

# Event Perception of Schema-Rich and Schema-Poor Video Sequences During fMRI Scanning: Top Down Versus Bottom Up Processing

ADI ZAIMI, CATHERINE HANSON, STEPHEN JOSÉ HANSON

Rutgers University, Newark, NJ 07102, USA



## Abstract

We use fMRI to compare brain activation for video clips of different levels of complexity: eight subjects watched schema-rich or schema-poor videos while making event boundary decisions. Brain activity was recorded during the event parsing task. We find different patterns of activity when subjects parsed schema-rich stimuli compared to when they parsed schema-poor stimuli. Schema-rich stimuli induced stronger activation in areas associated with planning, visual imagery and episodic memory as well as in areas associated with attention and low level sensory and motor functions (putamen, globus pallidus). We applied structural equation modeling to the brain activity covariances of the common ROIs and found distinctive patterns in the regional networks between schema-rich and schema-poor stimuli which may characterize differences in processing strategy.

## Introduction

In this study, we examine the brain correlates of “top-down” and “bottom-up” processing in an event perception study. Event perception tasks are ideal to study the dynamics and interactions of perception, memory, attention, schema, planning and decision making. Subjects watched scripted videos of people engaging in everyday activities, indicating whenever they saw an event change in these videos. We showed subjects stimuli with different level of complexity to elicit different networks depending on the top-down or bottom-up nature of the stimulus. We hypothesized that we would detect separate neural networks involved in parsing top-down versus bottom-up stimuli.

## Previous fMRI evidence

Zacks et. al (2001) is the first study to apply fMRI to event perception. They investigated brain areas active during the event changes for active and passive viewing and found stronger activation for important event changes compared to less important ones. For passive viewing they found activity in the following areas:

Posterior Inferior Sulcus (or V5), Precentral Sulcus (FEF), Left Fusiform Gyrus (BA 37), Precuneus (right BA 31 and left BA 19), Middle Occipital Gyrus (BA 18), Superior Temporal Gyrus (BA 22), and in Precentral Gyrus (BA 9).

In the active viewing condition Zacks et al. (personal communications, Zacks 2003) found a series of areas not present in the passive condition, such as:

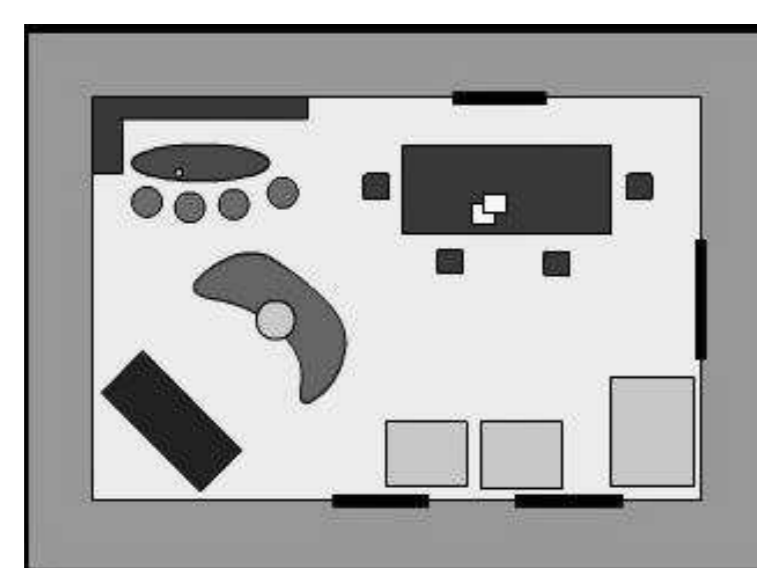
Middle Temporal Gyrus (BA 37), Medial and Middle Frontal Gyrus (BA 6, 9, and 37), Inferior Occipital Gyrus (BA 18), Cuneus (BA 17, 18 and 19), Lingual Gyrus (BA 18), Postcentral Gyrus (BA 2, 3 and 40), Insula (BA 13), Lentiform Nucleus (Putamen), Posterior Cingulate (BA 23 and 30) and Cingulate Gyrus (BA 31), Inferior Frontal Gyrus (BA 11 and 45), Parahippocampal Gyrus (BA 28 and 36), and in Anterior Cingulate (BA 32).

