw data is passed,
  # the assumed acquisition rate (TR) is 1 second. Since our data has TR = 2, expand
  # design matrix to account for this.
  expandedDesign = rbind(design, design)
  n = nrow(design); r <- rep(1:n, each = 2) + (0:1) * n
  expandedDesign <- expandedDesign[r,]; rownames(expandedDesign) <- c(1:(n*2))
  # Shift value in each column by the HRF
  expandedDesign = apply(expandedDesign, MARGIN = 2, hrfConvolve)
  # Down sampling - taking the value at the first second of onset as the instantaneous
  # value due to slice timing correction to the time point in preprocessing.
  design <- expandedDesign[seq(1,nrow(expandedDesign),2),]

  # Reassigning stimuli labels according to what has the max value
  maxStimIdx = apply(design, MARGIN = 1, which.max)
  newStim = levels(stimuli)[maxStimIdx]
  # Although convolution may assign higher values to other stimuli, early time points
  # should have the label of the first stim presented until it reaches it's first peak
  firstStimIdx = which(newStim == stimuli[1])[1] # Index of first occurrence of first stim
  newStim[1:firstStimIdx] <- levels(stimuli)[stimuli[1]]

  data = data.frame(newStim, design)

  return(data)
}

# Needed to generate correlation matrices
for (j in 1:96) {
  mydata = read.csv(file = files[j])
  stimuli = as.vector(unique(mydata[, 1]))
  for (i in stimuli) {
    data = mydata[which(mydata[, 1] == i), ]
    corr = cor(data[, c(2:117)])
    write.csv(corr , paste0(substr(filename[j], 1, 34), i, ".csv"))
  }
}
```

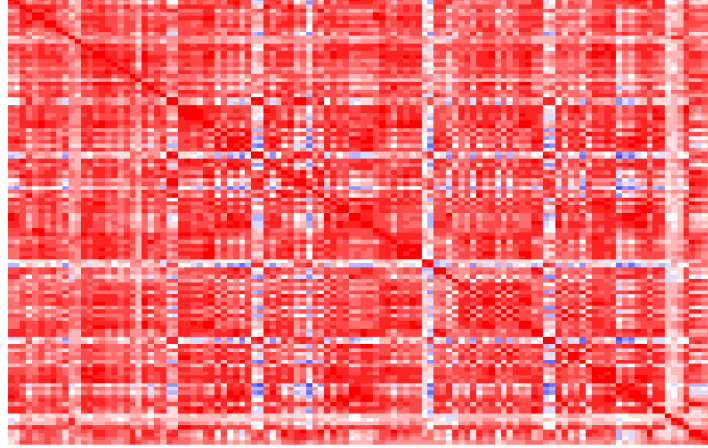
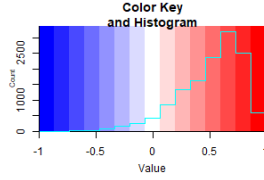


Figure 8: Heatmap for an example subject’s connectivity matrix of EmoLabelNeg

The stimuli with the max value at a given time point was taken as the singular stimuli contributing it’s effects. We recognize this presents a potential flaw in that we are not considering the full scope of stimuli impinging on brain activity, but this simplification led to more interpretability of results and is pointed to as an area in need of future development.

To create stimuli representative matrices, time points belonging to a given stimuli class (*e.g.* InfOwnPos, EmoLookNeg, ...) were serially concatenated to provide a composite timecourse of all neural contributions from the stimuli and a Pearson’s correlation calculated across regions to produce a connectivity graph. An example of one such matrix is highlighted in the accompanying figure. These matrices were then subjected to several different clustering approaches to examine which if any are suitable for use in brain clustering analyses.

K-Means Clustering

This project examined 3 alternate approaches to k-means clustering. The first two are represented in the literature and the third represented a novel approach. It should be noted that all three approaches calculate distance in terms of L_2 -norms, but the data is transformed into different domains to explore the informativeness of both local and global properties of the brain. All stimuli matrices

Clustering on Raw Correlation Matrices

First clustering performance was assessed in the untouched correlation matrices. This is perhaps one of the most well represented techniques in dynamic connectivity. Although we had *a priori* reasons for suspecting certain numbers of cluster to emerge aligning with our main stimuli grouping (valence, own/other, and fixation). We chose a naïve approach to cluster selection and examined the within-cluster (WC) distance of a range of cluster numbers in order to determine the optimal number via the elbow method. The WC

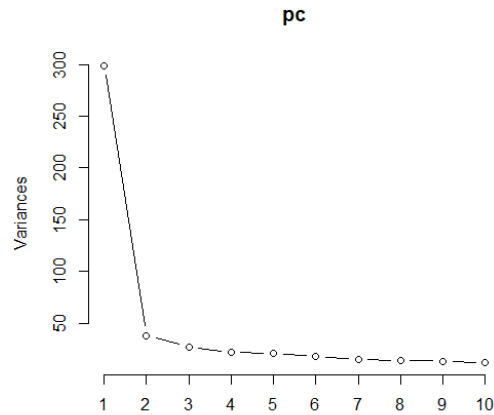


Figure 9: Amount of variance explained by top principle components

distance was determined under 50 random restarts in order to protect again the effects of initial cluster assignment and the final cluster under 2,000 restarts.

