

# Effective Connectivity Modeling with the euSEM and GIMME

Benjamin Zinszer  
PhD candidate, Department of Psychology  
MAS candidate, Department of Statistics  
Center for Language Science  
Pennsylvania State University

*Supported by Guangwai Brain & Language Lab*



*Director:* Prof. Yangping Dong 董燕萍  
*Associate Director:* Dr. Jing Yang 杨静  
[www.gwball.com](http://www.gwball.com)

## Section I: Connectivity Models

The present tutorial will describe the extended unified SEM (euSEM: Gates, Molenaar, Hillary, & Slobounov, 2011), a model for estimating relationships between multiple regions of interest in neuroimaging data. Before addressing the practical steps for preparing and analyzing datasets, it is important to have some background on the statistical principles underlying the euSEM. First, I introduce structural equation modeling (SEM) and vector auto-regression (VAR), two models commonly used to estimate connectivity and integrated under the euSEM. Next I outline the euSEM model and statistical assumptions necessary for applying it to neuroimaging data. Finally I describe some of the statistics used in evaluating and comparing euSEM models. The practical steps for implementing these statistical models are available in the next section.

**Structural equation modeling**, or SEM, is a statistical tool for comparing causal models between multiple variables by estimating a set of linear regression equations. In the case of neuroimaging data, these variables represent regions of interest (ROIs) that are each measured as a time series. SEM allows researchers to propose a set of connections between the ROIs and estimate the regression equations describing these connections. The proposed set of connections constitutes the causal model (that is, a hypothesis about what regions cause changes in activity for other regions), and this causal model is compared to the observed data. The best fitting set of coefficients describing the weights of the proposed connections is estimated by solving the matrix form of these regression equations. The resulting covariance is compared to the covariance of the observed data to obtain a  $\chi^2$  (chi-squared) statistic, a test for the significance of the model (covered in more detail later).

To illustrate SEM conceptually, consider the hemodynamic activity measured in three ROIs  $X$ ,  $Y$ , and  $Z$  repeatedly measured at time point  $t$ . The relationship between these three regions can be described as a set of three linear regression equations:

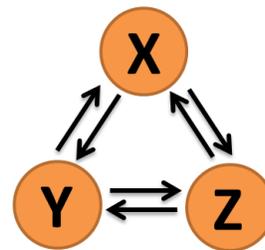
$$X_t = \beta_0 + \beta_1 Y_t + \beta_2 Z_t + \varepsilon$$

$$Y_t = \beta_3 + \beta_4 X_t + \beta_5 Z_t + \varepsilon$$

$$Z_t = \beta_6 + \beta_7 X_t + \beta_8 Y_t + \varepsilon$$

*Matrix notation:*

$$\begin{bmatrix} X_t \\ Y_t \\ Z_t \end{bmatrix} = \begin{bmatrix} \beta_{0X} \\ \beta_{0Y} \\ \beta_{0Z} \end{bmatrix} + \begin{bmatrix} 0 & \beta_{XY} & \beta_{XZ} \\ \beta_{YX} & 0 & \beta_{YZ} \\ \beta_{ZX} & \beta_{ZY} & 0 \end{bmatrix} \begin{bmatrix} X_t \\ Y_t \\ Z_t \end{bmatrix} + \begin{bmatrix} \varepsilon_X \\ \varepsilon_Y \\ \varepsilon_Z \end{bmatrix}$$



This model is the weakest (least specific) possible hypothesis. It assumes that every ROI has *some* causal influence on all of the other ROIs. A graphical representation of this model shows causal relationships (an arrow pointing from causal ROI to affected ROI) in every possible direction.

