Comments and Controversies

Understanding DCM: Ten simple rules for the clinician

Joshua Kahan*, Tom Foltynie

Sobell Department for Motor Neuroscience and Movement Disorders, UCL Institute of Neurology, London, UK

ARTICLE INFO

ABSTRACT

Despite almost a decade since the introduction of Dynamic Causal Modelling (DCM), there remains some confusion within the wider neuroimaging, neuroscience and clinical communities as to what DCM studies are probing, and what all the jargon means. We provide ten simple rules, and a theoretical example to gently introduce the reader to the rationale behind DCM analyses, and how one should consider neuroimaging data and experiments that use DCM. It is deliberately written as a primer or orientation for non-technical imaging neuroscientists or clinicians who have had to contend with the technical intricacies of understanding DCM.

© 2013 Elsevier Inc. All rights reserved.

Introduction

"As science becomes more complex, the argument goes, an ever-increasing amount of specialist jargon is required to describe it precisely. Even if this is true, however, technical terminology can be explained, and it need not present an insurmountable problem to the scientifically literate reader." (Anon., 2000)

As neuroimaging research has matured, researchers have adopted complex methods to help them infer the most from their datasets. A recent advance has seen authors attempt to construct realistic models of their data in order to best explain how they were generated. This piece aims to provide a short introduction to one such method; namely, dynamic causal modelling (DCM) (Friston et al., 2003), with a view to demystifying the technical language that may prevent a clinician (or anyone who forgot to do a graduate degree in engineering) from critiquing these studies. The only mathematical prerequisite to this report is that the reader understands the difference between a series of data points in time, and the rate of change of the data with respect to time (i.e. the gradient of the plot of data vs. time). This is important as DCM models the changes from one time point to the next to make sense of how brain regions impact on each other's neural activity (see Rule V). This piece aims to (i) frame how one should consider neuroimaging data (with a strong emphasis on fMRI), (ii) how this relates to the rationale of DCM, and (iii) gently introduce the reader to the, at first, obscure language that attends DCM. As an introductory text, we have chosen to focus on fMRI data as DCM for fMRI models neuronal interactions in the most generic form possible. However, we briefly discuss later how these concepts translate to electrophysiological datasets.

I. Neural activity causes behaviour

Functional neuroimaging studies rest on the assumption that measurable behaviours are caused by neural activity, and that experimental stimuli change neural activity. This activity occurs throughout the brain, especially so in *spatially distributed* regions that are specialised to process the stimuli or behaviour in question. These different regions are often considered to be distinct processing modules. This concept is referred to as "functional segregation". However, the neuronal processing is more than the sum of activity in these modules — the regions need to communicate; i.e., pass information to each other in an optimal way in order to process stimuli and execute optimal behaviour. This concept is known as "functional integration".

Publication statistics suggest that there is an increasing focus on connectivity analyses (Friston, 2011), as the field has turned from addressing questions of functional segregation to investigating functional integration. Quantifying connectivity based on the fMRI blood oxygen level dependent (BOLD) signal can be achieved through various methods. A dichotomy exists in the literature between "functional connectivity" and "effective connectivity". Functional and effective connectivity both describe measures that are purely based on the BOLD signal of distant regions. This is not the same as "structural connectivity" measured through "tractography" or diffusion tensor imaging (DTI) white matter assessments (for primer, see Mori and Zhang, 2006) — although structural and functional/effective connectivity are not independent as such (Stephan et al., 2009b). In order to best understand the difference, and how DCM works, one must first consider what the BOLD signal is and how data are generated in fMRI experiments.

II. Neuroimaging data is generated by downstream effects of neural activity

The BOLD signal has long been the subject of much discussion, the details of which are covered in depth elsewhere (Heeger and Ress,

---

* Corresponding author.
   E-mail address: joshua.kahan@ucl.ac.uk (J. Kahan).

2002). In short, it results from differences in magnetic field homogeneity induced by the differential magnetic properties of oxy- and deoxyhaemoglobin. The deoxygenation of haemoglobin is considered an indirect measure of synaptic activity, under the assumption that neurons require more oxygen, and thus more blood, when they are metabolically more active. However, the events leading to BOLD signal alterations are still not precisely defined. Therefore, physiologically validated models are used to guide understanding of the process (more in Rule V). It is also important to remember that these discrete cortical processing units are not simply composed of excitable principal cells; cortical columns are arranged in populations of both glutamatergic and GABAergic populations, simultaneously displaying excitation and inhibition on receipt of a distant input (sometimes referred to as "canonical microcircuits" (Bastos et al., 2012; Douglas et al., 1989) — the implications of this on electrophysiological DCMs are discussed in Box 1). The acknowledgement that the BOLD signal is not a direct measure of neural activity, although seemingly obvious, is fundamental (Logothetis, 2008), and can be confused in reporting results (the same would be true of any other indirect measure of neural activity — PET, SPECT etc.).