```
k.max = 15
wss <- sapply(1:k.max,function(k){kmeans(corr.array, k, nstart=50,iter.max = 30 )$tot.withinss})
plot(2:k.max, wss[2:15], type = "b", pch = 19)
```

```
%{r wcDistance, echo=FALSE, out.width="400px", fig.align="center", fig.cap="Within-Cluster  
Distance of different amounts of %clusters"} %knitr::include_graphics("./images/Heatplot  
for Correlation.png") %
```

There did not seem to be a well defined elbow (similar to the plot displayed for PCA) but beyond 4 clusters, there was seeming linear trend, so 4 was chosen by consensus. When examining the confusion matrix, it was clear that there was some limited success in accurately sorting connectivity matrices, but it largely preformed quite poorly.

Clustering on Principle Components

Because connectivity matrices represent the information from 116 anatomical regions, it represent extremely high dimensional data. Dimensions that don't provide much meaningful variance may be impeding accurate classification. To test this hypothesis, we applied principle component analysis to each matrix and extracted the 3 largest principle components to perform clustering on. See the following figures for a summary of this process.

As this PC variance plot shows, the first component carries the largest amount of variance. This large weight on the first component may be driven by collinearity among regions which is to be expected to some degree. Still, some variance was captured in smaller components. For ease of visualization 3 components were chosen for clustering variables.

Again, just as in the untouched correlation matrices, there was no clear elbow to guide choice of numbers of clusters but it appeared that 4 would be a reasonable choice.

Thus we executed the following:

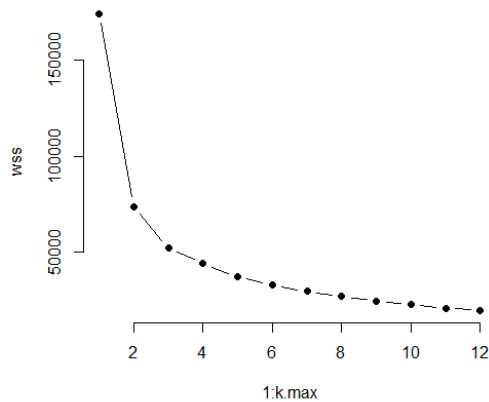


Figure 10: Elbow plot of PC clustering

