Use of fMRI and structural equation modeling for studying interconnected brain areas within a hypothetical network

Giovanni de Marco, Sidy Fall, Pierre Vrignaud

Use of fMRI and structural equation modeling for studying interconnected brain areas within a hypothetical network

de Marco, G.[a*] (PhD), Fall, S.[a] (Master), Vrignaud, P.[b] (PhD)

[a]Laboratoire de traitement d'images médicales, UPJV, CHU Nord, Amiens, France

[b]Équipe « Travail et Développement Professionnel ». EA 3984. Université de Paris 10, Nanterre France.

Abstract

Functional neuroimaging first allowed researchers to describe the functional segregation of regionally activated areas during a variety of experimental tasks. More recently, functional integration studies have described how these functionally specialized areas (i.e. areas whose activity is temporally modified) interact within a highly distributed neural network. When applied to the field of functional magnetic resonance imaging (fMRI), structural equation modeling (SEM) uses theoretical and/or empirical hypotheses to estimate the effects (path coefficients) of an experimental task within a putative network. Structural equation modeling represents a linear technique for multivariate analysis of fMRI data and has been developed to simultaneously examine ratios of multiple causality in an experimental design; the method attempts to explain a covariance structure within an anatomical (constrained) model. This method, when combined with the concept of effective connectivity, can provide information on the strength and direction of the functional interactions which take place between identified nodes of a putative network. After having provided a brief reminder of the principle of the blood oxygen level-dependent (BOLD) contrast effect, the physiological bases of brain activity and the concepts of functional integration and effective connectivity, we specify the various steps in the SEM analysis and the use of fMRI data to explore putative networks of interconnected active areas.

Keywords: fMRI, model, network, effective connectivity, SEM

Introduction

Functional magnetic resonance imaging (fMRI) has enabled great progress in the field of cognitive science. Indeed, over the last two decades, fascinating discoveries have been made in terms of the brain's structure, organization and function. Thanks to neuroscience, we can now better understand the way the brain operates when it recognizes a face [1-4], admires a work of art or learns a language [5]. Today, fMRI helps us locate the brain areas involved in the various forms of vision or emotion, in self-awareness and in awareness of others. On the basis of assumptions and theoretical and/or empiric models, we are now able to show how brain areas interact with each other and how external stimuli may modulate the latter's activity [6-11]. In order to understand the various stages which lead from fMRI-measured dynamic brain activity to interpretation of a large-scale network of brain functions, we first need to examine briefly the physiological bases of the hemodynamic blood oxygen level-dependent (BOLD) response. Secondly, we will describe the classic statistical approach (which enables measurement of brain activity in an exploratory way) and the concepts of functional and effective connectivity (which enable examination of the interactions between various brain areas). Thirdly, we will describe the principles behind the structural equation modeling (SEM) method which enables the analysis of interdependency relationships within a putative network of functional areas. Finally, we will emphasize the stages in SEM which must be thoroughly mastered in order to use this type of analysis correctly.

Physiology and brain activity measurements

Cortical activity measured by fMRI may correspond to a cascade of neurochemical events which occur within neurons and the glial cells (astrocytes, in particular) located near the synapses [12]. These cells play an essential role in reuptake of glutamate (the

neurotransmitter) released at the synapse following a change in membrane voltage. Glutamate is quickly transported within adjacent astrocytes where it is converted into glutamine for transfer back to the neuron. This synaptic and dendritic activity, when integrated over a few millimeters of neural tissue, would display relatively weak energy efficiency (production of ATP) but would roughly reflect the slow fluctuations in local field potentials, corresponding to electric entries and their intracortical treatment [13]. Unit recordings have been used to show that the frequency of neuronal discharge is related to the fMRI signal [12]. Thus, the BOLD contrast obtained in fMRI may represent the signature of the electrical, energetic and hemodynamic consequences of brain activity. By comparison, the signals measured in electroencephalography (EEG) and magneto-encephalography (MEG) are closely related to neuronal currents [14] and thus may be more sensitive to post-synaptic activity [15]. Nevertheless, the exact physiological process leading from neuronal activity to the BOLD signal is still under investigation [16-20]. Despite a certain level of uncertainty concerning the exact nature of the microscopic mechanisms at the origin of neurovascular coupling, a perceptive or cognitive stimulation carried out in the scanner gives rise to a cascade of metabolic processes associated with neuronal activity in various areas of the brain. A hemodynamic response to this metabolic demand involves an increase in cerebral blood flow and volume, which will vary in terms of its intensity in each activated voxel. The BOLD signal will change over time according to the metabolic demand and location of the involved cerebral areas and with respect to the oxygenation conditions, cerebral blood flow and volume.

Neural assemblies

Neural assemblies which provide a conceptual framework for the integration of distributed neural activity are defined as distributed local networks of neurons transiently linked reciprocal dynamic connections [7,21,22]. The emergence of a specific neuronal assembly is thought to underlie the operation of every cognitive act. Neurons that belong to a given assembly are linked by selective interactions; that is, they interact preferentially with a sub-ensemble of other neurons that are interconnected. These interactions are mediated through direct (monosynaptic) or indirect (polysynaptic) connections that are typically reciprocal [23]. On the one hand, there are reciprocal connections within the same cortical area or between areas situated at the same level of the network. On the other, there exist connections that link different levels of the network in different brain regions to the same assembly [24]. Connections of this type have been traditionally described as feedforward and feedback (or as bottom-up and top-down). These neural assemblies have a transient, dynamical existence that spans the time required to accomplish an elementary cognitive act (a fraction of a second). But, at the same time, their existence is long enough for neural activity to propagate through the assembly, a propagation that necessarily involves cycles of reciprocal spike exchanges with transmission delays that last tens of milliseconds. So, the relevant variable required to describe these assemblies is not so much the individual activity of the components of the system but the dynamic nature of the links between them [25].

Local and large scale functional integration

Emerging hypothesizes in neuroscience can use fMRI to study functional interactivity. By defining neural assemblies as distributed local networks which are transiently linked by large-scale reciprocal dynamical connections, Varela et al. [25] made a clear distinction between

local and large-scale networks. On one hand, a local network is defined as a patch of neural tissue that synchronizes its activity through the local cyto-architecture, i.e. groups of entities (neurons) acting together in a coherent fashion [26]. Local integration occurs over a local network distributed over an area of ~1cm through monosynaptic connections with conduction delays of typically 4-6 ms [27]. For example, in columns of the primary visual cortex separated by 2-7 mm, which have non-overlapping receptive fields, neurons that share similar feature properties tend to synchronize [28]. On the other hand, large-scale dynamic connections are defined as interactions based on large fiber pathways between regions that are located far apart from one another. Large-scale synchronization concerns neural assemblies which are further apart in the brain (>1cm; transmission delays>8-10ms over polysynaptic pathways) such as, for example, assemblies between occipital and frontal lobes or across hemispheres, which are separated by dozens of milliseconds in transmission time [27].

The dichotomy between local and large-scale networks serves as a neural basis for the key assumption that brain functional architecture abides by two principles: functional segregation and functional integration. While the segregation principle states that some functional processes specifically engage well-localized and specialized brain regions, it is now thought that brain functions are most likely to emerge through integration of information flows across distributed regions [25,29,30]. In this approach, it is not only isolated brain areas that are presumed to process information but rather a large-scale network, i.e. a set of brain regions interacting in a coherent and dynamic way. Hence, according to the functional integration concept, cortical areas and thus functions are integrated within specific dynamic networks. This concept supposes that there is dynamic interaction between interconnected, active areas and thus that the brain areas are expressed as networks within integrated systems. In such a system, localized areas are included in networks which become dynamic according to the cognitive task. Brain areas underlie several functions and can belong successively to several

different functional networks. In other words, a given brain area does not have a single function; its resources can be exploited in several different cognitive strategies. The principle of functional integration (which is also known in the field of electrophysiology) was used to analyze the event potentials obtained from multipolar recordings [31]. Thus, based on the functional integration principle, the relationships between several brain areas may be examined.