## Schema-Rich and Schema-Poor Stimuli

We showed subjects two kind of stimuli: (a) schema-rich stimulus (*study* video): highly familiar situation, lots of expectations and associations, more top-down driven; (b) schema-poor stimulus (*house* video): less familiar situation, more stimulus or bottom-up driven, weaker expectations (Heider and Simmel, 1944).



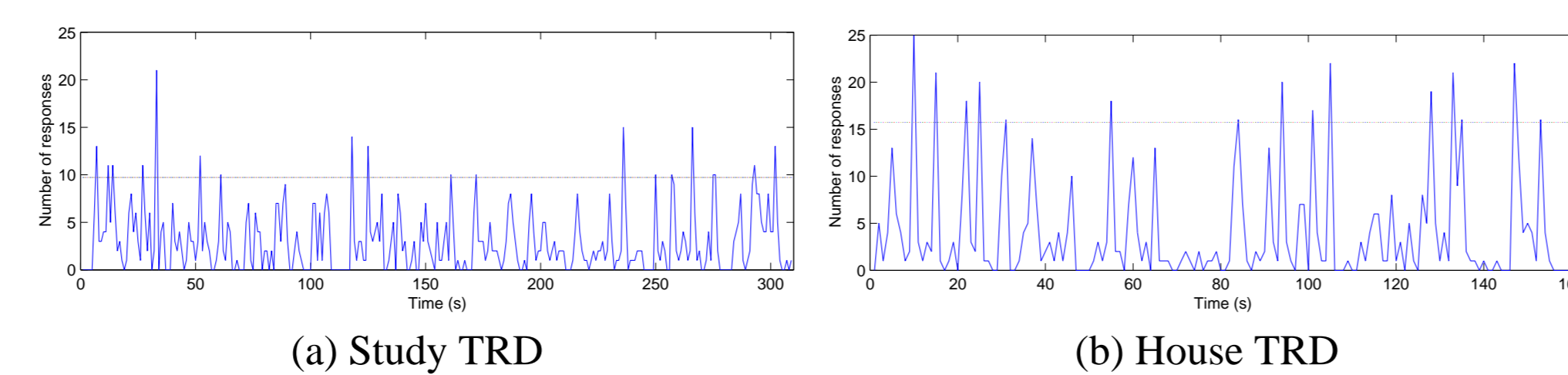
(a) Schema-rich — study



(b) Schema-poor — house

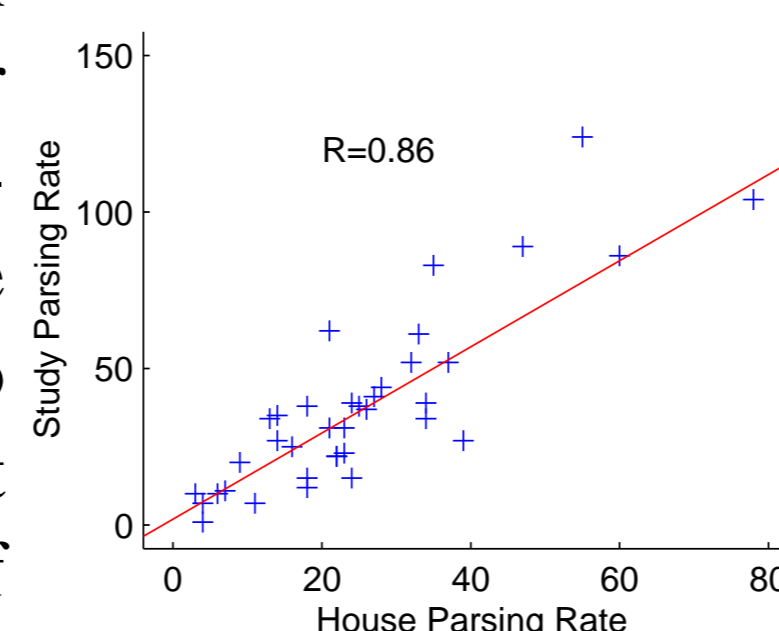
## Behavioral Responses

We collected event change judgements (using buttonpresses) from 38 subjects and counted for each video the momentary rate of parsing from all subjects in 