One could consider the BOLD signal a dependent, *measurable* (or observed) variable (y) of the underlying neural activity (z) that cannot be measured with fMRI (thus, the neural activity is an example of a "hidden state variable"). This concept has been used in all general linear model-based analysis of fMRI data, and is identical in DCM with one key difference (see footnote[1]).

III. Experimental manipulations can directly perturb neural activity

In a fictitious visual perception experiment, imagine a subject is placed in an MRI scanner and presented with a visual stimulus periodically [Fig. 1A]. The timings of stimulus-presentations are used to create an explanatory variable of the whole-brain BOLD data, and we find three clusters (that is, groups of voxels) are significantly explained by this explanatory variable (i.e. regions that show an altered response). Now imagine one knew the functional architecture of the highlighted brain regions; that is, how information propagates through these connected and functionally specialised regions to produce the observed results.

One such architecture could be as depicted in [Fig. 1B]. In this example, presentation of the visual stimulus u_1 may cause a change in the neural activity (z) of region 1 (z_1), which in turn causes a change in the measurable BOLD signal (y) of region 1 (y_1). Z_1 then causes an effect on z_2, mediated by the ("extrinsic" or between region) connection from 1 to 2 ($a_{2,1}$). Likewise, z_2 in turn causes an effect on z_3. In addition, these regions contain some self-inhibitory properties, mediated by self-connections ("intrinsic" or within region connections; e.g. $a_{1,1}$), preventing runaway outbursts of neural activity. Note that this chain of events only occurs with presentation of our visual stimulus, thus the model only explains the network dynamics in that instance. It does not tell us the dynamics when the stimulus is not present (e.g. at rest) or when the visual stimulus changes (to say, an emotional face). Obviously, this functional architecture is only one possible model of how our visual perception data are generated, there are a number of other equally plausible

---

[1] Standard general linear model (GLM) analyses rely on convolving a stimulus function (representing onsets and durations of stimuli — an assumed neural model) with a haemodynamic response function (HRF) to produce an explanatory variable (an exogenous input) that is used to identify brain regions related to the stimuli. This essentially treats voxels as isolated regions, testing to see which voxels are sensitive to the input. DCM also uses convolution models, with two key differences: Firstly, neural states (z) causing BOLD data (y) in DCM are sensitive to both exogenous inputs (as in GLM analyses) and afferents from other regions. Secondly, the convolution model implicit in DCM is nonlinear due to some mild nonlinearities in the haemodynamic response, dealt with in GLM analyses by including basis functions of the HRF (see Friston et al., 2003; Friston, 2002). In short, DCM is simply a generalisation of the convolution models used in all GLM analyses, the only important difference is that DCM allows the experimenter to additionally consider the effects of other regions on their neural model.

Box 1
Extending these concepts to EEG/MEG and LFP datasets.

Unlike fMRI data, EEG/MEG and LFP datasets have a rich temporal structure allowing much more complicated models of cortical function to be employed (David et al., 2005). Pioneered with evoked responses in EEG and MEG (David et al., 2006; Jansen and Rit, 1995), and later extended for steady-state responses and neural fields in EEG/MEG and LFP data (Moran et al., 2007, 2009), electrophysiological DCMs use physiologically plausible neural mass and neural field models as generative models of the observed data. These models embed each region with several neuronal subpopulations representing key constituents of the cortex. A variety of models differing with respect their structure (e.g. canonical microcircuit — Bastos et al., 2012), or how the model was formulated (see difference between convolution vs. conductance-based models — Moran et al., 2013) are available, dependent on the type of data collected and experimental setup. Most embed each region with excitatory pyramidal output neurons, inhibitory interneurons, and excitatory spiny input neurons (Fig. 4). Intrinsic connections are estimated for each region (in the case of Fig. 4, reciprocal connections between the pyramidal and inhibitory interneurons, reciprocal connections between the pyramidal and spiny input neurons, and a self-inhibitory connection in the inhibitory population). Extrinsic connections are also subdivided into forward, backward or lateral connections, each arriving as afferents to different subpopulations in accordance with primate connectivity patterns (Felleman and Van Essen, 1991). The firing rate of each subpopulation is treated as a hidden state, dependent on the average pre-synaptic inputs, post-synaptic membrane potential, constants summarising the biophysical membrane properties. The forward models used to map the neural activity to the observed data are discussed in depth elsewhere (see review — Moran et al., 2013). Putting aside the elegant mathematics, these DCMs can be used in exactly the same way as we have discussed. The coupling parameters (intrinsic or extrinsic) are estimated in the same manner, and models can be similarly compared with BMS. Importantly, given the additional complexity of the generative models employed, these DCMs can furnish estimates of much more subtle parameters than those estimated in DCM for fMRI. For example, these DCMs have been applied as a 'mathematical microscope' to probe neurotransmitter receptor function (Moran et al., 2011a), but can also be used in a similar way to DCM for fMRI to explore extrinsic coupling parameters (Marreiros et al., 2012; Moran et al., 2011b). For a practical guide to implementing DCM for these datasets, see (Litvak et al., 2011).