MR Imaging Processing

Individual analysis. In neuroscience, fMRI statistical results typically appear as statistical parametric maps (SPMs) which display relative activities (versus a control) according to the most commonly asked question: which brain regions are involved in the functional process under investigation? Hence, the majority of functional MR imaging studies try to locate cerebral functions by modeling the form of the hemodynamic response by a canonical function (the general linear model), which is evoked by a specific component of a given cognitive task [20,32]. The univariate analysis of BOLD signals served as a data reduction step in order to identify significant voxels based on the meaningful contrasts (conditions) according to the subtraction method. A single subject, fixed-effects model analysis is performed for each individual subject in order to prepare the extraction of the BOLD signal. In each single subject analysis, a minimum significance level of $p = 0.05$ is applied to detect activated voxels for the meaningful contrast. Local maxima are located within the predefined brain regions and then assigned to a Brodmann or anatomical area. This univariate analysis for revealing brain activity does not enable us to understand how brain areas interact with each other to modify their respective activities. This measurement method does not provide information on the interregional relationships in the brain and, currently, cannot tell us how

the regions interact. In other words, univariate statistics do not enable us to study the covariations (or causal relations) between various connected brain areas.

Regions of interest (ROIs) approach. Specific hypotheses for activation within regions of interest are located individually in each subject using for example Marsbar (http://marsbar.sourceforge.net/) or PickAtlas softwares (http://fmri.wfubmc.edu/cms/softwar) which provide a method for generating ROI masks based on the Talairach Daemon database. Talairach Daemon software (http://ric.uthscsa.edu/new/resources/talairachdaemon) and mni2tal tool are used to automatically define Talairach atlas labels and to convert coordinates from MNI brain for equivalent Talairach coordinates using a nonlinear transformation respectively. The mni2tal algorithm is used to ensure that the activated pixels actually corresponded to the predefined brain regions (http://www.mrc-cbu.cam.ac.uk/Imaging/Common/mnispace.shtml).

BOLD signal extraction for SEM. Local maxima within the predefined areas and BOLD signal extracted for the SEM analysis are identified for each subject and each ROI. The extraction of the weighted mean BOLD signal is performed separately in each area. BOLD signals are thereafter normalized and concatenated to allow comparison of BOLD signals across subjects and sessions.

Group analysis. While fMRI data are analyzed in a "whole-brain" fashion, the functional integration (connectivity) study is constrained to specific hypotheses concerning certain regions of interest, e.g. the inferior parietal cortex (BA 40), the middle frontal cortex (BA 46), the anterior cingulate cortex (BA 32), etc. At the group level, these regions were identified at the second (between -subjects) level using two one-sample t-test contrasts testing for

activation in both groups (or both conditions for one group) with p values statistically corrected for multiple comparisons. Local maxima were located within the predefined brain regions and then assigned to a Talairach coordinate and a Brodmann area.

Exploration of the effective connectivity

Effective connectivity is closer to the intuitive notion of a connection and can be defined as the influence that one neural system exerts over another, either at a synaptic level (synaptic efficacy) or a cortical level [11,33]. In electrophysiology, there is a close relationship between effective connectivity and synaptic efficacy [34]. Effective connectivity can be estimated from structural equation modeling (SEM), a linear technique that can be used to test whether a theoretical model seeking to explain a network of relationships can actually fit the relationships estimated from the observed data. In the case of fMRI, the theoretical model is an anatomical (constrained) model and the data are interregional covariances of activity [35]. To describe a functional network, network nodes (active areas) and anatomical connections (fibers) must be proposed in conjunction with SEM in order to model interregional covariances and determine the intensity of the connections. Effective connectivity will represent the dynamic influence that cortical and subcortical regions exert on each other via a putative network of interdependent areas. An increasing number of fMRI and PET studies have used the SEM method to show specific networks involved either in working memory [36-41], within certain attentional processes [42-44], face perception [4,45], movement processing [46-50] or verbal fluency (semantic task) [8].

Path model construction and Structural Equation Modeling (SEM)

9

The path model is based on a hypothesized network restricted to different topologically distinct brain regions. We used SEM to construct a pathway model that could account for fMRI data in all regions of interest. Adjusted signals extracted from the data set (see BOLD signal extraction for SEM) are entered as variables. In SEM analysis, the variables are considered in terms of the covariance structure (see below for the detailed method).

Softwares for SEM analysis

Structural Equation Modelling (SEM) can be performed with a number of different software programs. The most popular are LISREL (http://www.ssicentral.com/lisrel), EQS (http://www.mvsoft.com/products.htm) and AMOS (http://www.spss.com/amos/). Other, less widely used software includes CALIS (SEM procedure in SAS), EZPATH (in Statistica) and MPLUS (http://www.StatModel.com/), which works on non-normal variables (binary, censured, etc.).

<u>**Basic concepts**</u>

SEM is a technique of multivariate statistical analysis that can be used on fMRI data. Originally, SEM was developed to analyze psychometric and economic data. Structural models make it possible to analyze linear relationships between variables from the analysis of the covariance among the variables. Structural models evolved from two principal methods of analyses 1) factorial analysis (for a review: Bollen, 1989a) and 2) multiple regression or causal path analysis (a method developed in the 1930's by Wright e.g., (for a review: Wolfe, 1999)). Structural models examine, simultaneously, multiple sources of influence on the dependent variable in an experiment [51,52]. When applied to fMRI data, SEM allows

modelling of paths of connection between cortical areas and reveals relations, interdependencies and covariance among the various cortical areas. Given an anatomical model, SEM shows the effects of an experimental task on specific network of connections [33,53-55]. In this type of statistical analysis, normalized signals (variables) are considered in terms of the structure of their covariances. Thus, SEM allows one to infer interregional dependencies (interactivities) between various cerebral cortical areas.

SEM methods, in comparison with classical approaches such as linear regression, allows one to simultaneously analyze several types of interrelationships between variables in an experiment [6,11,20,56]. Variables can be directly observable or indirectly observable (latent), i.e., variables that are inferred from multiple indicators[1]. This paper discusses only directly observable variables.

- Variables may be directly related or indirectly related. For example, the effect of A on C may be direct, (A->C), or it may be indirect or mediated by B, (A -> B -> C). SEM can distinguish direct from indirect relations.

- A relation between variables may be recursive. That is, the effect of A on B, (A->B), may be different than the effect of B on A, (B->A).

- Relations between variables may be reciprocal. For example, an important question often asked is, "Which of two variables acts more on a third one?" SEM answers the question by analyzing the relations on a hierarchical basis.

- The nature of the relation between variables is given by the regression coefficient; it describes how much the dependent variable changes when an independent variable changes by one unit. SEM directly integrates the errors of measurement into a statistical model, by doing so the estimates of regression coefficients are more precise than they are with classical

methods such as multiple regression, factorial analysis, analysis of variance or discriminant analysis. The older methods examine only one linear relation at the same time between independent and dependant variables and do so only within a range of values set by the researcher [33][1].

Contrary to classical methods, SEM is interested in a structure of variances/covariances in a dataset of observed variables and it will try to predict dependencies among the variables (see Figure 1). In other words, SEM seeks to explain as much of the variance in dependant variables as it can from the simultaneous measurement of the variances of independent variables that are included in the model. Similarly, SEM incorporates errors of measurement (residual variances) of the independent variables into the calculation of estimate, which reinforces the statistical power of the method and provides more precise estimates of b coefficients (coefficients of regression). Thus, one can validate a model of measurement from a theoretical model or empirical data [41].