```
#PCA and Kmean. Data should be the combined corr. matrix.
pc = prcomp(sapply(data[, -c(1:2)], as.numeric), center= TRUE)
```

Plots of the cluster assignment were dubious. Although they seemed to make clean divisions along the PC dimensions, there did not appear to be any natural grouping of the data to drive this results and it was suspected that large amounts of this may be spurious.

However, on closer inspection, there did appear to be some slight natural grouping of the data in terms of cluster assignment seemingly falling along rough heteroskedastic levels along the first principle component. This was very interesting because it was suggestive of the thought that variability in connectivity might have an important role in distinguishing brain states.

But the clean cut separation of connectivity states translated only marginally in terms of it's accuracy in representing mental states. There appeared to be a high amount representation of InfOtherNeg in cluster 1, cluster 2 had no clearcut grouping, cluster 3 perhaps had a slight over representation of passive viewing, and again cluster 4 had no clear groupings. This leaves room for much susicion in the inferences neuroscientists are currently drawing on these techniques.

Clustering on Graph Metrics

The preceeding clustering approaches relied on local differences in neural coordination. However, it is well known that the brain exhibits global properties that are important for cognition (32). Therefore we attempted to cluster matrices on two widely used graph metrics in neuroscience, rich club structure which provides a measure of the number of highly connected nodes and the mean path length which is related to communication efficiency. Graphs were constructed from thresholding of the correlation matrices setting all correlation coefficients less than $\rho = .85$ to 0. For visualization, network shown below have been thresholded at $\rho < .93$. Examples of the resultant graphs are provided.

Rich club coefficients and mean path lengths were calculated as follows:

```
library(brainGraph)
library(igraph)
richclub_coef = numeric(432)
mean_dist = numeric(432)
for (i in 1:length(x)) {
```

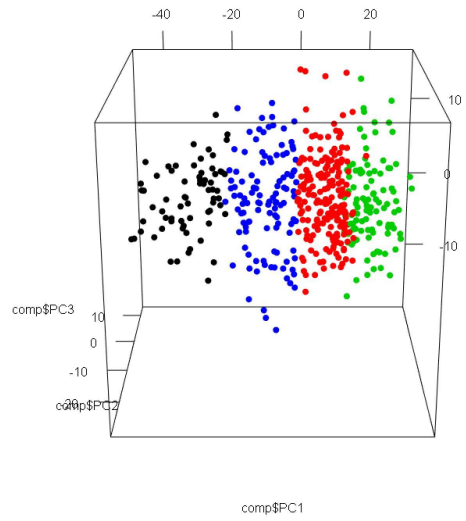


Figure 11: Plots of Principle Component and Cluster Assignment

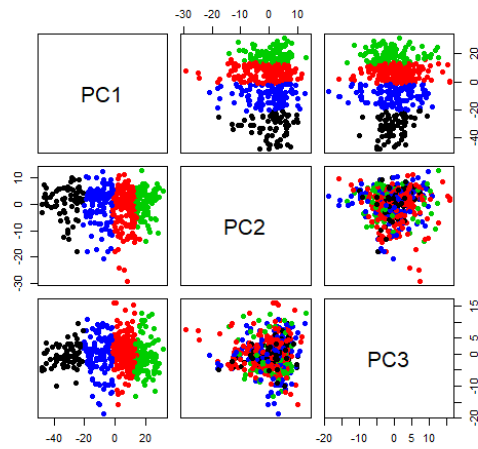


Figure 12: Plots of Principle Component Values

	1	2	3	4
EmoLabelNeg	6	22	9	11
EmoLabelPos	4	24	12	8
EmoLookNeg	7	18	13	10
EmoLookPos	7	16	13	12
fixation	6	22	6	14
InfoOtherNeg	12	16	10	10
InfoOtherPos	7	13	18	10
InfoOwnNeg	8	17	12	11
InfoOwnPos	8	21	4	15
ISI	4	26	5	13

Figure 13: Confusion Matrix PC Components

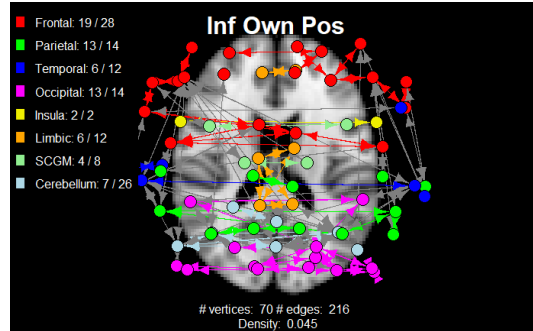


Figure 14: Network generated for InfOwnPos connectivity

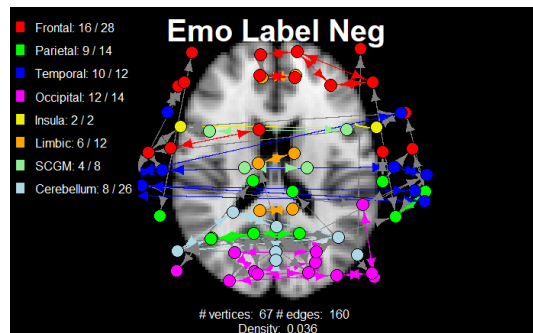


Figure 15: Network generated for EmoLabelNeg connectivity