models that could underlie this network (this will be discussed in Rule IX). In a clinical population, the investigator here might be interested in understanding how the model architecture changes with the onset of acute disease (e.g. modelling responses following a stroke (Rehme et al., 2011)), explore how neurodegenerative processes or neurorehabilitation impacts on BOLD responses (Abutalebi et al., 2009).

IV. Functional connectivity describes statistical dependencies between regions

Having considered a mechanism by which the results in the visual perception example may have come about, let us return to the dichotomy between functional and effective connectivity. Functional connectivity
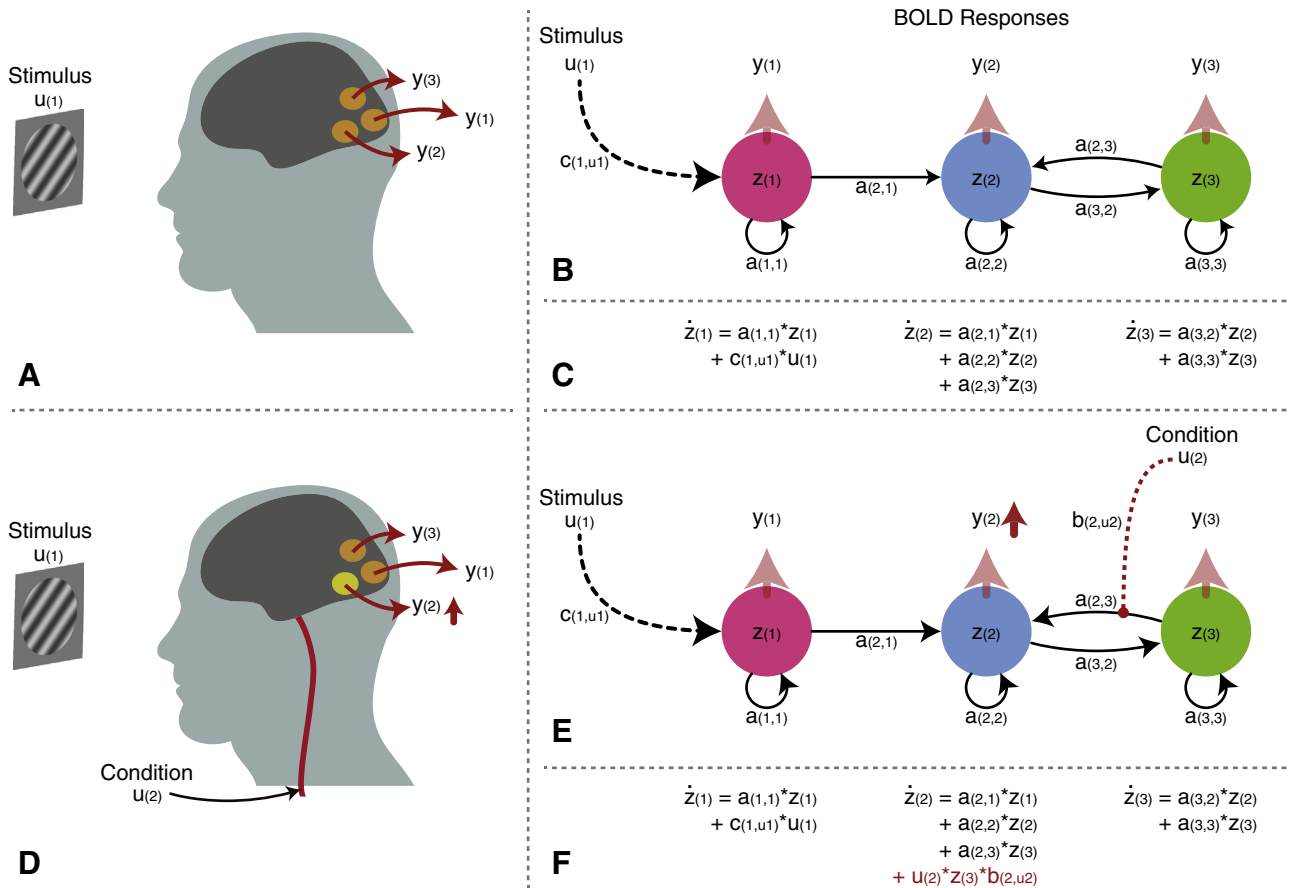
Fig. 1. Our fictitious visual perception fMRI experiment. This experiment consists of two experimental manipulations or 'inputs' (u); $u_{(1)}$ = visual stimulus input, $u_{(2)}$ = drug input. Coupling parameters denoted by $a_{(k,j)}$ represents the effective connectivity from node j to k, in the context of stimulus $u_{(1)}$. Parameters labelled $c_{(k,u)}$ denote the effect of the input on target node k. Parameters labelled $b_{(k,u)}$ denote the modulatory effect of the input on target connection. When the subject views the stimuli in the control state (A), three brain regions are shown to be related to the onset of the stimuli. The BOLD signal from each of these regions is represented by $y_{(1)}$, $y_{(2)}$ and $y_{(3)}$. (B) The BOLD signal from a single region is produced by a change in its underlying neural activity z, which has not been measured during the experiment. The underlying neural activity of region 1 has an effect on the underlying neural activity of region 2 (an *extrinsic* connection); the strength of this effect is determined by the value $a_{(2,1)}$. The underlying activity of region 1 also has some self-inhibitory dynamics (*intrinsic* connection), determined by the value $a_{(1,1)}$. The equations of motion (C) show the rate of change of each of the regions' underlying neural activity, $\dot{z}$. During the drug condition, an interaction is found in region 2 (D). This is modelled as a modulatory drug input having an effect on the strength of the extrinsic connection $a_{(2,3)}$ (E). The strength of this effect is determined by the value $b_{(2,u2)}$. (F) The equations of motion have thus changed to accommodate this additional modulatory input.