Constructing and Specifying the Theoretical Model

Relationships between each variable should be supported by the hypotheses of a theoretical framework and/or empirical observations. It is important not to voluntarily omit significant explanatory variables in the model of analysis (error of the model specification) at the risk of under or over-estimating (biasing) the effects of the other variables that are retained in the model. The stronger the correlation between these variables, the more the coefficients of regression will be biased. The principle of parsimony requires that we identify and retain

---

[1]Latent variables play a double role: in the first part of causal path they act as dependent variables or variables to be explained and in the second part they act as independent or explanatory variables. The concept of latent variables is important because it permits one to study non-directly or indirectly observable phenomena; phenomena for which we have no direct measurements. So, SEM uses observable variables that are strongly associated with the latent variable as indicators of the latent variable itself; the result is the latent structure- the relations among the indicator variables, which can then be analyzed

explanatory variables and the most relevant relations or paths among them to explain a phenomenon. We can do so by first testing an initial, economical, model and then making it more complex, if necessary. The objective is to retain the fewest number of variables and interrelationships (paths) that explain as much as the variance in the phenomenon, as possible.

The specification step consists of constructing a diagram of relations (e.g., Figure 1A) and then determining the structural model by answering various questions: Which linear relations between variables must be tested? What is the nature of their relationship (one-way, reciprocal/recursive)? What are the different parameters to be estimated (coefficients of regression, covariances)? Which parameters are fixed, free or constrained?

Figure 1, adapted from [4]illustrates an example of SEM. We examined effective connectivity within an emotional network composed of three brain areas involved in processing fearful faces: Orbito-frontal (OFC or VAR1), Anterior Cingulate Cortex (ACC or VAR2) and Amygdala (AMY or VAR3). Their experimental manipulation consisted of two tasks, an incidental perception task (gender identification) (n=14 subjects) and an intentional detection task (effortful discrimination) (n= 10 subjects), each task performed with three facial expressions, fearful, neutral and ambiguous. Subjects were scanned while performing the tasks. Effective connectivity between the three brain areas was assessed using SEM.

Results show that the hypothetical network fits the experimental data for both tasks, in both hemispheres. The comparison between task 1 and 2 reveals significant differences in strength and direction of the connectivity patterns for the left hemisphere and, to a lesser extent, for the right hemisphere. The path coefficients analysis suggests that during incidental perception, the fearful information generated in AMY, reaches the OFC through the ACC, while in the intentional perception task fearful information generated in AMY follows a reverse route,

from the OFC to the ACC. The findings confirm differential brain connectivity during incidental and intentional processing of fearful faces.

Insert figure 1 about here

Free parameters are those whose value is unknown and which the algorithm has to estimate. Constrained parameters are free parameters on which the model imposes a constraint, often a constraint of equality between their values. Fixed parameters indicate potential paths which will not be estimated; they are fixed at zero and so, do not enter in the procedure of calculation. This diagram of relations (Figure 1A) is then transposed into linear equations as:

$$VAR1 = \textbf{pV1\_V3} * VAR3 + \textbf{pV1\_V2} * VAR2 + 1 * e\_var1$$
$$VAR2 = \textbf{pV2\_V3} * VAR3 + \textbf{pV2\_V1} * VAR1 + 1 * e\_var2$$

The parameters are named, in bold, by the variables that are related: It is convenient to put the influenced variable (effect) first and the influencing variable (cause) second.

From the system of linear equations we can calculate the proportion of variance in the observed variables (e.g. VAR1) that is explained by the variables on the right side of the equation. The more closely the proportion of the explained variance approaches 1.0, the greater our confidence that we have identified the variables that are associated with the observed variable. Conversely, when the proportion of explained variance approaches zero we have little confidence that the variables are linearly related.

Estimates of the theoretical model

SEM works with a matrix of covariances rather than with a matrix of correlations. This choice implies that the value of estimated coefficients will depend on the unit of measure of the variables (like the b coefficients in the multiple regression analysis). Standardizing the values of the coefficients gives them all the same metric, 1.0, which makes them easier to interpret. The standardized values are equivalent to beta coefficients of simple or multiple regressions.

Calculating the model parameters. Only relationships proposed by the theoretical model are examined. The estimate of the model parameters (covariances, correlations, coefficients of regression) by maximum likelihood (robust estimator) is achieved with an iterative process guided by the decrease in the value of the difference between the observed covariance matrix and the theoretical covariance matrix calculated from the parameters included in the model. An iterative procedure adjusts the predictive values with the observed values. The estimator is consistent when the predictive values converge towards the observed, true, values of parameters.

Testing hypothesis. At the end of the estimation process, a test of fit is applied: In this test, the value of the function representing the difference between the observed and predicted matrix is used to compute an indicator called χ^2 which follows a chi-square distribution. The null hypothesis is that there is no difference (a gap of zero) between the predicted and observed matrix in the parent population (test of exact fit). Hence, a good fit between the observed and predicted matrices is indicated by the impossibility of rejecting the null hypothesis (Ho). This chi-square test is the only indicator of adequacy which follows a known distribution : the distribution of chi-square with [p (p+1)/2-q] degrees of freedom; p = number of observed variables, q = number of estimated parameters). A general consensus in the scientific community is to consider that the model fits the data when the probability of the Chi-square test has a value of $p > 0.05$.

The estimation algorithm (ML) and the Chi-square test rely on the assumption of the multinormality of the variables,. The multinormality assumption can be tested by calculating either coefficients of symmetry and concentration, or the coefficient of Mardia, which takes all the variables into account simultaneously.

Sample size. In general, it is advisable to have a sample size including at least about ten observations by estimated parameter [57]. If we take into account that the fit of the model is assessed by a chi2 statistics that is calculated using the sample size (n) and the fit function, there can be a lack of sensitivity (not rejecting the null hypothesis) when the sample size is small (i.e. less than one hundred), at the contrary, the test can become too much sensitive (arousing the type II error: rejecting wrongly the null hypothesis) when the sample size is higher than one thousand. It is henceforth recommended as, a rule of thumb, to work with sample including at least the number of observations corresponding to the number of observations necessary to get a reliable estimation of the parameters. But, according to the sensitivity of the test of fit a sample size of several hundred will be optimal to avoid a lack – reciprocally an excess – of sensitivity [57]. In fMRI, it means that the number of observations will depend on the subject and scans number. All these information are based upon recommendations in the framework of the Maximum Likelihood algorithm. Recent developments in the use of Bayesian algorithm (Markov Chain Monte-Carlo with Gibbs sampler and/or Metropolis Hastings) to estimate SEM parameters have proved to give consistent estimations even with samples smaller than one hundred observations [58]

Identifying a theoretical model

The identification of a theoretical model is a determining issue in the resolution of an analysis by SEM. A problem of model identification means that the algorithm used is unable to generate single estimates of the parameters, which will lead it to reject the model. Thus, to

identify a model implies that it is theoretically possible to calculate a single estimate of each one of its parameters (coefficient of correlation, covariance or regression). This question is partially important with the complexity of the tested model. The more the theoretical model becomes complex, the more the difficulty of its identification increases. The first condition necessary to identify a model is that there are at least as many sample moments as parameters to be estimated. The second is the absence of multi-colinearities between the observed variables. If some of them are too strongly correlated, that reduces the number of observations. More largely, the problem of the identification of a model is relied on the basic concept of degrees of freedom.

As previously noted, we want a structural model that is able to reproduce the theoretical covariance matrix (predicted) from the observed covariance matrix. Thus, the model must be identifiable. For that, we need to compare the number of parameters to be estimated (unknown) with the number of observed variables (known). In general, the simplest structural model will be always preferred because it has fewer parameters to be estimated, it is more economical, more stable and it will have a greater number of degrees of freedom, which makes it more difficult to reject the null hypothesis. We thus will seek to estimate each parameter in a single way (a unique solution). In the model just identified, the number of observed variables equals the number of parameters to be estimated, the model, called the saturated model, gives a trivial solution for the estimation of the model parameters and thus is non-informative. In the model under identified, the number of observed variables is lower than the number of parameters to be estimated, the model tries to estimate more coefficients than there is information available in the matrix of data; the estimation is impossible because there are too many unknown parameters. In the model over identified, the number of observed variables is higher than the number of parameters to be estimated. This model is optimal since we have fewer parameters to estimate than we have observed variables; thus, the model is

identifiable. To identify a model, it is often necessary to fix or constrain certain parameters, to fix variances of error of observed variables, to fix coefficients of regression measuring the known effects with certainty, to make the model identifiable.