1 second-bins constructing the Temporal Response Density (TRD) functions. The horizontal line in the figure denotes two standard deviations above the mean response, and an excursion above such line will be called a *significant event change*. We used the significant event changes as onsets (subject induced) to design the GLM analysis.



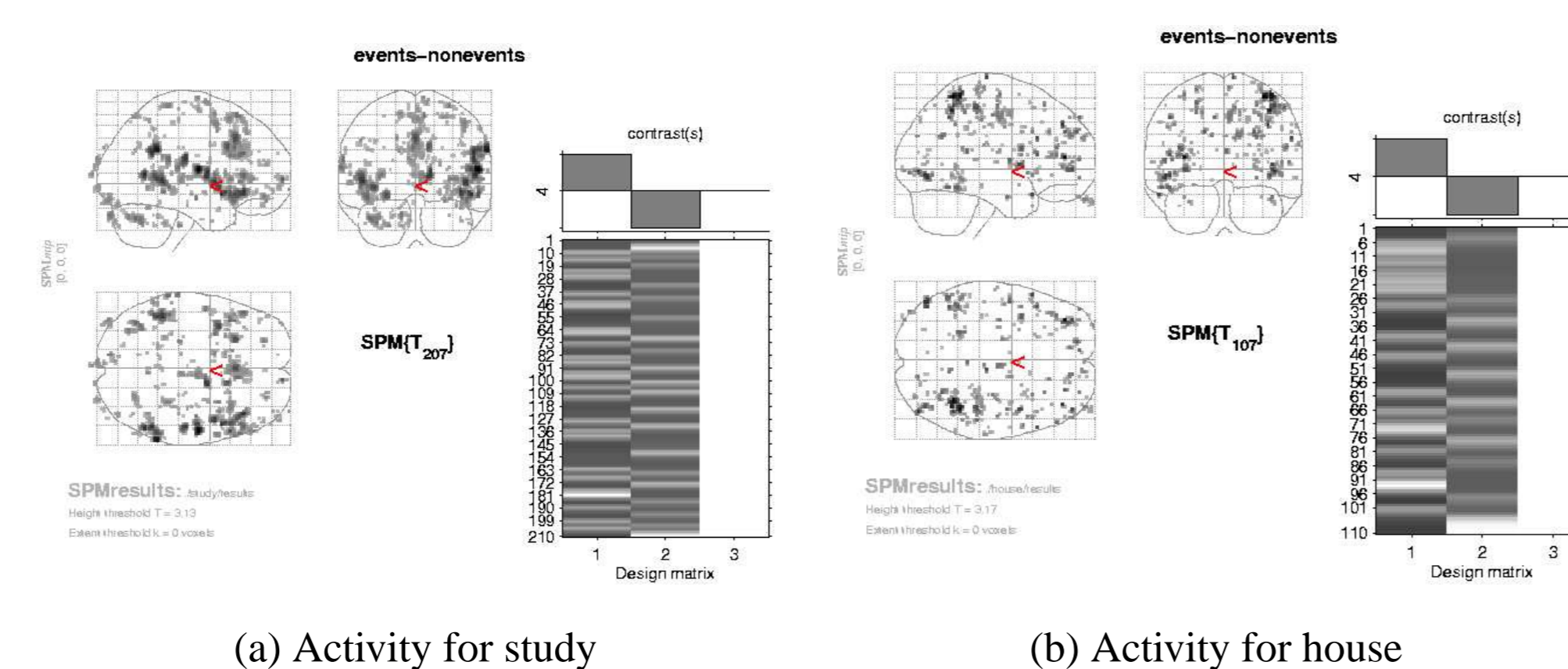
## Parsing Rates

There was a high correlation between the parsing rates of the subjects for the house and study stimuli, suggesting that subjects maintained the same level of parsing rates across the two stimuli. Subjects have an inherent parsing rate which is independent of video type.