describes statistical dependencies between regions, i.e. correlations. In other words, which voxels in the brain display similar BOLD signal fluctuations over the course of a scan. BOLD data extracted from a seed region have been shown to be significantly correlated with BOLD data at distant regions that show similar functional specialisation, regardless of behavioural state (e.g. motor (Biswal et al., 1995) and visual networks (Lowe et al., 1998)). So called 'data-driven' decomposition analyses (independent component analyses etc.) have similarly been able to document consistent spatiotemporal relationships between brain regions of similar function in healthy subjects (Damoiseaux et al., 2006). Furthermore, changes in these measures have been reported with onset of disease and may give rise to biomarkers of neurological and psychiatric disease (Greicius et al., 2007; Zhang and Raichle, 2010). While these consistent findings are interesting, these analyses document a feature of fMRI data; namely, that similarly specialised regions display a similar y at a given time. However, functional connectivity analyses do not consider causal interactions within the network, i.e. how information propagates through brain regions (Friston, 1994, 2009, 2011; Smith, 2012). Returning to the visual perception example, imagine the BOLD signal at regions 1, 2 and 3, (y_1, y_2 and y_3,) are all strongly correlated; this could mean that they are connected in a number of different ways, or not connected at all, rather are all driven by a common input [Fig. 2]. In that respect, while of cartographic interest – in our example, identifying regions that may be engaged in visual processing – we are still

no closer to understanding how information is being exchanged between these regions. For example, in Parkinson's disease (PD), functional connectivity analyses have been able to document widespread changes in BOLD spatiotemporal relationships within motor and non-motor regions (Hacker et al., 2012; Jech et al., 2013; Wu et al., 2009). These results are interesting and may serve as reliable biomarkers, however conclusions from these studies provide limited insight into how such changes come about, or how they might be resolved.

V. Effective connectivity is defined by a model and corresponds to the *directed* influence that one region exerts on the rate of change of activity in another

In contrast, effective connectivity is defined as the directed influence one region has over another (Friston, 2009), and thus can supplement functional connectivity analyses in a complementary manner.