A more reasonable model might make use of some causal relationships that the researcher hypothesizes. For instance, perhaps activation is predicted to “flow” from X to Y to Z. This structural equation model would be described by the following regression equations and graphic:

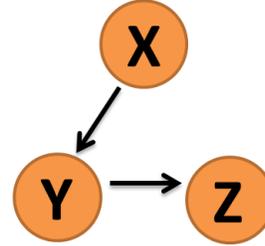
$$X_t = \beta_0 + \varepsilon$$

$$Y_t = \beta_3 + \beta_4 X_t + \varepsilon$$

$$Z_t = \beta_6 + \beta_8 Y_t + \varepsilon$$

*Matrix notation:*

$$\begin{bmatrix} X_t \\ Y_t \\ Z_t \end{bmatrix} = \begin{bmatrix} \beta_{0X} \\ \beta_{0Y} \\ \beta_{0Z} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ \beta_{YX} & 0 & 0 \\ 0 & \beta_{ZY} & 0 \end{bmatrix} \begin{bmatrix} X_t \\ Y_t \\ Z_t \end{bmatrix} + \begin{bmatrix} \varepsilon_X \\ \varepsilon_Y \\ \varepsilon_Z \end{bmatrix}$$



Estimating the  $\beta$  coefficients in these regression equations yields the estimated connectivity between each region. Specifically  $\beta_4$  is the influence of X on Y, and  $\beta_8$  is the influence of Y on Z. Real connectivity studies, however, are unlikely to begin with such specific models. In this case, model comparison is necessary to determine which model (from either a set of competing hypotheses or through a search process) best fits the observed data. There are a number of statistics useful for comparing models which will be discussed in a subsequent section.

One further benefit of SEM is the possibility of estimating exogenous influences on the ROIs. Exogenous variables are defined as those which may cause changes in other variables but are not themselves influenced by any other variable in the system. In the previous model, the ROI labeled A is technically an exogenous variable, but it is unconventional to think of brain measurements as exogenous because we want to maintain the possibility that A is influenced by some other region or by a truly exogenous variable.

In event-related neuroimaging studies, a typical exogenous variable would be a stimulus, such as a checkerboard image projected into the participant’s visual field. (In a block design study, the exogenous variable is not necessary, as each block type may be independently modeled and compared). In this case, the variable would have a value 1 when the checkerboard is present and 0 when the checkerboard is absent. If the checkerboard has a continuous property of some sort (such as contrast or luminance), its value might be anywhere between one and zero. The regression equations and graphical representation might then be as follows:

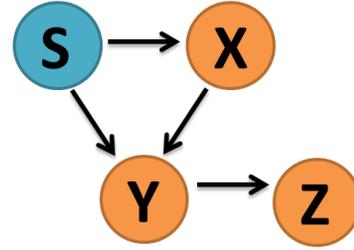
$$X_t = \beta_0 + \beta_9 S_t + \varepsilon$$

$$Y_t = \beta_3 + \beta_4 X_t + \beta_{10} S_t + \varepsilon$$

$$Z_t = \beta_6 + \beta_8 Y_t + \varepsilon$$

Matrix notation:

$$\begin{bmatrix} X_t \\ Y_t \\ Z_t \\ S_t \end{bmatrix} = \begin{bmatrix} \beta_{0X} \\ \beta_{0Y} \\ \beta_{0Z} \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & \beta_{XS} \\ \beta_{YX} & 0 & 0 & \beta_{YS} \\ 0 & \beta_{ZY} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} X_t \\ Y_t \\ Z_t \\ S_t \end{bmatrix} + \begin{bmatrix} \varepsilon_X \\ \varepsilon_Y \\ \varepsilon_Z \\ \varepsilon_S \end{bmatrix}$$



While SEM is an extremely valuable tool for estimating connectivity for known or predicted models, it is critical to understand that connectivity estimates are based on causal *assumptions* of the estimated model. For example, in the third model, the researcher has explicitly assumed that activity in X causes the activity in Y and uses SEM to estimate the magnitude of that connection. If, in fact, activity in B is the cause of activity in A, SEM may still yield a significant connection weight ( $\beta$  coefficient) in the wrong ( $X \rightarrow Y$ ) direction. Two approaches for reducing these errors are searching a wide variety of models to compare which direction best explains the observed data (comparing the model  $X \rightarrow Y$  to the model  $Y \rightarrow X$ ) or using time sequence as a causal constraint by regressing  $Y_t$  over  $X_{t-1}$  since activity in Y could not cause a change in X backwards in time. This second approach is known as vector autoregression.