## fMRI Analysis

We scanned 8 subjects who were making event boundary judgements on the study and house videos. We defined two conditions: (a) events with subjects' own responses as onsets, and (b) non-events defined as the background. The GLM analysis compared the event versus the non-events conditions:

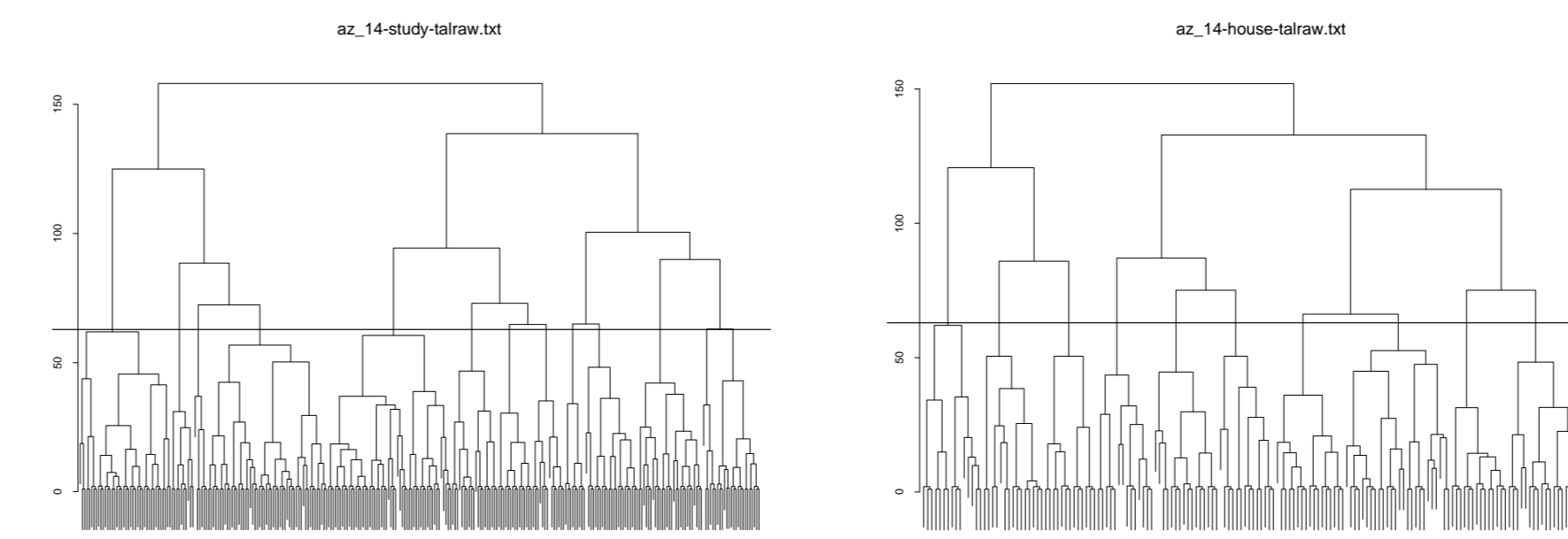


We found significant activity ( $p < 0.001$ ) in areas congruent with those of Zacks et al. (2001), such as:

Inferior Temporal Gyrus, Middle Occipital Gyrus, Superior Parietal Lobule, Caudate, Cingulate Gyrus, Claustrum, Lentiform Nucleus, Middle Temporal Gyrus, Cuneus, Insula, Medial Frontal Gyrus, Precuneus, Thalamus, Inferior Frontal Gyrus, Inferior Parietal Lobule, Middle Frontal Gyrus, Postcentral Gyrus, Precentral Gyrus, Superior Frontal Gyrus, and Superior Temporal Gyrus.