```
file = read.csv(x[i])
cor_mat = as.matrix(file)
data = cor_mat[, -1]
rownames(data) = data[, 1]
for (j in 1:116) {
  for (k in 1:116) {
    if(colnames(data)[j] == fin$Name[k])
      colnames(data)[j] = fin$Code[k]
  }
}
rownames(data) = colnames(data)
cor_g <- graph_from_adjacency_matrix(data, mode='directed',
                                     weighted = 'correlation')
cor_edge_list <- igraph::as_data_frame(cor_g, 'edges')
only_sig <- cor_edge_list[abs(cor_edge_list$correlation) > .85 &
                          abs(cor_edge_list$correlation) < 1, ]
new_g <- graph_from_data_frame(only_sig, T)
bg = make_brainGraph(new_g, atlas = "aal116")
richclub_coef[i] = richclub_coef(bg)$phi
mean_dist[i] = mean_distance(bg)
}

results = data.frame("Data" = x, "Rich_Club_Coefficient" = richclub_coef,
                    "Avg_path_length" = mean_dist)
```

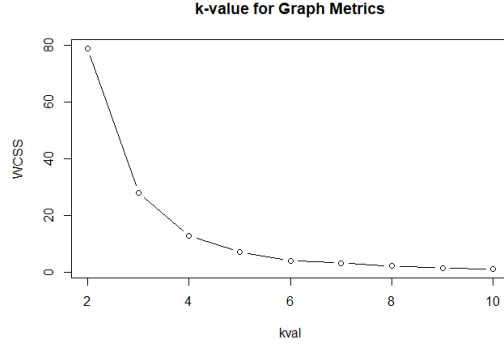


Figure 16: Within Cluster distances for graph poperties clustering

cluster	EmLabelNeg	EmLabelPos	EmLookNeg	EmLookPos	fixation	InfoOtherNeg	InfoOtherPos	InfoOwnNeg	InfoOwnPos
1	4	9	5	7	3	8	3	6	5
2	15	12	20	20	17	23	12	17	18
3	10	12	8	11	15	15	19	9	9
4	19	15	15	9	13	12	14	16	16

Figure 17: Confusion matrix for graph poperties clustering

Within-cluster distances again indicated that 4 clusters may be an appropriate choice.

The graph metrics were treated as if they were $\langle x, y \rangle$ coordinates and subjected to k-means clustering.

A more wholistic measure of the network also performed poorly, although there does seem to be some differential splitting among clusters 2 and 3 along valence lines. No other category seemed to provide such diffiential splitting suggesting differences in emotional valence processing may show distinct patterns of connectivity which would be consistent with the broader literature of affective connectivity.

Although it was planned to make quantitative assessments of purity post-clustering, given the consistently poor clustering results across all approaches, this did not seem needed until more accurate approaches to brain network clustering are devised. Hence we deviate from our proposal to only provide qualitative assessment of clustering accuracy.

Alternate Classification Approaches

Additionally we decided to cluster our correlation matrix using SVM and Random Forest to see if it would yield any acceptable performance. Afterall, SVM and Random Forest seem to be heavily used algorithm to tackle classifications. Our initial result was utter failure with 88% Error rates when we used Random Forest with 100 trees and number of variables sampled at each split 5000. Initial results of SVM seemed little better than chance. We decided not to troubleshoot further due to lack of time and the computational demands of these approaches on such a large dataset.

Additionally we decided to cluster our correlation matrix using Random Forest to see if it would yield any acceptable performance. Afterall, Random Forest seems to be heavily used to tackle classification problems. Our initial result was an utter failure with when we used Random Forest with 100 trees and number of variables sampled at each split 5000. The error rate of 88% was calculated by the in-built algorithm. We hypothesize that the failure to classify our correlation matrix into 10 categories is caused by lack of sufficient variance within individual correlation. This suspicion seem warranted when we observe the graph of Within Sum Squared (WSS) values of K-Mean clustering, the decrease in WSS was empirically negligible for clusters considered beyond four clusters. Therefore, it seemed reasonable to suspect the failure in attempts to cluster the correlation matrix into 10 groups would fail since additional clusters (after 4th) would not capture sufficient variance. We decided not to troubleshoot further due to lack of time.