Error Variance (or residual variance). Error variances represent unmeasured influences that are outside the model. They indicate a proportion of the total variance of a variable not taken into account by the other variables of the model [37,59,60]. Finding that a certain brain area has no residual variance would imply that all variance in this area is taken into account by connections with other areas in the model of measurement, which seems unlikely in neural systems. This error variance generally comes from the influence of brain areas not included within the model or even from the influence of an area on it. In practice, error variance can be fixed using rough estimations obtained via multiple regression, which reduces the number of parameters to be estimated in the model.

Evaluating the model's adequacy to the data

The Chi square test presented above is the only test based on frequentist inference (Neyman-Pearson approach) available in SEM. The major drawback of Chi square test is its dependency on sample size, and to a lesser extent on the number of parameters introduced in the model. There are other indexes of fit of the model which are based on error of approximation, which is indicated by difference between the matrix of observed covariance and the matrix of estimated covariance. These indexes can adapt to the sample size and to the nature of the data; they are based on the total population or the sample and they can support a simple model or not penalize a complex one.

Contrary to models just perfectly identified ($N_{obs} = N_{par}$), models over identified ($N_{obs} > N_{par}$) never perfectly adjust the observed data. Thus, the degree of adequacy of the models must be estimated. For that, we use indexes of adequacy or indicators. However, such indicators often

evaluate an overall fit based on averaged values, thus it is extremely probable that certain parts of the model must be slightly adjusted. In this case, a set of indicators will be necessary to validate the adequacy of the model. Each indicator will generally reflect a particular aspect of the fit of the model; this is why several indicators are often used to test the fit of two matrices. By the way, numerous studies based on simulated data have shown that different indicators can be more or less sensitive to different types of misfit or to different types of bias (sample size, deviation from multinormality), it's henceforth recommended to establish the decision of model acceptation (or rejection) by considering a set of indicators rather than one or two indicators isolated [61,62]. Thus, from theoretical model we will estimate path coefficients which will be validated on the basis of a chi-square test and by the means of various indicators [50,57,61,63,64].

Briefly, absolute fit indexes evaluate how well a theoretical model correctly reproduces observed data. Incremental fit indexes (or indexes of comparison) measure the fit improvement by comparing the tested model with a more restrictive basic model. Indexes of parsimony allow one to avoid under or over-estimating a given model, and to determine the most suitable model.

Absolute fit indexes. There is no perfect absolute index of measurement (gold standard) for evaluating the improvement of fit [61,64,65]. All indexes presented in this section are approximation indexes that test the quality of fit.

Goodness-of-fit indexes. In principle, goodness-of-fit indexes compare the matrix of covariance predicted by the model with that of a "*saturated*" model, which would perfectly predict observed data. We first obtain a value of chi-square (χ^2):

$$\chi^2 = (N-1)F_{\min}$$ Relation 1, where N is the sample size and F_{min} is the minimal value of the minimization function obtained through the estimation algorithm used (in general maximum likelihood).

We use the test based on the minimization function χ^2 or test of exact fit. Ideally the test result should be non-significant. This index follows a chi square distribution so it is possible to determine a p-value associated with the chi-square value and the number of degree of freedom (df) of the tested model. The Chi-square is biased by the sample size because sample size (n) enters into the computation of this index. So, this index presents two disadvantages: its value is affected by (1) sample size, and (2) the number of parameters in the tested model. Furthermore, the chi-square is sensitive to the magnitude of the correlations. The stronger the correlations the greater the values of chi-square, and thus the larger the difference between observed correlations and those predicted by the model [55].

During the last thirty years new indexes have been designed that avoid the dependencies on sample size and the number of parameter. Their authors have tried to find indexes which have know intervals of variation and for which threshold values can be found that allow one to objectively decide whether to accept or reject the model. In the last ten years, the limitations of theses indexes, demonstrated through simulation studies, resulted in the approach of using several references and comparing their respective values for tested model [61,62] This point will be discussed after presenting the main indexes, regrouped under the rationale of their construction principle.

Approaches based on the comparison of the fit of reference models. Two reference models are the best model (saturated model) and the worst model (model of independence). For the independence model we assume that there is no relation between variables. Covariance values

are fixed at zero and only the values of variances are estimated, thus this model can be considered as the model presenting the worst fit to the data. The ratio of the model tested and model of independence is particularly instructive because it shows improvement made by the model tested. These indicators base on the relationship between the χ^2 of the tested model and the χ^2 of the model of independence. Ratio values can theoretically vary from 0 to 1: We consider that values about 0.90 indicate a good adjustment. NFI ("Normed Fit Index") and RFI ("Relative Fit Index") [64]correspond to the ratio between the χ^2 of the considered models, RFI is corrected by the number of degrees of freedom (see formulas in appendix). IFI ("The Incremental Fit Index") [66] and TLI ("Tucker-Lewis Index") [67] are based on the ratio between the difference of χ^2 of the model of independence and the tested model and the difference of χ^2 of the basic model and the number of degrees of freedom of the model. Moreover, TLI is corrected for the number of degrees of freedom. We will be able to neglect the CFI ("Comparative Fit Index") in this first approach. NFI and TLI are two indexes that can be used to judge the fit improvement when comparing two models designed by the researcher. We discuss this issue when presenting the model comparison.

The GFI ("Goodness of Fit Index") can be considered to estimate the proportion of variance of the observed matrix that is explained by the model [68]. AGFI ("Adjusted Goodness of Fit Index") corrects GFI for the number of degree of freedom (df). These indicators vary theoretically between 0 and 1. Values of at least 0.90 indicate a good adjustment. PGFI ("Parsimony Goodness of Fit Index") is the product of GFI and the ratio accounting for the economic aspect of the model, i.e., (df of the model) /(df of the basic model). PGFI supports the more economic models.

The RMR ("Root Mean square Residual) estimates the average of residues. Its interpretation is made possible if its initial matrix is a correlation matrix. SRMR corresponds to Standardized RMR it is interpretable if its initial matrix is a variance/covariance matrix.

Approaches based on the gap between observed and estimated. The following indicators try to standardize the indicators that are calculated from the differences between the observed matrix and the estimated matrix. These indexes test the "badness" of fit of the model. FMIN is the minimum value of the difference function between in the sample reached at the time of the estimate and the estimated matrix. F0 is the value of the difference function in the parent population. It is especially used in the form of the RMSEA ("Root Mean Square Error of Approximation") presented in Table 1. RMSEA (Root Mean Square Error Approximation) represents the average difference, by degree of freedom, expected in the population and not in the sample (Relation 2). RMSEA is particularly important because it is relatively independent of sample size and the number of parameters used in the model [69]. It is given with a confidence interval of 90%.

RMSEA combines several properties. It is an index of parsimonious adequacy because RMSEA includes a correction for the complexity of the model. It will not seek to approach a centred distribution (as does chi-square), rather, to approach a non-centred distribution, which does not require having a true null hypothesis. It is a test of close fit based on the hypothesis that in the parent population, the gap between the model and the data can be different from zero (a hypothesis of exact fit that is very restrictive). This means that an adequate fit between the model and the observed sample need not be perfect (close fit). It is an index of badness of fit, i.e. a bad specification of the model: A zero indicates a good fit, whereas higher values indicate a bad fit. Thus, the smaller RMSEA, the better. Otherwise, the model will be only a rough approximation of reality and not an exact copy. Because RMSEA is relatively

independent of sample size and the number of parameters, its values can generally be considered independently of the data and model characteristics. These properties allowed Browne and Cudeck (1993) to give general reference thresholds, now widely accepted by the SEM community: RMSEA < 0.05 indicates a good adequacy of the model and the null hypothesis will not be rejected; 0.05 < RMSEA < 0.08 constitutes a reasonable error of approximation and for RMSEA > 0.10, the model's adequacy is considered to be weak.