## Cluster Analysis and ROI Selection

From the activation t-maps, we submitted the gray area voxels above threshold ( $p < 0.001$ ) to hierarchical clustering of voxel euclidean distances. We examined the scree plots (plots of merge history as a function of distance merged) for each dendrogram, and chose a cut point (60 mm—the maximal distance between cluster leafs) across all subjects. This cut point maximized the distance between clusters and minimized the distance within clusters, and corresponded to the point of maximal change in most of the scree plots, indicating optimal cluster density.



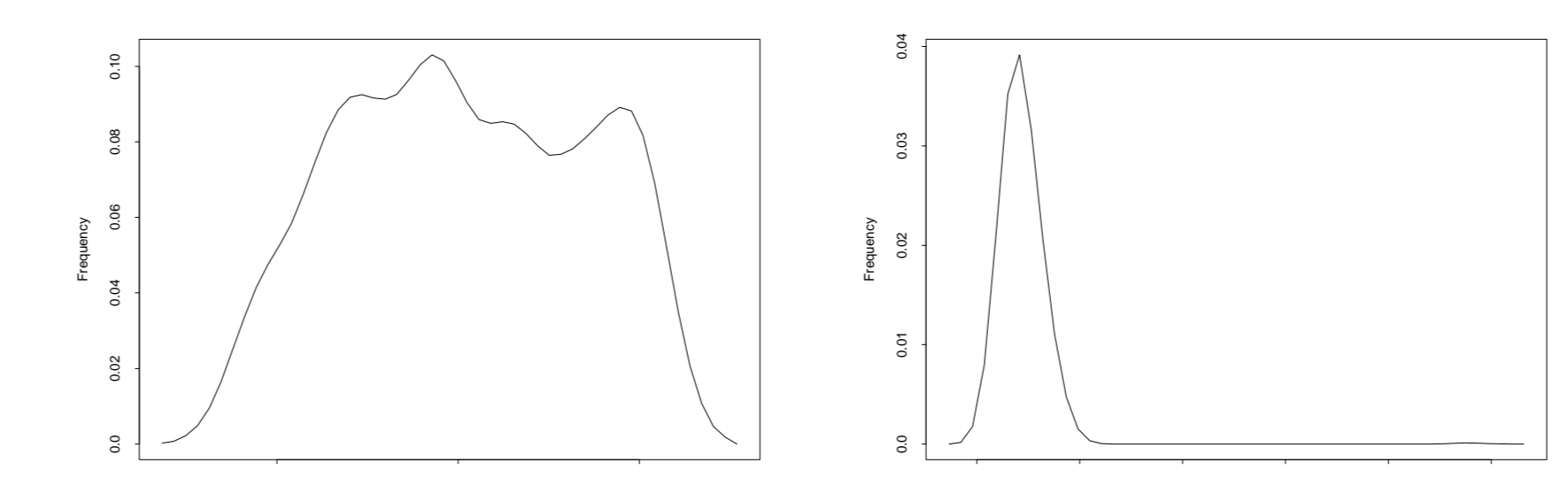
The analysis produced between 10-20 clusters per subject. We computed the centroid of each cluster as the mean of the ( $x, y, z$ ) coordinates of its component voxels and used the Talairach Atlas to identify the area for each centroid location. The areas in common to all subjects for the study and the areas in common for the house video were:

Study	House
Right Middle Frontal Gyrus	Right Middle Frontal Gyrus
Right Inferior Frontal Gyrus	Right Inferior Frontal Gyrus
Right Precuneus	Left Cingulate Gyrus
Right Lentiform Nucleus	Left Inferior Parietal Lobule
Left Middle Frontal Gyrus	
Left Superior Temporal Gyrus	

These areas are responsible for cognitive functions such as attention, memory, spatial imagery, etc., involvement of which we would expect in an event perception task.