```
Call:
  randomForest(x = x_val, y = y_val, ntree = 100, mtry = 5000)
  Type of random forest: classification
    Number of trees: 100
No. of variables tried at each split: 5000

  OOB estimate of  error rate: 88.54%
Confusion matrix:
```

	EmoLabelNeg	EmoLabelPos	EmoLookNeg	EmoLookPos	fixation	InfOtherNeg	InfOtherPos
EmoLabelNeg	2	12	8	7	2	1	3
EmoLabelPos	9	4	6	7	1	1	3
EmoLookNeg	5	5	9	8	3	2	1
EmoLookPos	7	3	9	3	0	7	1
fixation	0	2	2	2	10	9	7
InfOtherNeg	4	0	0	2	8	2	9
InfOtherPos	2	1	1	1	9	13	5
InfOwnNeg	3	3	3	4	3	5	10
InfOwnPos	2	1	1	1	16	8	7
ISI	14	6	6	7	0	1	1

	InfOwnNeg	InfOwnPos	ISI	class.error
EmoLabelNeg	2	2	9	0.9583333
EmoLabelPos	1	1	15	0.9166667
EmoLookNeg	1	1	13	0.8125000
EmoLookPos	3	1	14	0.9375000
fixation	2	13	1	0.7916667
InfOtherNeg	7	15	1	0.9583333
InfOtherPos	4	9	3	0.8958333
InfOwnNeg	4	6	7	0.9166667
InfOwnPos	5	5	2	0.8958333
ISI	0	2	11	0.7708333

Figure 18: Results of Random Forest Clustering

Random Forest

The results of random forest classification are as follows:

```
data=read.csv("corr.matrix.combined.csv", header = TRUE)

head(data)
x_val=data[,-1]
y_val=data[,1]
library(randomForest)

model_rf=randomForest(x_val,y_val,mtry=5000, ntree = 100)
```

Support Vector Machines

SVM was calculated in the following manner:

```
library(caret)
ctrl = trainControl(method = "cv", savePred=T, classProb=T)
SVM_mod = train(y_val~as.matrix(x_val), method = "svmLinear", trControl = ctrl)
```

Discussion

Although no method performed particularly well, there were some winners and losers. SVM and Random Forests should not be used to classify imaging data as they performed by far the worst. PCA decomposition, although not stellar, provided slightly better results. Graph theoretic measures did not perform substantially

better but it is entirely plausible that the particular graph metrics were not well suited and therefore we recommend that this area receive more exploration to assess its importance.

There were several limitations to this particular dataset that impede generalization more broadly. The subjects in this cohort had extremely high levels of head motion. This is a well known cause of noise in fMRI data (33) and although efforts were made to screen for unacceptable levels of head motion, there still were elevated levels which may have impeded clustering. Although the immediate implication is that these analyses should be conducted in a cleaner dataset, it may also serve to highlight the role of noise in dynamic connectivity and serve as a reminder to the field to make all efforts to minimize other sources of noise such as heartbeat and respiration. Another limitation is our deviation from more standard dynamic connectivity pipelines in disrupting the temporal order of our stimuli matrices, but this was necessitated by the data and not a point that we believe detracts substantially from the interpretability of results.

Until the point however that more conclusive studies are conducted there should be a strong suspicion kept around drawing conclusions about cognition from the results of k-means or any other clustering technique. To our knowledge, this is the first study of its kind to take a systematic approach to testing the relationship between clustered transient brain connectivity and inferred mental contents. On the basis of our work the assumption that such clusters represent unitary cognitive states is untenable.

It is the hope of the authors of this class project that the presumed relationship between mind states and brain states is viewed with a weary eye.

Appendix A.

Philosophical Assumptions

It is imperative to establish the core philosophical understanding from which this data is being approached from. The core tenants to follow are neither radical nor unreasonable, but are too often left implicit in discussions of brain science. If however we accept these tenants to be true, then the analysis we present follows quite naturally.

Assumption #1 – Brain states represent cognitive states

This is the idea that when we discuss a particular configuration of correlation among brain regions, it reflects a particular internal representation experienced by the individual. Although the correctness of this assumption may be challenged in some extreme cases (*e.g.* persistent vegetative states), this is the driving ontological stance of neuroscience for typically functioning human beings. If we did not believe the brain was representative of the mind, there would be little point to studying its function. This is not to say that all brain function maps well onto mind function and in fact there is evidence that much of it may be attributable to basal metabolic functions (34). Yet despite this significant confound, it is assumed that as long as a human is functioning in a state that we would reasonably believe they possess a mind, the network configuration of their brain represents some element of the brain-to-mind mapping.

Assumption #2 – Different brain states represent different cognitive states (and same represents same)

Again this is an assumption that cuts to the core of neuroscience but an essential one to articulate. It perhaps can best be summarized by the statement that if we observe “Brain Connectivity A” at time 1 then (according to Assumption 1) we can ascribe to it the “Mental Contents A”. If we then later observe “Brain Connectivity A” at time 2, then we can also ascribe “Mental Contents A” to the individual at this later time. The converse is also assumed true that if two brain states differ then so do their mental contents. This is not to say that there is absolutely no overlap in either brain states or cognitive states, but that if we can