**Vector autoregression (VAR)**, like SEM, uses a system of linear regression equations to estimate the connections between ROIs. However, unlike SEM, VAR compares the lagged (or delayed in time) relationships between ROIs. To illustrate this method statistically, consider only one ROI X. Because X's activation has been measured repeatedly through time, it could be described by a time series. Many time series have a property called autocorrelation which means that the value of X at time point  $t$  is correlated with the value of X at a previous time point. In the following discussion, we will look only at the immediately preceding time point  $t-1$ , denoted as a VAR(1) model. Autoregression reveals the *systematic* changes that occur in a variable over time, and this relationship is represented by the following regression equation:

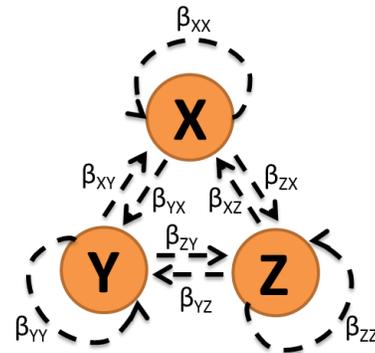
$$X_t = \beta_0 + \beta_1 X_{t-1} + \varepsilon$$

Vector autoregression generalizes this equation to allow the regression of several time series over themselves and each other. Extending the regression equation for  $X_t$ , we could describe X in relation to all three ROIs:

$$X_t = \beta_0 + \beta_1 X_{t-1} + \beta_2 Y_{t-1} + \beta_3 Z_{t-1} + \varepsilon$$

The same autoregression equations can also be estimated for each ROI, and vector autoregression provides a tool for solving them all simultaneously:

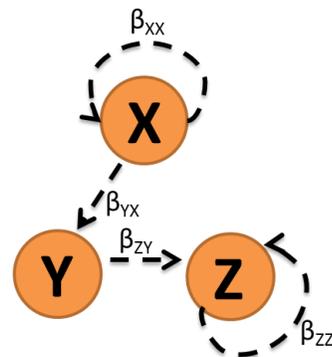
$$\begin{bmatrix} X_t \\ Y_t \\ Z_t \end{bmatrix} = \begin{bmatrix} \beta_{0A} \\ \beta_{0B} \\ \beta_{0C} \end{bmatrix} + \begin{bmatrix} \beta_{XX} & \beta_{XY} & \beta_{XZ} \\ \beta_{YX} & \beta_{YY} & \beta_{YZ} \\ \beta_{ZX} & \beta_{ZY} & \beta_{ZZ} \end{bmatrix} \begin{bmatrix} X_{t-1} \\ Y_{t-1} \\ Z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_X \\ \varepsilon_Y \\ \varepsilon_Z \end{bmatrix}$$



The graphical representation of this vector autoregression (depicted above) highlights that each ROI has a time-lagged relationship (denoted by the dashed lines) with every ROI, including on itself.

As was the case in SEM, the  $\beta$  matrix in the foregoing VAR equation is not very useful because it allows each ROI to have a lagged influence over every ROI (including itself), making practical interpretation difficult. Also like SEM, VAR is suitable for comparing competing hypotheses or may be searched iteratively to add only connections which significantly improve the model (as will be described in detail for the euSEM model). In this way, the  $\beta$  matrix may be sparsely populated with only the hypothetically relevant or statistically best fitting connections. The following regression equation and graphic represent one possible model:

$$\begin{bmatrix} X_t \\ Y_t \\ Z_t \end{bmatrix} = \begin{bmatrix} \beta_{0A} \\ \beta_{0B} \\ \beta_{0C} \end{bmatrix} + \begin{bmatrix} \beta_{XX} & 0 & 0 \\ \beta_{YX} & 0 & 0 \\ 0 & \beta_{ZY} & \beta_{ZZ} \end{bmatrix} \begin{bmatrix} X_{t-1} \\ Y_{t-1} \\ Z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_X \\ \varepsilon_Y \\ \varepsilon_Z \end{bmatrix}$$