## Structural Equation Modeling

To study the neural interaction of areas, McIntosh (1999) did confirmatory analysis on predefined models by doing path analysis on the covariances of selected regions. A more general way to study the dynamics of the interactions between areas, although a computationally expensive one, is to perform an exhaustive search of the graph space. This can be done by fitting *all* possible graphs by Structural Equation Modeling and then selecting the best model or best set of edges (by voting). This strategy<sup>d</sup> can work only if the goodness of fit distribution is single peaked with a positive skew.



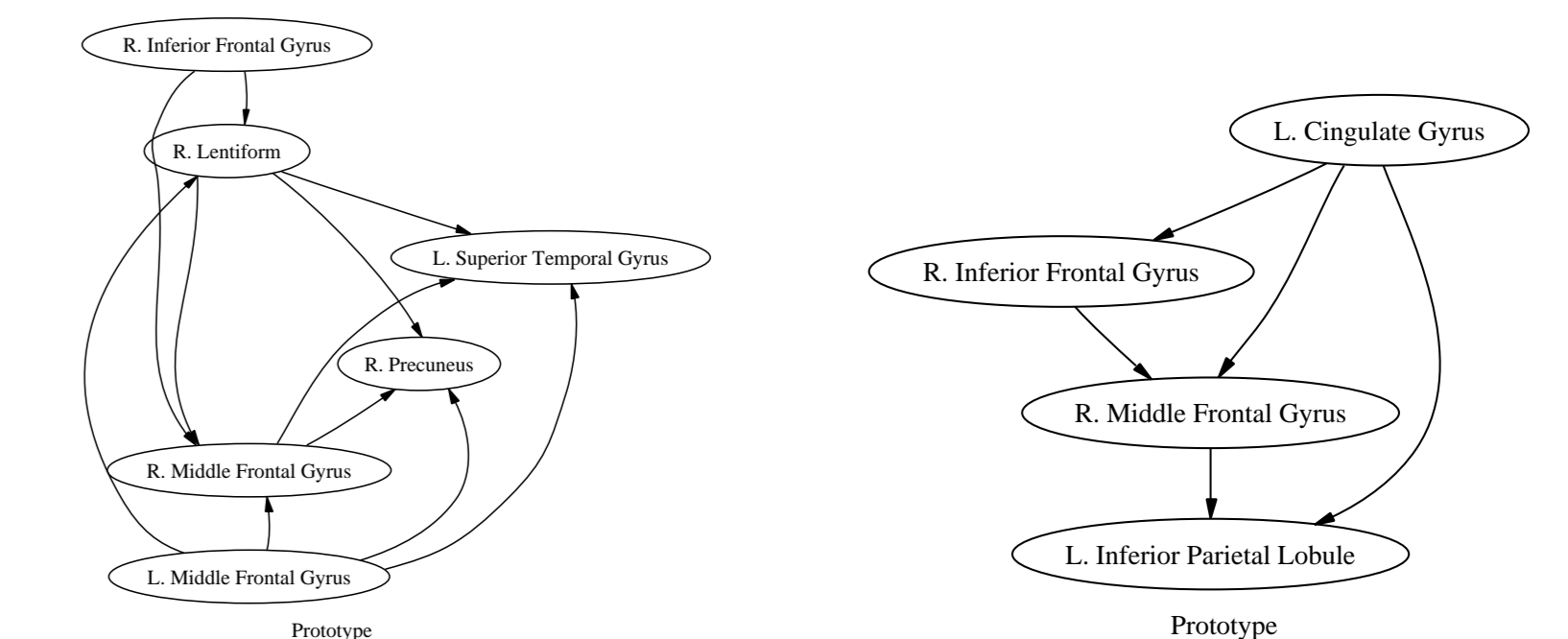
Distribution of AIC values: six node model (left), an eight node model (right)