Indexes of parsimony (absolute and incremental corrected indexes). Indexes of parsimony balance two opposite constraints: they minimize the number of coefficients to be estimated and maximize the degree of fit of the model. They are used to prevent models from having too many parameters; they detect if a bad fit of the model does not arise from a lack of free parameters (too many fixed parameters). Thus, these indexes can be used to find the most parsimonious of several equivalent models. They measure the lack of fit due to the model constraints. For example, the scaled chi-square of Jöreskog [68] can detect models that are over or under adjusted: It indicates which of several alternative models is the most parsimonious. Table 1 shows key values for interpreting the fit of the model to the data.

Insert table 1 about here

Comparing models

The current canonical approach of SEM, of testing only one model, has been evolving towards a model comparison approach. The models in competition can be defined before the experimentation (theory driven process) or can be constructed through an improvement process while analyzing the data (data driven process). This way of handling the models tends

to reduce the distinction between a purely confirmatory approach (only one model tested) and exploratory approach (like in classical factor analysis where the aim is to explore the data structure without imposing any constraint on the number of factors or on the relationships between the observed variables and the factors). The interest in comparing several models and/or the existence of equivalent models (models for which the fit indicators have exactly the same values hence the relationships between variables represent different causal networks) has lead to a new paradigm that compares a set of plausible models. This new paradigm will be discussed after describing the classical model comparison approach starting with the data driven approach: the improving of a model.

Re-specifying the theoretical model (indexes of modification).Questions arise when we try to improve the fit of the structural model to empirical data or to the theory which underlies it. Relations supported by theoretical assumptions can not be modified while relations supported by empirical assumptions, based on previous results, can be modified if theoretically justified. Indexes exist to identify variables that are worth the effort of re-specification. Indeed, it is possible to add or withdraw paths based on empirical criteria. Maximum Lagrangian multipliers estimate of how much the chi-square test would decrease if a particular path that was fixed at zero were introduced in the model. However, maximum Lagrangian multipliers are sensitive to the sample size; changes are more likely to be significant with large samples. If the amplitude of the expected effect is low, the statistical significance of this indicator might reflect size rather than the amplitude of required effect. Thus, modification indexes represent the amount of decrease in chi square if one or more additional parameters of covariance were considered. A value of the modification index higher than 3.84 (the value of the chi square distribution with one degree of freedom corresponding to a probability of .05) would show a significant chi-square reduction if we added a specified relation in model.

Indexes of modification are sometimes delicate even if their significance is easy to understand,. Initially, such indicators are provided for all the parameters not introduced within the model, including aberrant parameters such as covariances between latent variables and error variances. More dangerous are the parameters that seem plausible but cannot be interpreted. Thus, error covariances often have high values. In certain cases, the covariances make sense, for example when testing and retesting, when conducting consecutive tests, halo effects, etc. In such cases they can be introduced into the model. However, the user often cannot give a reliable interpretation of the error covariance between observed variables. Being able to interpret the error covariance between observed variables is more important than decreasing the χ^2 value in order to accept the model. In addition, the values of the indexes are not independent: if we add one parameter not yet introduced into the model, the indexes of modification of the remaining parameters will be modified. Adding parameters whose indexes of modification are not very high may not make much of a change in the quality of fit of the model. Moreover, adding addition parameters will adversely affect indicators that asses the economy of the model. Thus, it is important:

- To introduce only parameters which are justified on theoretical and/or methodological grounds.

- Not to introduce several parameters at once, but rather, to introduce one by one new parameters for which the index of modification is high and interpretable, in order to see their effect on the remaining indexes of modification of the other parameters that were not yet introduced [65].

Examining the standardized residues of the covariance matrix might suggest that there is an error in the predictions of the model. Residues exceeding $\pm$ 2.58 (the value of the Z distribution for p= .01), should be regarded as significant. Residues with high values indicate

that the relationship between the two variables is insufficiently accounted for by the model. Such high residual values can suggest the relationship may be improved by adding parameters (representing direct and/or indirect effects) to the model.

The main drawback of these data driven techniques is that they result in strong biases in the fit indicator values and their interpretation. A strong test of the re-specification would be to test the improved model with data from a new sample, a difficult and costly effort. One could instead, divide the sample in two sub-samples: use one to get information for improving the model and the other to test the improved model.

Comparing models defined *a priori*. The simplest situation is the comparison of nested models. Model A is said to be nested in a model B when it is possible to pass from model A to model B only by freeing some parameters. SEM can explicitly compare the adequacy of nested models. The two, nested, models can be easily specified by most software by introducing constraints on the parameters. For example, if we want to compare two models, one of which contains a parameter and the other does not, one simply leaves the parameter free in the one model and constrains is to equal zero in the other. The significance of the difference between the χ^2 of the two nested models is tested. The number of degrees of freedom of the test is the difference between the numbers of df of the two models. If the test is significant, the Ho of equivalence can be rejected and one can conclude that one of the models more adequately fits the data: Notice, this conclusion is true only if the nesting model fits the data, i.e., the nesting model is correct.

Comparing independent groups. Between-group comparisons are a precious feature of SEM for fMRI studies [36,39,43,53,65,70-72], as it provides the means to test the effect of different experimental conditions when the data come from independent groups. We invite the reader to consult the article of de Marco et al.[4], noted above, for an example. That study concerned

changes in effective connectivity that occurred during incidental and intentional perception of the same fearful faces. Briefly, SEM was used to construct a pathway model that could account for fMRI data in all regions of interest for the perception of fearful faces. Adjusted fMRI signals in the regions of interest extracted from the data set were entered as variables in the structural model, which was assessed by minimizing the difference between observed and predicted covariances of the fMRI data, according to the maximum-likelihood algorithm. In the analysis, the variables are considered in terms of their covariance structure.

Typically, in SEM, between-group comparisons use a model comparison approach . Instead of analyzing only one covariance matrix, the analysis uses stacked matrices, one for each experimental group. In a non-constrained model, there is one value for each path coefficient and each group. Also in this approach, constraints are introduced on the estimations of the parameters, usually constraints of equality; there is only one value for each parameter and it is common to all groups. For example, the estimate of the values of a path coefficient can be constrained to be equal in the experimental groups. The between-group comparison, therefore, amounts to a comparison between a constrained (fixed parameters) and a non-constrained (free parameters) model. The model-comparison proceeds as follows. First, the algorithm estimates the specific parameter value for each path and each group, and the model is examined to ensure that the hypothesized causality network fits the data. Then, the parameters for each group are forced to equality. Finally, the significance of the difference between the two models is tested. Since the constrained model is nested in the non-constrained model, the χ^2 of the difference between the χ^2 of the two models can be used to test the significance of the difference between them. If the models differ, then it must be that the experimental groups differ in terms of path connectivity. In other words, if the fit of the constrained model (using common values for the parameters) is significantly different from the fit of the unconstrained model (using specific/different values for the parameters in each group) then we can conclude

that there is a difference between the parameters (i.e. the causal pathways) of the experimental groups and the differences must be taken into consideration to model the covariance matrix.

In order to account for the overall difference between the experimental groups, each pair of parameters from the experimental groups can be compared (pair-wise comparison). The test, which follows a Z-distribution, is based on the differences between the parameter values divided by the standard error of measurement of these differences [73].