One considerable advantage to examining lagged relationships is the additional causal evidence provided. Autoregression utilizes a form of causal inference known as Granger Causality which simply states that if two events happen sequentially in time, the causal relationship cannot occur backwards in time and must occur forwards in time. For example, if activity in ROI Y is correlated with activity in ROI X at a previous time point, the activity in Y could not have caused the past activity in X. Consequently, one may infer that activity in X is the cause of activity in Y. This inference ignores the third-variable problem (that X and Y may both be caused by something else and do not influence each other at all), but it remains a stronger inference than correlation of simultaneous measurements (as inferred in SEM).

**Extended Unified SEM** (euSEM) is a functional connectivity modeling technique designed to integrate the properties of both structural equation modeling (SEM) and vector autoregression (VAR). To this end, euSEM estimates both contemporaneous connections (as in SEM) and time-lagged connections (as in VAR) for both endogenous variables (other ROIs) and exogenous variables (stimuli). Finally, euSEM also provides for interactions between exogenous and endogenous variables, such that stimuli may modulate the connections between ROIs (as opposed to modulating activity within an ROI).

As in the preceding time series models, euSEM relies on the assumption of stationarity for each of the time series measured at each ROI. Stationarity entails a mean value for the time series that is constant over time. In graphical terms, the time series should not drift up or down but should center around a single mean value. Variance must also be constant for the time series to maintain stationarity. Graphically, the width of the distribution of observations around the mean should be equally wide over time. If either of these conditions is not met, the assumptions for performing the regression over a time series are invalid and may produce spurious results.

The matrix form of the regression equation for euSEM becomes quite large, and some shorthand will be necessary to present it on the page. Let the endogenous variables (ROIs such as  $X$ ,  $Y$ , and  $Z$ ) be represented by the vector  $\eta$  at time points  $t$  and  $t-1$ . Exogenous variables (like  $S$ ) will be represented by the vector  $u$  also at time points  $t$  and  $t-1$ . The  $\beta$  matrix will be composed of several sub-matrices:

- The lagged relationships (previously described in VAR) for endogenous variables will be described in the  $\Phi$  (phi) matrix.
- The contemporaneous relationships (previously described in SEM) for endogenous variables will be described in the  $A$  (alpha) matrix.
- $\Gamma_1$  will describe the lagged relationships between the endogenous and exogenous variables, and  $\Gamma_0$  will describe the contemporaneous relationships.
- The bilinear (interaction) terms between the exogenous and endogenous variables will be described as  $\tau_0$  and  $\tau_1$  for time points  $t$  and  $t-1$ , respectively.

Finally, the  $\varepsilon$  matrix will be described as  $\zeta$  matrices for each  $\eta$ ,  $u$ , and bilinear term.

$$\begin{bmatrix} \eta_{t-1} \\ \eta_t \\ u_{t-1} \\ u_t \\ \eta_{t-1}u_{t-1} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ \Phi & A & \Gamma_1 & \Gamma_0 & \tau_{11} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \eta_{t-1} \\ \eta_t \\ u_{t-1} \\ u_t \\ \eta_{t-1}u_{t-1} \end{bmatrix} + \begin{bmatrix} \zeta_{\eta(t-1)} \\ \zeta_{\eta(t)} \\ \zeta_{u(t-1)} \\ \zeta_{u(t)} \\ \zeta_{\eta u(t-1)} \end{bmatrix}$$

Notice, however, that the number of free parameters in the  $\beta$  matrix is relatively small compared to the size of the matrix. Since we hold  $S$  as exogenous (a stimulus, unaffected by brain activity), all regressors of  $S$  are set to zero. Similarly, since we maintain that causality cannot take place backwards in time, all regressors of an ROI at time point  $t-1$  (e.g.,  $X_{t-1}$ ) over ROIs at time point  $t$

(e.g.,  $X_t$ ) are also set to zero. That is, activity at  $X_t$  could not possibly influence activity at  $X_{t-1}$ . What remains is only predictors of  $\eta_t$  (the vector of observations at the ROIs for time point  $t$ ).