## SEM Results

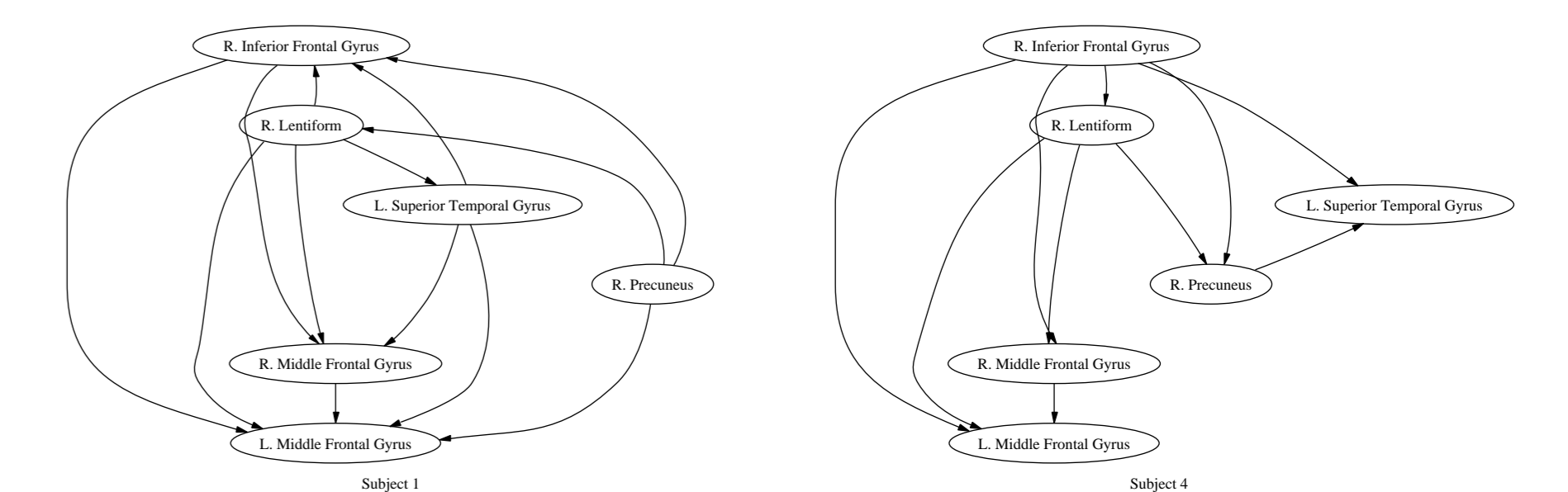
From the centroids of the areas common to all subjects, we extracted the voxel time series and computed their covariances (for each subject and each condition separately). We also produced an average covariance for each condition by taking the harmonic mean of the covariances from each subject. We fit the covariance matrices in SEM (using Lisrel) for all possible directed acyclic graphs.

<sup>d</sup>Although studying subject individuality may unveil the different networks that are present, for instance, in fast parses versus slow ones (different strategies), it would require to search networks with as many as 20 nodes (areas) in them. Analysing the full search space for such graphs is an intractable problem since the number of models increases factorially with the number of nodes, but we are currently investigating MCMC methods to sample the graph space of these networks to reduce the amount of SEM analysis.