Measures of effective connectivity in DCM consider the rate of change of neural activity with respect to time ($\dot{z}$) in response to some incoming signal (be it from another brain region or an exogenous environmental stimulus). Most (possibly all) connections that exist in the brain are reciprocal, thus in order to mathematically dissect the effect of one region on another, we must consider that the
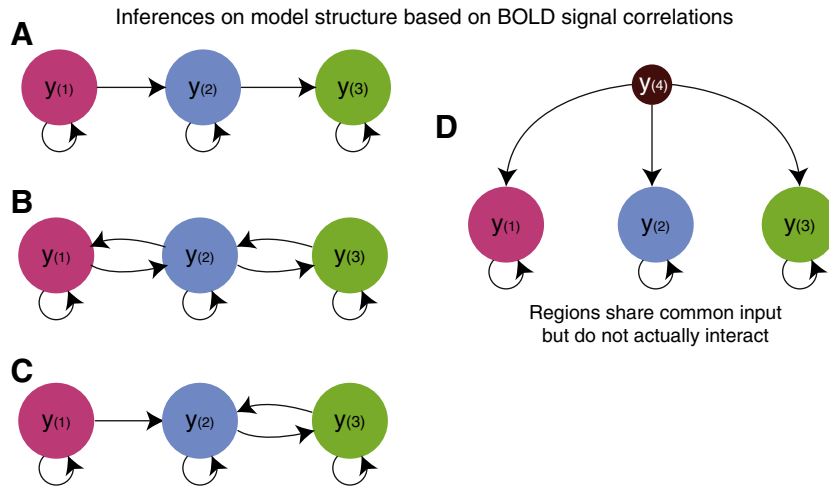
Inferences on model structure based on BOLD signal correlations

Fig. 2. Potential models of functional connectivity between regions: These three regions express highly correlated BOLD signal fluctuations, but little is known about the underlying functional architecture. Note that inferences are made at the level of the BOLD data, y, not the underlying neural activity, z.

impact of one region on another is delayed in time (utilising the principles of "dynamical systems theory") (Daunizeau et al., 2011a,b).

In order to quantify effective connectivity, one must create a plausible model of how the observed BOLD signal y may be generated by these influences (this is known as a "generative model"), just as was done in [Fig. 1B]. This departure from analysing y, to considering how y is generated is an important distinction to understand. When we model data it is important to consider that we are rarely able to fully explain how observed data are generated by complex systems — the aim of modelling is really to provide a parsimonious and plausible mechanism with which to better understand such systems.

DCM has become the method of choice for modelling causal interactions in neuroimaging data (note, while we have concentrated here on fMRI data, adapted DCM methods for local field potentials (LFP), electroencephalography (EEG) and magnetoencephalography (MEG) datasets are similarly increasingly used — see Box 1). The most common implementation of DCM (standard/deterministic DCM) assumes that the system modelled is fully deterministic, meaning that the evolution of hidden neural activity over time in a given region ($\dot{z}$) is purely due to its afferent (incoming) connections and experimental inputs (as described in Fig. 1C). Note the distinction between an afferent and an input. An afferent arrives from a node (distant or self), whereas an input is introduced by the experimenter (for example visual input from the retina, via the lateral geniculate nuclei – neither of these neural populations are themselves included in the model, rather it is assumed that they convey the convolved stimulus function coding visual input to a region – see footnote to Rule II). Theoretically therefore, if all the afferents and inputs to a region (including intrinsic connections) were removed, the rate of change of neural activity z ($\dot{z}$), would be zero, thus activity would remain constant. Such assumptions may not always be ideal as they forgo autonomous dynamics that may characterise certain brain regions (e.g. substantia nigra neurons), or periods where there are no experimental inputs (such as resting state fMRI — see Rule X), or regions not showing experimentally evoked effects.

The effect an afferent has on the dynamics of the node depends on the connection strength (or "coupling parameter" — e.g. the size of value $a_{2,1}$) and the activity at the source of the afferent. This is similarly true of inputs; the influence of an experimental input is determined by the value of $c_{1,u1}$.

VI. Experimental manipulations can also change the effective connectivity strength of a connection to produce *bilinear effects*

Now imagine the visual experiment is repeated in the same subject, this time after they have received a drug infusion (for pure theoretical perfection, imagine that this drug had no effect on neurovascular coupling, and had very slow pharmacokinetics) [Fig. 1D]. New fictitious findings now show the same regions to be active again, and the BOLD signal y_1 and y_3 remains unchanged. However, an interaction between visual stimulation and drug was noted in the observed responses of region 2; under drug, y_2 consistently increases.