As in the previous models, solving this regression equation is unlikely to yield an interesting result if we attempt to estimate a parameter at every allowable connection. Instead, we can search the model iteratively by opening a single parameter at a time and comparing the resulting model to a preceding model without that parameter. When a new model ceases to yield a significant improvement by the addition of new open parameters, we conclude that the most parsimonious model has been generated. An example model may look like this:

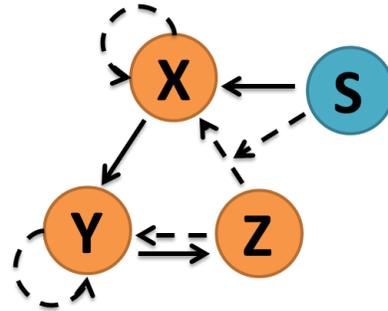
$$\Phi = \begin{bmatrix} \beta_{XX} & 0 & \beta_{XZ} \\ 0 & \beta_{YY} & \beta_{YZ} \\ 0 & 0 & 0 \end{bmatrix}$$

$$\tau_{11} = \begin{bmatrix} 0 & 0 & \beta_{XZS} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$A = \begin{bmatrix} 0 & 0 & 0 \\ \beta_{YX} & 0 & 0 \\ 0 & \beta_{ZY} & 0 \end{bmatrix}$$

$$\Gamma_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Gamma_0 = \begin{bmatrix} \beta_{XS} \\ 0 \\ 0 \end{bmatrix}$$



Diagnostic statistics for euSEM model provide a quantitative evaluation of the model's improvement over a baseline model, such as the model immediately prior to freeing an additional parameter. The chi-squared test compares the likelihood of two models to test the null hypothesis that two models are equally fit to the data. If we reject this null hypothesis, we can conclude that the model with the new parameter significantly improves the model. The  $\chi^2$  test is performed by the LISREL software package each time a model is estimated. If you want to estimate a connectivity model without using the GIMME package, you will have to manually update your model by freeing individual parameters, re-running LISREL, and checking the  $\chi^2$  test for the significance of the improvement.

Other diagnostics computed by LISREL such as the Root Mean Square Error of Approximation (RMSEA), Standardized Root Mean square Residual (Standardized RMR), Non-Normed Fit Index (NNFI), and Comparative Fit Index (CFI) also provide information on the fit of the model. In general, a well-fit model should produce RMSEA and Standardized RMR values around or below 0.05, and NNFI and CFI values should be above 0.95. It is possible to continue improving a model (as indicated by significant chi-squared scores) that satisfies these other diagnostic

criteria, but it remains in the judgment of the researcher to what extent to pursue increased fit versus parsimony of the model.

Using GIMME, however, this process of model evaluation is automated. The script will automatically search the parameters, freeing one at a time until the model is no longer significantly improved. GIMME performs this analysis first at the group level (searching for parameters that, when freed, will improve the fit of the majority of participants models) and then refine each participant's model to account for individual differences. The procedure performing this search is outlined in the next section. A brief GIMME user's manual written by the software author (Prof. Kathleen Gates, [kmgates@unc.edu](mailto:kmgates@unc.edu)) is also available.

## **Section II: Data Analysis**

Connectivity analysis of functional neuroimaging data is conducted in several steps. Here we will outline the procedure for fMRI data using the SPM software package (Friston & Holmes, 1995). This procedure is conceptually applicable across different software packages, but given the ubiquity of SPM, we will focus on this example.

The three main steps which will compose this analysis are:

1. Extraction of the time series data
2. Preparation of the data set for connectivity analysis
3. Estimation of the connectivity map with the GIMME package

### **Extraction of Time Series Data (using Marsbar)**

Time series data, that is, repeated measurements of the same region of interests (ROI, one or more voxels in an anatomically circumscribed region of the brain) through time, are the basic observations used in the connectivity analysis. The locations of these ROIS are arbitrary, defined by the researcher depending on her interests. In order to extract the time series data, it will be necessary to define these ROIs in SPM.