Procedure for model specification

This approach was first mentioned by McCallum (Mac Callum, 1986; MacCallum, Wegener, Uchino, & Fabrigar, 1993). The model specification approach considers a set of models comprising common or obligatory parameters (they will be present in all the studied models) and optional parameters. Optional parameters may or may not be present in some of the models. If the number of optional parameters is p, then the total number of different models will be 2^p: for one parameter there are two models (with and without the parameter), for two parameters there are four models, etc. Once the not-identified models or models comprising inadmissible values are eliminated, we obtain a set of "plausible" models that contain the obligatory and optional parameters. To sort them out we will use the Kullback and Leibler (Burnham & Anderson, 1998) approach that is based on information theory. The issue is to retain the best model in terms of the trade off between the quality of the adequacy of fit of the model to the data and the economic character (number of parameters) of the model. To do so we will use the indicators built within the framework of the information theory, in particular the criterion of information of Akaike (Akaike Information Criterion; about the use of this indicator and model selection in SEM, see Raftery, 1993). Because they are "badness of fit indicators", high values indicate bad models. The approach is particularly heuristic: It bridges clearly exploratory approaches (there is no *a priori* model) and clearly confirmatory

approaches (there is only one *a priori* model) by exploring "plausible" models. The essential idea of the information-theory data analysis is that, although the "real" model behind the data is not knowable, we can construct models that could plausibly generate the data. Plausibility models are assessed by the information that they take into account.

Choosing an algorithm of estimate. The main softwares propose different algorithms to estimate the parameters. The maximum likelihood algorithm is most frequently used in structural modelling because its parameter estimates are highly robust. Nevertheless, questions of estimate bias caused by variations in the conditions of validity (i.e., multinormality of the data), reports of the absence of convergence and the occurrence of aberrant estimates (e.g., negative variance) have resulted in the development of other estimation algorithms. Some of them provide reliable starting values for other more sophisticated algorithms, such as maximum likelihood algorithms (e.g., Generalized Least Squares). Others circumvent the necessity of the multinormality assumption, e.g., skewed distributions are not problematical. The "Weighted Least Squares" or "Asymptotic Distribution Free" algorithm estimates weights for each covariance to correct for asymmetric characteristics of the distributions of each pair of variables: It gives a more reliable estimate of the parameters and of the tests of fit when the distributions are asymmetrical (Browne, 1984). Weighted algorithms require large sample sizes. The difficulties and biases of estimate noted earlier, as well as the absence of tests of inference for certain indicators (for example the coefficient of determination R), resulted in including bootstrapping procedures within the software of structural models. Bootstrapping procedures have the advantage of providing distributions for all the parameters, distributions taking into account the specificities of the data. As Bollen and Stine (1993) showed, the relevance of bootstrapping must be clear. Finally, 20 years ago, the Bayesian approach appeared to provide a pragmatic improvements in the procedures for estimating parameters within the framework of complex models (e.g.,

multilevel models for example) (Gelman, Carlin, Stern and Rubin, 2004). In the case of the structural models, the Bayesian approach can yield robust estimates, especially with small samples. It also allows a priori constraints to be placed on the distributions in order to avoid problems of estimate (e.g., requiring the variances to be positive). We have such a Bayesian procedure in version 6 of AMOS (the "Metropolis" algorithm). The algorithm obtains an *a posteriori* distribution of all the usually estimated parameters. We can also obtain, in supplement, the distribution of statistics of complex parameters, such as direct and indirect effects, which are not provided by the usual algorithms, for example as the maximum likelihood algorithm.

Conclusion

It should be remembered that SEM is a confirmatory model of analysis and/or a theoretical model: Thus, each relation in the model that is tested must be grounded in a theory or on empirical bases. The theoretical framework must be rigorously defined, the research question must be clearly posed, the issues correctly argued and the hypotheses no falsifiable.

Insert figure 2 about here

Figure 2 summarizes the main steps of the SEM method, along three paths. The first one (left) indicates the theoretical process: from the theory and the anatomical hypothesis to the formalization of a structural model and its implantation in SEM software using the different languages available (path diagram, matrix language or equations). The second one (top)

shows the data collection steps, from fMRI data collection to the implantation of these data in a file readable by the software used. The third path (right) indicates the parameter estimation and the interpretation process: from the model and data to the algorithm chosen to estimate parameters that allow computing the theoretical covariance matrix that is compared to the observed matrix. The gap between the matrices serves as indicator to compute the χ^2 test that, together with other indicators noted in the text, are used to decide whether the model adequately describes the data. Acceptance of the model confirms the theory. Sometimes, the value of the indicator can lead to an improvement of the model. If so, it is important to go back to the different steps, especially step one. Rejection of the model, while probably not a publishable outcome, can be quite heuristic as it can give insight about why a theory is not working in a specific case. The approach presented in Figure 2 is the canonical one, aimed at testing one model or comparing few models defined *a priori* (for example in a group comparison approach). New approaches, such as model specification, may add new perspectives but they will not modify the path one must take from theory construction, to data collection, to decision making.

Acknowledgments

References

[1]   Morris JS, Friston KJ, Buchel C, et al. A neuromodulatory role for the human amygdala in processing emotional facial expressions. Brain 1998;121 (Pt 1):47-57.

[2]   Killgore WD, Yurgelun-Todd DA. Activation of the amygdala and anterior cingulate during nonconscious processing of sad versus happy faces. Neuroimage 2004;21(4):1215-1223.

[3]   Wild B, Erb M, Eyb M, Bartels M, Grodd W. Why are smiles contagious? An fMRI study of the interaction between perception of facial affect and facial movements. Psychiatry Res 2003;123(1):17-36.

[4]   de Marco G, de Bonis M, Vrignaud P, Henry-Feugeas MC, Peretti I. Changes in effective connectivity during incidental and intentional perception of fearful faces. Neuroimage 2006;30(3):1030-1037.

[5]   Pallier C, Dehaene S, Poline JB, et al. Brain imaging of language plasticity in adopted adults: can a second language replace the first? Cereb Cortex 2003;13(2):155-161.

[6]   Lee L, Harrison L, Mechelli A. A report of the functional connectivity. Workshop,Dusseldorf 2003;457-465.

[7]   Mesulam MM. Large-scale neurocognitive networks and distributed processing for attention, language, and memory. Ann Neurol 1990;28(5):597-613.

[8]   Fu CH, McIntosh AR, Kim J, et al. Modulation of effective connectivity by cognitive demand in phonological verbal fluency. Neuroimage 2006;30(1):266-271.

[9]   Vitali P, Abutalebi J, Tettamanti M, et al. Generating animal and tool names: an fMRI study of effective connectivity. Brain Lang 2005;93(1):32-45.

[10]   Friston KJ, Buechel C, Fink GR, Morris J, Rolls E, Dolan RJ. Psychophysiological and modulatory interactions in neuroimaging. Neuroimage 1997;6(3):218-229.

[11]   Friston K. Functional and Effective Connectivity in Neuroimaging: A Synthesis. Human Brain Mapping 1994;2:56 - 78.

[12]   Heeger DJ, Ress D. What does fMRI tell us about neuronal activity? Nat Rev Neurosci 2002;3(2):142-151.

[13]   Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A. Neurophysiological investigation of the basis of the fMRI signal. Nature 2001;412(6843):150-157.

[14]   Hamalainen M, Hari R, Ilmoniemi R, Knuutila J, Lounasmaa O. Magnetoencephalography - theory, instrumentation, and application to noninvasive studies of the working human brain. Reviews of Modern Physics 1993;65:413-497.

[15]   Nunez PL, Silberstein RB. On the relationship of synaptic activity to macroscopic measurements: does co-registration of EEG with fMRI make sense? Brain Topogr 2000;13(2):79-96.

[16]   Attwell D, Iadecola C. The neural basis of functional brain imaging signals. Trends Neurosci 2002;25(12):621-625.

[17]   Li TQ, Haefelin TN, Chan B, et al. Assessment of hemodynamic response during focal neural activity in human using bolus tracking, arterial spin labeling and BOLD techniques. Neuroimage 2000;12(4):442-451.

[18]   Shulman RG, Rothman DL, Behar KL, Hyder F. Energetic basis of brain activity: implications for neuroimaging. Trends Neurosci 2004;27(8):489-495.

[19]   Aubert A, Costalat R. A model of the coupling between brain electrical activity, metabolism, and hemodynamics: application to the interpretation of functional neuroimaging. Neuroimage 2002;17(3):1162-1181.