Now, returning to our generative model, we can say that because the observed y_1 and y_3 are unaffected by the drug, it is likely that z_1 and z_3 have also remained largely unaffected. The changes noted in y_2 can only be explained by changes in z_2, and – under deterministic assumptions – this must be explained by a change in one (or more) of the coupling parameters of its afferents. In order to explain interactions, separate experimental inputs can enter the model as modulatory effects on a connection, as shown in [Fig. 1E]. In our example, the effect of drug (u_2) is modulating the connection $a_{2,3}$, according to the strength of the modulatory effect, $b_{2,u2}$, culminating in changes in $\dot{z}_2$. Notice that these are changes to a rate of change ($\dot{z}_2$), thus are called "second order" or "bilinear effects". This is because the modulatory input and the afferent input interact to affect the rate of change of neuronal activity. The astute reader will notice that changes in $\dot{z}_2$ could have been caused by modulatory effects on any of the other afferents into node 2 ($a_{2,1}$ or $a_{2,2}$ or both). This highlights the problem of competing hypotheses (explanations) in data modelling and will be discussed later.

VII. DCMs estimate the coupling parameters given the structure of the model, the experimental inputs, and the observed data

The purpose of a DCM analysis is to estimate the coupling parameters of a model (the a, b and c values), and evaluate how well a particular model explains the observed data. This allows the experimenter to make inferences about the structure of the network (e.g. whether the data is best explained by bottom-up or top-down transfer — see model comparison, Rule IX), as well as/or quantify the coupling strength and direction of coupling between regions. Importantly, neither of these questions can be answered by looking at correlation strengths as in functional connectivity.

Returning to the earlier example of PD, DCM studies have been able to account for differences between patients and controls, on vs. off medication states (Rowe et al., 2010), and on vs. off deep brain stimulation states (Kahan et al., 2012), specifically by considering the presence/absence of connections, and/or the strength of certain connections within a generative model. The same is true of electrophysiological datasets; Marreiros et al. (2012) have been able to demonstrate that the effect of levodopa on beta synchrony observed in motor regions in PD can be explained in terms of changes in causal influences between the

motor cortex, pallidum, and the subthalamic nucleus. Such studies have emphasised the role of neuronal synchrony over rate-based models of disease, and explain the observed changes in oscillatory dynamics in terms of changes in extrinsic effective connectivity (see also Moran et al., 2011b). The conclusions of these studies are arguably more informative about the disease/therapeutic process, highlighting key differences at specific causal influences, potentially providing targets for future therapeutic interventions.

So, how does DCM estimate the parameters of the generative model? Inevitably, this is where it gets a bit complicated. In our example, we have three equations that define $\dot{z}$ for each of our nodes (under deterministic assumptions); these are known as the "equations of motion" because they described the motion or rate of change of the hidden neuronal states (displayed in [Figs. 1C & F]).

DCM uses an expectation maximisation (EM) algorithm to produce probabilistic estimates of the expected value of each parameter, in addition to its variance (under the assumption that the parameter values conform to Gaussian assumptions, i.e. the parameter values are normally distributed). These values ("posterior estimates") are conditional on the model structure, thus models with a different architecture (different connections), will have different posterior estimates. In other words, posterior estimates are the most likely parameters, given a particular model and data. The intricacies of this estimation process are beyond the scope of this report; for those with a (strong) mathematical background, see (Supplementary Table 1 & Do and Batzoglou, 2008; Friston, 2002; Friston et al., 2003). The aim of the estimation process is to refine the model parameters, so that it produces a predicted signal that is as close as possible to the observed BOLD data.

In order to do this, one would need to understand the relationship between the z values for each node and the observed BOLD signal y. The relationship between neuronal state and observed signal can be modelled as a series of nonlinear biological processes depending on the neuroimaging modality utilised (see Box 1 for application to EEG/MEG/LFP data). For fMRI data, this originates with neuronal activity, causing an increase in the vasodilatory signal. This results in a proportional increase in flow into the region with concomitant changes in blood volume and deoxyhaemoglobin content, causing the observed change in BOLD response (Buxton et al., 1998; Friston et al., 2000). This is called the "haemodynamic forward model", mapping the hidden state to the observed data. This mapping is dependent on a number of parameters ("the haemodynamic parameters" in the case of fMRI) that, like the coupling parameters, need to be estimated. Once we have potential values for the coupling parameters, and the parameters that determine the forward model, a predicted y can be calculated.