ROIs can be defined in SPM by selecting anatomical regions using the Pick Atlas tool and saving the resulting mask file. Alternately, ROIs can be defined manually by inputting MNI coordinates and defining the radius of a sphere surrounding the voxel at those coordinates. In either case, the Matlab package Marsbar is used to extract these time series data from your SPM files.

Marsbar is launched by selecting "Marsbar" from SPM8's Toolbox menu. At the initial screen, open "ROI definition" and select "Build". At this stage, you will decide whether to define the ROI anatomically (based on the a mask you defined using the Pick Atlas) or manually (based on a sphere around MNI coordinates). In the third menu (Type of ROI) select "Image" for an anatomical mask or select "Sphere" for a manual ROI.

For anatomically defined ROIs: Select the saved mask for the contrast you have defined in SPM. After selecting the mask, respond "Yes" to "Maintain binary image" and "No" to "Apply function to image."

For manually defined ROIs: Enter the MNI coordinates for the center of your sphere and the radius of the sphere in millimeters (mm).

In both cases, save the file by entering a name in “Label for ROI” field, opening the Save dialogue and clicking Save. To extract the time series data, select “Extract ROI Data (default)” from the Data menu. Select the ROI mask that you just saved, and then select the SPM.mat (design file) for the participant whose time series you are trying to extract. Once the data have been extracted you can select “Export data” from the Data menu, choose “Summary times series”, and save the time series to an MS Excel format file.

### Preparation of the Data Set

Having acquired the time series from Marsbar, some work will be required to insure these files follow a format compatible with the GIMME software. In brief, GIMME requires each participant's time series data to be saved in a separate file as a  $t$ -by- $r$  matrix, where the vertical dimension  $t$  is the set of time points (the length of a single time series) and the horizontal dimension  $r$  is the number of ROIs. It's most convenient to edit these files in Microsoft Excel (or LibreOffice Calc) and save them to one of several file formats accepted by GIMME: csv (comma separated variable), txt (text), or xls/xlsx (Excel).

Marsbar outputs time series data in the  $t$ -by- $r$  matrix format, correct dimensions for use in GIMME. However, Marsbar starts every column with a label for the ROI from which measurements are taken. For that reason, it is necessary to open the Marsbar output in Excel, remove the first row (containing labels) and save the data again in an acceptable format, one file for each participant.

If you are performing an *extended* unified SEM analysis (including stimulus events in the connectivity analysis) you will also need to build an onset matrix for these events. This matrix must be saved in \*.mat format (Matlab workspace) in a file called `onsets.mat` containing a variable `onsets1`. This variable is a  $t$ -by- $n$  matrix, where the vertical dimension  $t$  is the set of time points (in either units of either scans or seconds) and the horizontal dimension  $n$  is the number of participants. Each column of this matrix should be filled with zeros, except for ones at the time-points when a stimulus was presented for that particular participant. If a second stimulus type is included in the analysis, a second variable named `onsets2` maybe be included in the `onsets.mat` file, following the same convention as `onsets1`. This `onsets.mat` file should be stored in another directory *outside* the `source` directory, titled `output`.

### Estimation of the Connectivity Map

With all of the time series data and stimulus data coded into matrices and saved in the proper locations, GIMME makes estimation of a connectivity map very easy. Before enumerating the steps for doing this estimation, though, it is valuable to have some understanding of the underlying process. GIMME is a set of Matlab functions which assist the user by composing LISREL syntax to describe the time series data and guide the iterative search process for the best fitting model over the group data. To complete this process, GIMME will use Matlab's operating

system interface to call the LISREL program and execute the commands that GIMME has composed for LISREL. It is therefore necessary to have LISREL installed in the location where GIMME expects to find it.

In the following steps, the procedure for estimating the connectivity map is outlined. For additional technical details about the design and use of GIMME, see the GIMME users' manual (Gates, 2012).