ascribe different mental states to individuals then we can naturally expect their brain connectivity to differ across these mental states. We cannot neglect some of the differences to be attributable to differences in the metabolic considerations mentioned above. But for the relatively homogenous environment of an fMRI scanner, this will be, rightly or wrongly, assumed negligible in comparison to cognition induced shifts.

Assumption #3 – Different environmental inputs induce different brain states

This is the idea that the brain must act adaptively to its environment such that if two environmental stimuli are sufficiently divergent on some salient property, then they will induce different configurations in the brain (and hence different mental contents by way of assumption 2). This assumption is left deliberately vague on what constitute different stimuli, but given the abundance of neuroimaging literature observing different activity patterns in brain states across different tasks, this is no stretch of the imagination.

References

- [1] J. R. Binder, J. A. Frost, T. A. Hammeke, P. Bellgowan, S. M. Rao, and R. W. Cox, “Conceptual processing during the conscious resting state: A functional mri study,” *Journal of cognitive neuroscience*, vol. 11, no. 1, pp. 80–93, 1999.
- [2] R. Buckner, J. Andrews-Hanna, and D. Schacter, “The brain’s default network: Anatomy, function, and relevance to disease,” *Annals of the New York Academy of Sciences*, vol. 1124, no. 1, pp. 1–38, 2008.
- [3] K. Christoff, Z. Irving, K. Fox, R. Spreng, and J. Andrews-Hanna, “Mind-wandering as spontaneous thought: A dynamic framework,” *Nature Reviews Neuroscience*, vol. 17, no. 11, pp. 718–731, 2016.
- [4] M. Mason, M. Norton, J. Van Horn, D. Wegner, S. Grafton, and C. Macrae, “Wandering minds: The default network and stimulus-independent thought,” *Science*, vol. 315, no. 5810, pp. 393–395, 2007.
- [5] J. Smallwood and J. Schooler, “The science of mind wandering: Empirically navigating the stream of consciousness,” *Annual review of psychology*, vol. 66, pp. 487–518, 2015.
- [6] K. Fox, R. Spreng, M. Ellamil, J. Andrews-Hanna, and K. Christoff, “The wandering brain: Meta-analysis of functional neuroimaging studies of mind-wandering and related spontaneous thought processes,” *Neuroimage*, vol. 111, pp. 611–621, 2015.
- [7] J. D. Teasdale, “Negative thinking in depression: Cause, effect, or reciprocal relationship?” *Advances in Behaviour Research and Therapy*, vol. 5, no. 1, pp. 3–25, 1983.
- [8] E. Burdwood, Z. Infantolino, L. Crocker, J. Spielberg, M. Banich, G. Miller, and W. Heller, “Resting-state functional connectivity differentiates anxious apprehension and anxious arousal,” *Psychophysiology*, vol. 53, no. 10, pp. 1451–1459, 2016.
- [9] S. Nolen-Hoeksema, “Responses to depression and their effects on the duration of depressive episodes,” *Journal of abnormal psychology*, vol. 100, no. 4, pp. 569–582, 1991.
- [10] S. Nolen-Hoeksema, “The role of rumination in depressive disorders and mixed anxiety/depressive symptoms,” *Journal of abnormal psychology*, vol. 109, no. 3, pp. 504–511, 2000.
- [11] A. Kucyi and K. D. Davis, “Dynamic functional connectivity of the default mode network tracks day-dreaming,” *Neuroimage*, vol. 100, pp. 471–480, 2014.
- [12] R. M. Hutchison, T. Womelsdorf, E. A. Allen, P. A. Bandettini, V. D. Calhoun, M. Corbetta, S. Della Penna, J. H. Duyn, G. H. Glover, J. Gonzalez-Castillo, and others, “Dynamic functional connectivity: Promise, issues, and interpretations,” *Neuroimage*, vol. 80, pp. 360–378, 2013.
- [13] E. A. Allen, E. Damaraju, S. M. Plis, E. B. Erhardt, T. Eichele, and V. D. Calhoun, “Tracking whole-brain connectivity dynamics in the resting state,” *Cerebral cortex*, vol. 24, no. 3, pp. 663–676, 2014.
- [14] B. W. Mooneyham, M. D. Mrazek, A. J. Mrazek, K. L. Mrazek, D. T. Phillips, and J. W. Schooler, “States of mind: Characterizing the neural bases of focus and mind-wandering through dynamic functional

- connectivity,” *Journal of cognitive neuroscience*, vol. 29, no. 3, pp. 495–506, 2017.
- [15] R. E. Beaty, Q. Chen, A. P. Christensen, J. Qiu, P. J. Silvia, and D. L. Schacter, “Brain networks of the imaginative mind: Dynamic functional connectivity of default and cognitive control networks relates to openness to experience,” *Human brain mapping*, vol. 39, no. 2, pp. 811–821, 2018.