In short, EM tunes the coupling and haemodynamic parameters so as to maximise the concordance between predicted and observed BOLD signal in a way that avoids using unlikely parameters (large parameters that render the model unstable or those that deviate substantially from prior assumptions on both coupling and haemodynamic parameters — the latter based on those obtained in (Friston, 2002)).

VIII. *Model inversion* allows one to compute the *evidence* for each model

The estimation procedure (sometimes called "model inversion") additionally scores the model in terms of how well it explains the data (i.e. how similar the predicted and observed responses are). It is important to note that this scoring is corrected for how many free parameters were estimated; mathematically, a model that has modulatory effects at two connections has the potential to explain the data better (i.e. it could be more accurate) than a model with modulatory effects on one connection. This would introduce a bias towards complex, "over-parameterised" models guilty of "over-fitting" (i.e. fitting too much of the observed data with the generative model — this is problematic because we know that whatever signal processing is implemented,

the BOLD signal is still noisy and in part due to extra-physiological variance) (Penny et al., 2004). Models that capture noise and over-fit data are less generalizable (i.e. they are unlikely to model a future dataset), thus have limited mechanistic value. This scoring is in terms of something called "model evidence" (or more accurately, the 'Free energy', which is an approximation of the model evidence), and is a compromise between model accuracy and model complexity, avoiding bias towards over-parameterised models.

IX. Equipped with the evidence, one can then compare models representing different *a priori* hypotheses of the functional architecture using Bayesian model selection

The model evidence can subsequently be used to compare a series of models to assess which of a number of plausible models is the most likely to have generated the observed data. This necessitates the investigator to have a series of equally likely competing hypotheses of the underlying functional architecture (or "model space") to test *a priori*. This model space may take the form of two models that do and do not possess an *a* connection between two regions, or models in which that connection is, or is not, modulated by an experimental manipulation.

Differences in relative "log-evidences" (the logarithm of model evidence) can then be summarised as a conditional probability (or "posterior probability") for each competing model, representing the probability of that model, given the observed data. This process of comparing the evidence for different models is known as Bayesian model selection (BMS) and has been extended to compare models in group studies, and compare different families of similar models (Penny et al., 2004, 2010; Stephan et al., 2009a,b). Model comparison is common practise in DCM studies and can shed light on pathological mechanisms. For example, BMS has been used to demonstrate that patients with primary progressive aphasia performing semantic and phonological processing tasks utilise the same functional architecture as control subjects, but with significantly reduced connectivity between Wernicke's and Broca's areas (Sonty et al., 2007).

Research into experimental design of DCM studies has revealed that subtle changes to an experiment, such as inter-stimulus interval and block length, can have significant effects on the statistical power of subsequent Bayesian model comparison. However, unlike GLM analyses, only a few studies have addressed issues of optimising DCM experimental design (Daunizeau et al., 2011a,b; Goulden et al., 2012).

X. 'Classical' DCM studies attempt to answer three types of question

When reading a study employing DCM, always try to understand the scientific question the author is trying to answer. Often this will fall into one (or more) of three categories: (1) What is the underlying functional architecture of a network of brain regions? (2) Which connections are modulated by experimental manipulation? (3) Are the coupling parameters of a network of brain regions different in two groups of people (e.g. patients vs. healthy controls)? Examples of questions that could be answered with an appropriately designed model space for our visual perception experiment are illustrated in [Fig. 3].

Question 1 can be answered by specifying a number of competing (but all equally plausible) models differing with respect to the presence/absence of a connection, and then comparing them using BMS to determine which is the most likely to have generated the data. For example, as alluded to in Rule VII, one may want to know whether the data is best explained purely by bottom-up transfer, or in combination with top–down mechanisms. A model space as shown in [Fig. 3A] (assuming a hierarchical organisation) could test this hypothesis; here, the models differ with regards the presence of feedback connections.