1. Add GIMME to Matlab's executable path.

```
addpath('C:\path\to\GIMME\scripts\')
```

2. Start the top-level script by typing “GIMME” into the Matlab command line.
3. A window will appear with several parameters that need to be defined:

**Number of ROIs:** Enter the number of regions that you have specified in the time series data for each participant. If you participant time series data have 3 columns, you have 3 ROIs (one time series for each region). In that case, you would enter “3”.

**Indicate directory where your data are:** This is the directory which contains the time series data files for each participant (may be \*.csv, \*.xls, or \*.txt format). Make sure that every participant's data file is in that directory (no additional directories), and make sure there are *no other* files in that directory, as GIMME will attempt to open each of the files it finds.

**Indicate directory where your results should go:** This directory will be populated with \*.mat files for each participant. Critically, if you are estimated an euSEM, your `onsets.mat` file must also be stored in this directory before you start GIMME.

**TR:** The value you enter here should be the length of a single scan (one TR) measured in seconds.

**What type of files are your data in?** Here you will select the integer value that corresponds to the data type you saved your participants' time series data files in. The correct choice here is the file type that you used in “Preparation of the Data Set” when you saved the time series matrices.

**Put “0” to conduct a uSEM, “1” for a euSEM:** If you have a stimulus with specified onsets (in the `onsets.mat` file) that you would like to include in the estimation of the connectivity map, select the euSEM by entering a 1. If you do not want to include stimulus data, enter 0.

**If this is an euSEM, how many inputs are there?** If you selected “1” for the previous question, you must now enter the number of inputs (stimuli) which you wish to estimate.

At this time, GIMME only supports 1 or 2 stimuli, so these are the only possible stimuli. If you selected “0” for the previous question, leave this question blank.

**If euSEM, are you input vectors in seconds (1) or TRs (2)?** If you are including stimuli (inputs) in the connectivity map, you must specify whether the variables in onsets.mat were defined by seconds (one observation per second) or TRs (one observation per TR). Another way to ask this question is whether the input matrix’s vertical dimension  $t$  is the number of seconds in the study or the number of scans. If you are estimating a uSEM, and therefore have no stimuli to input, enter 0 for this value.

**Would you like to start with the autoregressive terms estimated?** The connectivity analysis is expedited by first estimating the effect of each ROI on itself over time by performing an independent autoregression at each ROI. This speeds up the later estimation of the connectivity map between ROIs, but the effects of doing so are not fully understood (i.e., whether this will introduce error into the connectivity analysis).

After selecting all of the appropriate options for your analysis, click **GIMME**

4. GIMME may take several minutes to complete the individual and group level analyses. When it is finished, the specified output folder will be populated with:
  - `finalALL.mat` describing the best fit group connectivity map
  - `final[subj#].mat` files for each individual participant, describing that participant’s individual connectivity map

Within each output file, you will find data structures containing the regression weights described for Extended Unified SEM in Section I.

### Section III: References

- Friston, K. J., & Holmes, A. (1995). Statistical parametric maps in functional imaging: a general linear approach. *Human Brain Mapping*, 2, 189–210.
- Gates, K. M. (2012). Group Iterative Multiple Model Estimation (GIMME) User's Manual. Retrieved from [www.personal.psu.edu/kmg311/GIMME/Using%20GIMME.pdf](http://www.personal.psu.edu/kmg311/GIMME/Using%20GIMME.pdf).
- Gates, K. M. & Molenaar, P. C. M. (2012). Group search algorithm recovers effective connectivity maps for individuals in homogeneous and heterogeneous samples. *NeuroImage*, 63, 310-319.
- Gates, K. M., Molenaar, P. C. M., Hillary, F. G., & Slobounov, S. (2011). Extended unified SEM approach for modeling event-related fMRI data. *NeuroImage*, 54(2), 1151–8.
- Kim, J., Zhu, W., Chang, L., Bentler, P.M., Ernst, T., (2007). Unified structural equation modeling approach for the analysis of multisubject, multivariate functional MRI data. *Human Brain Mapping*, 28, 85–93.