[20]   Horwitz B, Friston KJ, Taylor JG. Neural modeling and functional brain imaging: an overview. Neural Netw 2000;13(8-9):829-846.

[21]   Goldman-Rakic PS. Topography of cognition: parallel distributed networks in primate association cortex. Annu Rev Neurosci 1988;11:137-156.

[22]   Varela FJ. Resonant cell assemblies: a new approach to cognitive functions and neuronal synchrony. Biol Res 1995;28(1):81-95.

[23]   Van Essen DC, Anderson CH, Felleman DJ. Information processing in the primate visual system: an integrated systems perspective. Science 1992;255(5043):419-423.

[24]   Phillips WA, Singer W. In search of common foundations for cortical computation. Behav Brain Sci 1997;20(4):657-683; discussion 683-722.

[25]   Varela F, Lachaux JP, Rodriguez E, Martinerie J. The brainweb: phase synchronization and large-scale integration. Nat Rev Neurosci 2001;2(4):229-239.

[26]   Hebb D. The organization of Behavior: A Neurophysiological Theory. Wiley, New York 1949.

[27]   Girard P, Hupe JM, Bullier J. Feedforward and feedback connections between areas V1 and V2 of the monkey have similar rapid conduction velocities. J Neurophysiol 2001;85(3):1328-1331.

[28]   Gray CM. The temporal correlation hypothesis of visual feature integration: still alive and well. Neuron 1999;24(1):31-47, 111-125.

[29]   Tononi G, Edelman GM, Sporns O. Complexity and coherence:integrating information in the brain. Trends in Cognitive Sciences 1998;2:474-484.

[30]   Sporns O, Chialvo DR, Kaiser M, Hilgetag CC. Organization, development and function of complex brain networks. Trends Cogn Sci 2004;8(9):418-425.

[31]   Gerstein GL, Perkel DH. Simultaneously recorded trains of action potentials: analysis and functional interpretation. Science 1969;164(881):828-830.

[32]   Friston KJ. Imaging neuroscience: principles or maps? Proc Natl Acad Sci U S A 1998;95(3):796-802.

[33]   McIntosh A, Gonzalez-Lima F. Structural equation modeling and its application to network analysis in functional brain imaging. Human Brain Mapping 1994;2:2-22.

[34]   Aersten A, Preissl H. Dynamics of activity and connectivity in physiological neuronal networks. In:Non Linear Dynamics and Neuronal Networks. (Schuster HG, ed) 1991;New York, VCH Publishers Inc.:pp 281-302.

[35]   Buchel C, Friston K. Assessing interactions among neuronal systems using functional neuroimaging. Neural Netw 2000;13(8-9):871-882.

[36]   Au Duong MV, Boulanouar K, Audoin B, et al. Modulation of effective connectivity inside the working memory network in patients at the earliest stage of multiple sclerosis. Neuroimage 2005;24(2):533-538.

[37]   Glabus MF, Horwitz B, Holt JL, et al. Interindividual differences in functional interactions among prefrontal, parietal and parahippocampal regions during working memory. Cereb Cortex 2003;13(12):1352-1361.

[38]   Schlosser RG, Wagner G, Sauer H. Assessing the working memory network: studies with functional magnetic resonance imaging and structural equation modeling. Neuroscience 2006;139(1):91-103.

[39]   Schlosser R, Gesierich T, Kaufmann B, et al. Altered effective connectivity during working memory performance in schizophrenia: a study with fMRI and structural equation modeling. Neuroimage 2003;19(3):751-763.

[40]   Kondo H, Morishita M, Osaka N, Osaka M, Fukuyama H, Shibasaki H. Functional roles of the cingulo-frontal network in performance on working memory. Neuroimage 2004;21(1):2-14.

[41]   Krause JB, Taylor JG, Schmidt D, Hautzel H, Mottaghy FM, Muller-Gartner HW. Imaging and neural modelling in episodic and working memory processes. Neural Netw 2000;13(8-9):847-859.

[42]   Rowe J, Friston K, Frackowiak R, Passingham R. Attention to action: specific modulation of corticocortical interactions in humans. Neuroimage 2002;17(2):988-998.

[43]   Mottaghy FM, Willmes K, Horwitz B, Muller HW, Krause BJ, Sturm W. Systems level modeling of a neuronal network subserving intrinsic alertness. Neuroimage 2006;29(1):225-233.

[44] Erickson KI, Ringo Ho MH, Colcombe SJ, Kramer AF. A structural equation modeling analysis of attentional control: an event-related fMRI study. **Brain Res Cogn Brain Res** 2005;Mar;22(3):349-357.

[45] Rajah MN, McIntosh AR, Grady CL. Frontotemporal interactions in face encoding and recognition. Brain Res Cogn Brain Res 1999;8(3):259-269.

[46] Rogers BP, Carew JD, Meyerand ME. Hemispheric asymmetry in supplementary motor area connectivity during unilateral finger movements. Neuroimage 2004;22(2):855-859.

[47] Grafton ST, Sutton J, Couldwell W, Lew M, Waters C. Network Analysis of Motor System Connectivity in Parkinson's Disease:Modulation of Thalamocortical Interactions After Pallidotomy. Human Brain Mapping 1994;2:45-55.

[48] Toni I, Rowe J, Stephan KE, Passingham RE. Changes of cortico-striatal effective connectivity during visuomotor learning. Cereb Cortex 2002;12(10):1040-1047.

[49] Taniwaki T, Okayama A, Yoshiura T, et al. Functional network of the basal ganglia and cerebellar motor loops in vivo: different activation patterns between self-initiated and externally triggered movements. Neuroimage 2006;31(2):745-753.

[50] Zhuang J, LaConte S, Peltier S, Zhang K, Hu X. Connectivity exploration with structural equation modeling: an fMRI study of bimanual motor coordination. NeuroImage 2004;25(2):462-470.

[51] Kline R. Principles and Practice of Structural Equation Modeling (Methodology In The Social Sciences). The Guilford Press 2004;2nd edition.

[52] Byrne B. Structural Equation Modeling With Amos: BASIC Concepts, Applications, and Programming. Deliver, Lawrence Erlbaum Associates 2001;Reprint edition.

[53] Buchel C, Friston KJ. Modulation of connectivity in visual pathways by attention: cortical interactions evaluated with structural equation modelling and fMRI. Cereb Cortex 1997;7(8):768-778.

[54] Horwitz B. The elusive concept of brain connectivity. Neuroimage 2003;19(2 Pt 1):466-470.

[55] Gonzalez-Lima F, McIntosh A. Analysis of neural network interactions related to associative learning using structural equation modeling. Mathematics and Computers in Simulation 1995;40,(1-2):115 - 140.

[56] McIntosh A. Towards a network theory of cognition. Neural Networks 2001;13:861-876.

[57] Bollen K. With new incremental structural index for general equation models made. Sociological Methods and Research 1989;17:303-316.

[58] Arbuckle J. Amos 6.0 Update to the Amos User' S Guides. SmallWaters Corporation, Chicago. 2005.

[59] Gavrilescu M, Stuart GW, Waites A, Jackson G, Svalbe ID, Egan GF. Changes in effective connectivity models in the presence of task-correlated motion: an fMRI study. Hum Brain Mapp 2004;21(2):49-63.

[60] Goncalves MS, Hall DA. Connectivity analysis with structural equation modelling: an example of the effects of voxel selection. Neuroimage 2003;20(3):1455-1467.

[61] Hu L, Bentler P. Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives. Structural Equation Modeling 1999;6:1-55.

[62] Fan X, Thompson B, Wang L. Effects of sample size, estimation methods and model specification on structural equation modelling. Structural Equation Modeling 1999;6:56-83.

[63] Bentler P, Freeman E. Tests for stability in linear structural equation systems. Psychometrika 1983;48:143-145.

[64]   Bentler P, Bonett D. Significance tests and goodness off made off in the analysis covariance structure. Psychological Bulletin 1980;88:588-606.