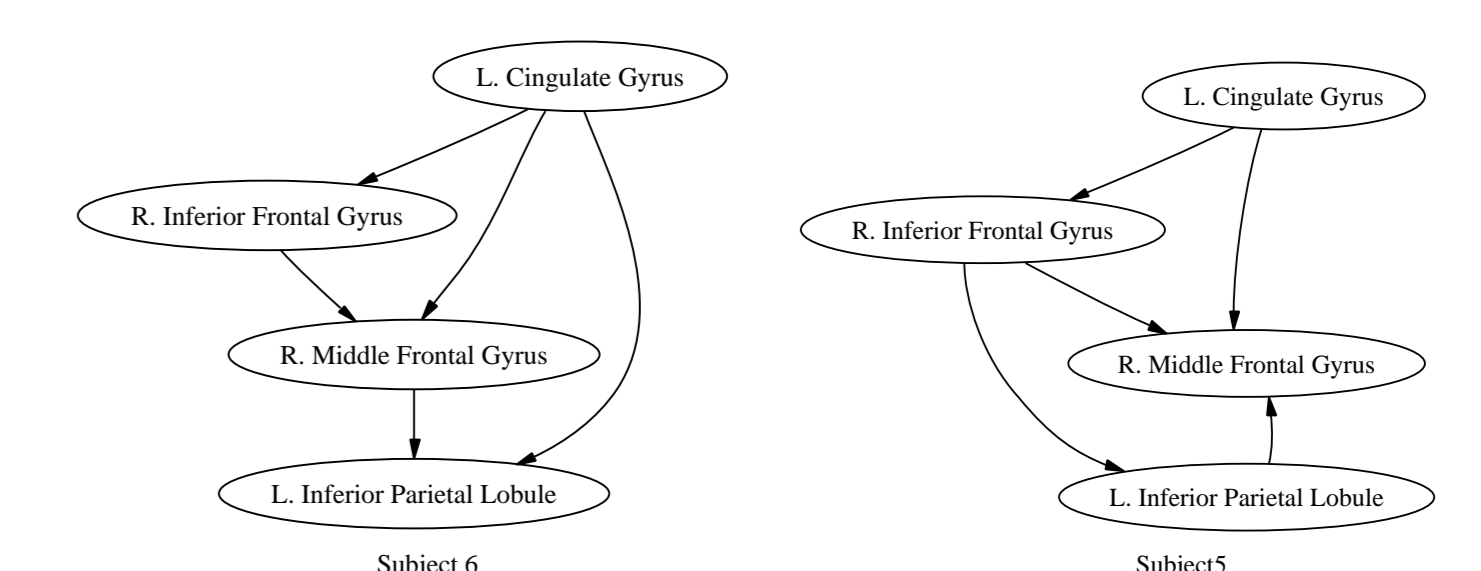
We investigated 24 models for house and 720 models for study and we selected the model with the lowest AIC (Akaike—a model selection criteria that trades bias vs. variance) value corresponding to the best fit model for a covariance matrix. We show the prototype graphs for study (left) and house (right):



For the study condition, the distribution of the Hamming distances of the prototype from the best AIC-value graphs of the four subjects is {4, 5, 7, 9}. Here we show the graphs most similar to the Prototype (left) and least similar to the prototype (right):



For the house condition, the distribution of the Hamming distances of the prototype from the best AIC-value graphs of the four subjects is {0, 0, 2, 3}. Here we show the graphs most similar to the Prototype (left) and least similar to the prototype (right):



## Discussion

The networks which we found share common brain areas with those from attentional networks (alerting system—right frontal), from the where pathway (inferior parietal). It is not surprising that event perception tasks involve attentional networks and other decision making networks. Furthermore, it is not surprising that we found two different networks: one typical to a schema-rich stimulus which is more guided by schemata and planning—a top-down stimulus, and another network typical to a schema-poor stimulus, which is more bottom-up controlled. We think that these different networks may represent different processing systems involving the recruiting of more top-down areas in the schema-rich stimulus, and fewer areas in the bottom-up stimulus; we have designed future experiments which will shed more light on this issue.

## References

Heider, F. and Simmel, M. (1944). An experimental study of apparent behaviour. *Am. J. of Psych.*, 57(2):243-59.  
 McIntosh, A. R. (1999). Mapping Cognition to the Brain Through Neural Interactions. *Memory*, 7(5/6), 523-48  
 Zacks, J., Braver, T. S., Sheridan, et al. (2001). Human brain activity time-locked to perceptual event boundaries. *Nature Neuroscience*, 4(6):651-655.

## Acknowledgements

Supported by NSF (EIA-0205178), from James S. McDonnell Foundation Thanks to Yaroslav Halchenko, Donovan Elford, and Toshihiko Matsuka