- [16] J. Jovicich, S. Czanner, D. Greve, E. Haley, A. van Der Kouwe, R. Gollub, D. Kennedy, F. Schmitt, G. Brown, J. MacFall, and others, “Reliability in multi-site structural mri studies: Effects of gradient non-linearity correction on phantom and human data,” *Neuroimage*, vol. 30, no. 2, pp. 436–443, 2006.
- [17] H. K. Laurent, D. Wright, and M. Finnegan, “Mindfulness-related differences in neural response to own infant negative versus positive emotion contexts,” *Developmental cognitive neuroscience*, vol. 30, pp. 70–76, 2018.
- [18] U. Moscardino and G. Axia, “Infants’ responses to arm restraint at 2 and 6 months: A longitudinal study,” *Infant Behavior and Development*, vol. 29, no. 1, pp. 59–69, 2006.
- [19] N. Tottenham, A. Borscheid, K. Ellertsen, D. Marcus, and C. Nelson, “Categorization of facial expressions in children and adults: Establishing a larger stimulus set,” in *Journal of cognitive neuroscience*, 2002, pp. 74–74.
- [20] J. Y. Chiao, T. Iidaka, H. L. Gordon, J. Nogawa, M. Bar, E. Aminoff, N. Sadato, and N. Ambady, “Cultural specificity in amygdala response to fear faces,” *Journal of Cognitive Neuroscience*, vol. 20, no. 12, pp. 2167–2174, 2008.
- [21] J. Ashburner, G. Barnes, C. Chen, J. Daunizeau, G. Flandin, K. Friston, S. Kiebel, J. Kilner, V. Litvak, R. Moran, and others, “SPM12 manual,” *Wellcome Trust Centre for Neuroimaging, London, UK*, 2014.
- [22] A. Eklund, T. E. Nichols, and H. Knutsson, “Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates,” *Proceedings of the National Academy of Sciences*, p. 201602413, 2016.
- [23] N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, and M. Joliot, “Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain,” *Neuroimage*, vol. 15, no. 1, pp. 273–289, 2002.
- [24] M. Brett, J.-L. Anton, R. Valabregue, J.-B. Poline, and others, “Region of interest analysis using an spm toolbox,” in *8th international conference on functional mapping of the human brain*, 2002, vol. 16, p. 497.
- [25] T. Pearlstein, M. Howard, A. Salisbury, and C. Zlotnick, “Postpartum depression,” *American journal of obstetrics and gynecology*, vol. 200, no. 4, pp. 357–364, 2009.
- [26] M. C. Lovejoy, P. A. Graczyk, E. O’Hare, and G. Neuman, “Maternal depression and parenting behavior: A meta-analytic review,” *Clinical psychology review*, vol. 20, no. 5, pp. 561–592, 2000.
- [27] A. Stein, R. M. Pearson, S. H. Goodman, E. Rapa, A. Rahman, M. McCallum, L. M. Howard, and C. M. Pariante, “Effects of perinatal mental disorders on the fetus and child,” *The Lancet*, vol. 384, no. 9956, pp. 1800–1819, 2014.
- [28] B. Gelaye, M. B. Rondon, R. Araya, and M. A. Williams, “Epidemiology of maternal depression, risk factors, and child outcomes in low-income and middle-income countries,” *The Lancet Psychiatry*, vol. 3, no. 10, pp. 973–982, 2016.
- [29] K. L. Wisner, D. K. Sit, M. C. McShea, D. M. Rizzo, R. A. Zoretich, C. L. Hughes, H. F. Eng, J. F. Luther, S. R. Wisniewski, M. L. Costantino, and others, “Onset timing, thoughts of self-harm, and diagnoses in postpartum women with screen-positive depression findings,” *JAMA psychiatry*, vol. 70, no. 5, pp. 490–498, 2013.
- [30] J. B. Henriques and R. J. Davidson, “Regional brain electrical asymmetries discriminate between previously depressed and healthy control subjects,” *Journal of abnormal psychology*, vol. 99, no. 1, p. 22, 1990.
- [31] H. A. Demaree, D. E. Everhart, E. A. Youngstrom, and D. W. Harrison, “Brain lateralization of emotional processing: Historical roots and a future incorporating ‘dominance’,” *Behavioral and cognitive neuroscience*

reviews, vol. 4, no. 1, pp. 3–20, 2005.

[32] M. A. Bertolero, B. T. Yeo, and M. D’Esposito, “The modular and integrative functional architecture of the human brain,” *Proceedings of the National Academy of Sciences*, vol. 112, no. 49, pp. E6798–E6807, 2015.

[33] K. R. Van Dijk, M. R. Sabuncu, and R. L. Buckner, “The influence of head motion on intrinsic functional connectivity mri,” *Neuroimage*, vol. 59, no. 1, pp. 431–438, 2012.

[34] M. E. Raichle and A. Z. Snyder, “A default mode of brain function: A brief history of an evolving idea,” *Neuroimage*, vol. 37, no. 4, pp. 1083–1090, 2007.