[65]   Bullmore E, Horwitz B, Honey G, Brammer M, Williams S, Sharma T. How good is good enough in path analysis of fMRI data? Neuroimage 2000;11(4):289-301.

[66]   Bollen KA. A new incremental fit index for general structural equation models. Sociological Methods and Research 1989b;17:303-316.

[67]   Tucker LR, Lewis C. The reliability coefficient for maximum likelihood factor analysis. Psychometrika 1973;38:1-10.

[68]   Jöreskog KG, Sörbom D. LISREL 8.5 User's Reference Guide. Chicago, Ill.: Scientific Software International. 2000.

[69]   Browne HW, Cudeck R. Alternative ways of assessing model fit. In KA Bollen & J S Long (Eds), Testing Structural Equations Models Newbury Park, CA : Sage Publications 1993.

[70]   Erickson KI, Ringo Ho MH, Colcombe SJ, Kramer AF. A structural equation modeling analysis of attentional control: an event-related fMRI study. Brain Res Cogn Brain Res 2005;22(3):349-357.

[71]   McIntosh AR, Gonzalez-Lima F. Structural Equation Modeling and Its Application to Network analysis in Functional Brain Imaging. Hum Brain Mapp 1994;2:2-22.

[72]   Penny WD, Stephan KE, Mechelli A, Friston KJ. Modelling functional integration: a comparison of structural equation and dynamic causal models. Neuroimage 2004;23 Suppl 1:S264-274.

[73]   Arbuckle J. Amos 5.0 Update to the Amos User' S Guides. SmallWaters Corporation, Chicago 2003.

A)
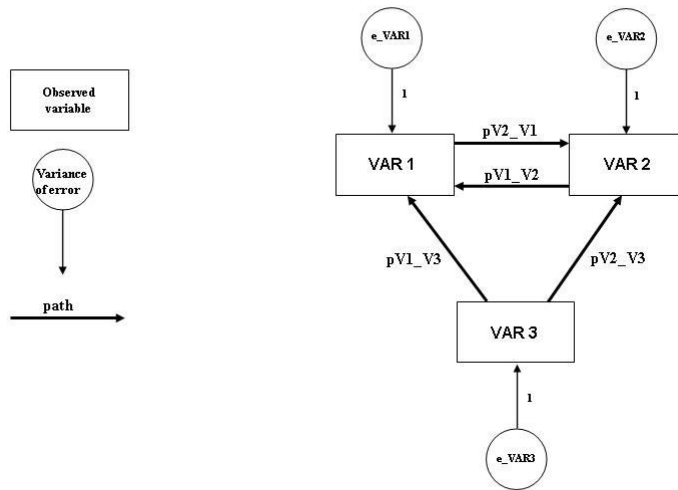
B)
```
VAR1 <---------- VAR3 (pV1_V3)
VAR1 <---------- VAR2 (pV1_V2)
VAR2 <---------- VAR3 (pV2_V3)
VAR2 <---------- VAR1 (pV2_V1)
VAR1 <--------- e_var1 (1)
VAR2 <--------- e_var2 (1)
```

Figure 1: A) Model with three regions. The three observed variables (VAR1, VAR2, and VAR3) are represented by rectangles; each of them is connected with an error term (variance not explained by the model). B) The relationships between the observed variables are described by an arrow pointing towards the dependent variable. Brackets after the relation represent the constraint on the parameter, e.g. (pV1_V2)- to indicate a free parameter and (1)-to indicate a fixed parameter. Some parameters are constrained or fixed to equal 1.0 (in this example the parameter relating the error variance to its variable) so that the system of equations can be identified. (Adapted from de Marco, et al. (2006). "Changes in effective connectivity during incidental and intentional perception of fearful faces." Neuroimage **30:** 1030-1037).
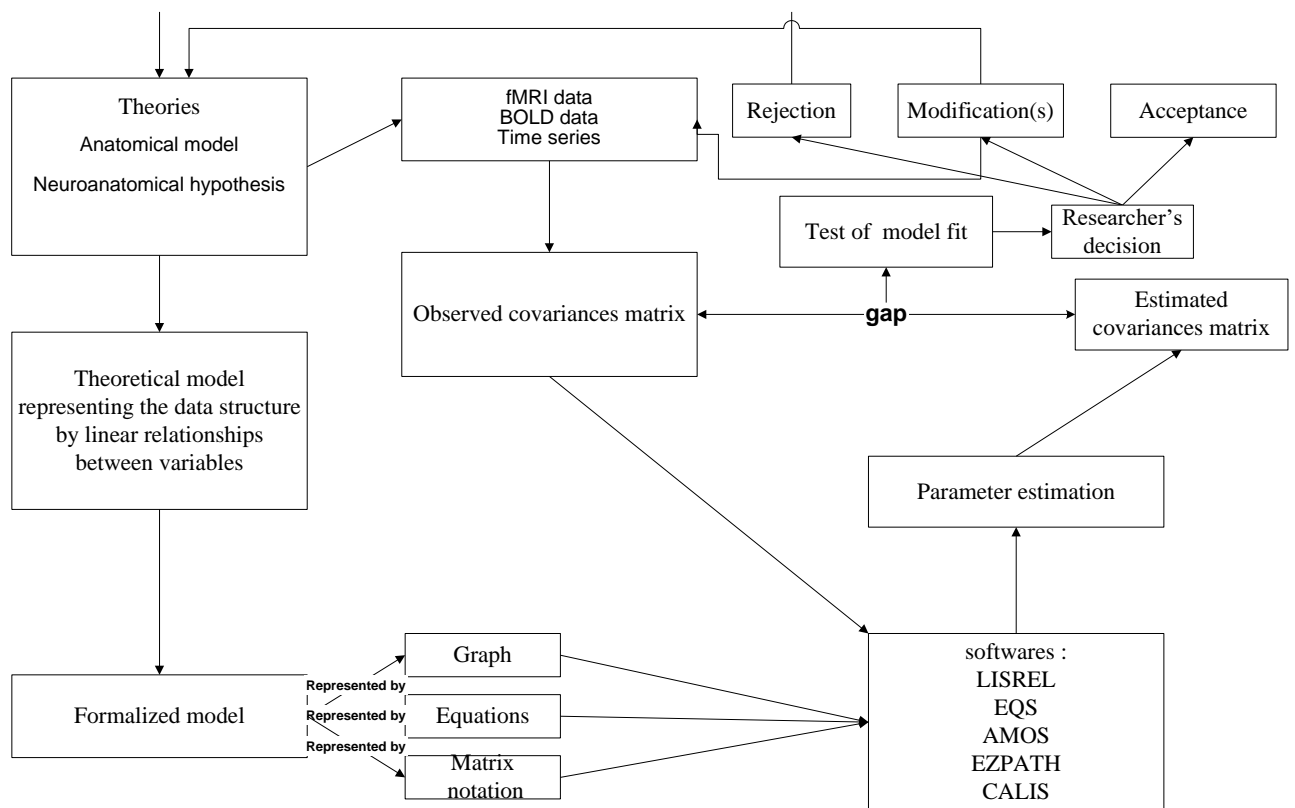
Figure 2: Diagram synthesizing the main steps of a SEM analysis.

Types of Indexes	Indexes	Prefered Values
Absolute fit indexes	χ^2 et scaled χ^2	Smaller possible (see p value)
	GFI, AGFI	> 0.90
	-	-
	PNNI	> 0.95
	RMR et SRMR	Smaller possible (close to 0)
	RMSEA	< .08 or better < .05
Incremental fit indexes	NFI and ρ of Bollen	> .90
	TLI (NNFI)	> .90
	CFI and RNI	> .90
Parsimony indexes	Scaled χ^2	Smaller possible, between 1, 2-3, 5
	AIC	Smaller possible (comparison)
	PNFI and PGFI	Stronger possible (comparison)

Table 1: key values of interpretation of indexes