Similarly, question 2 would use BMS, however this time the competing models would differ with respect to the connections which an
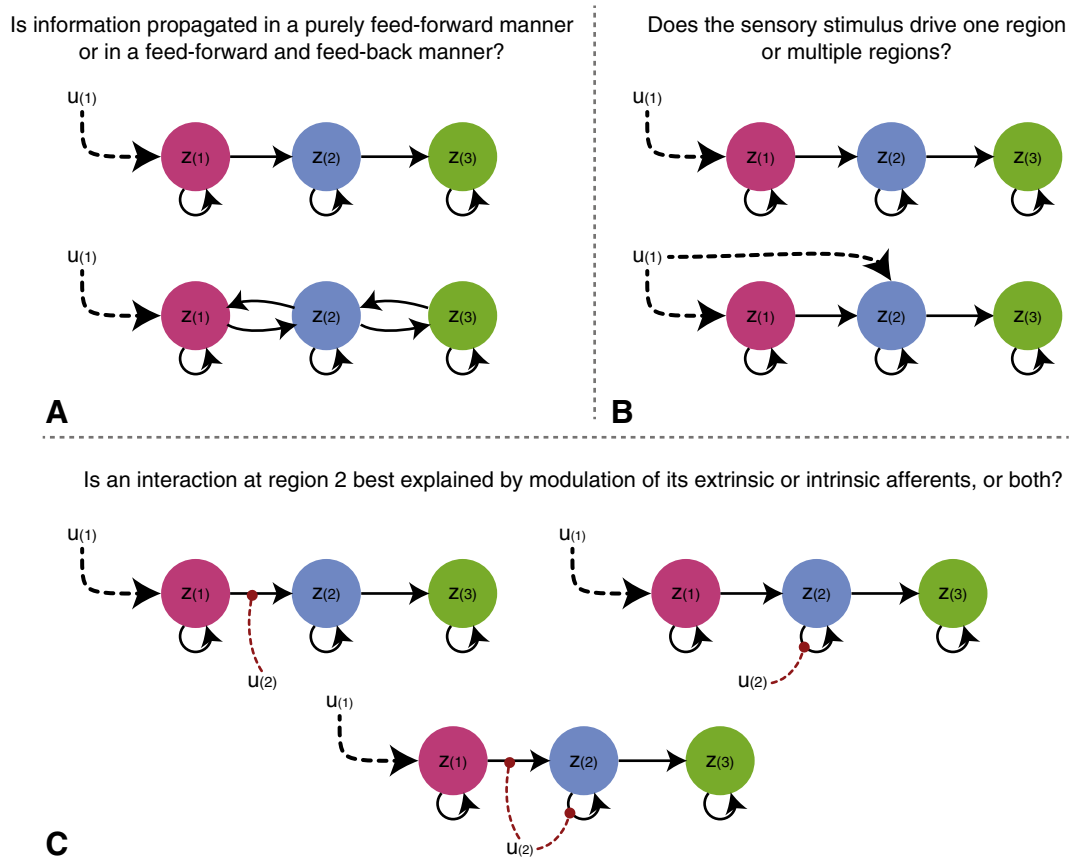
Fig. 3. Examples of possible model spaces testing different hypotheses: (A) Here, the two competing models differ by the presence of feedback connections, allowing the investigator to compare the evidence for purely feedforward transfer vs. both feedforward and feedback mechanisms. (B) These models differ by which nodes are directly driven by the experimental stimulus. (C) These models differ by which connections within the model are modulated by experimental factor u(2). Specifically, the investigator can ask whether the extrinsic connection from region 1, or the intrinsic connection in region 2, or both, are modulated.

experimental manipulation modulated. This design has been used to explore pathological processes (Daunizeau et al., 2012a,b), treatment effects (Kahan et al., 2012), as well as physiological cognitive processes (Urner et al., 2013; Vossel et al., 2012). Studies such as these often utilise a 2×2 factorial design, whereby one factor is used as a driving input, and another as a modulatory input (as in the visual perception example — the visual stimuli acted as a driving input to region 1, and the drug effect acted as a modulatory input on connection $a_{2,3}$).

Questions like (3) are typically answered by comparing (using conventional Student t tests) estimated (posterior) coupling parameters between two matched groups — where, crucially, the models per se were exactly the same for the two groups. DCMs attempt to model observed data at a number of regions, thus it is essential (especially with deterministic assumptions) that comparisons of coupling parameters (say, in a patient vs. control group) only occurs between models of the same architecture (models that possess the same regions, inputs and connections).

Since its conception, the DCM framework has been extended, allowing studies to address more specific questions. For example, "stochastic DCM" no longer relies on deterministic assumptions and accounts for random (stochastic) fluctuations in the evolution of regional activity (Li et al., 2011). This extension can thus be used to assess the relative contributions of experimental stimulation and endogenous fluctuations (Daunizeau et al., 2012a,b), or simply model periods of no task or experimental stimulation (for example, see Urner et al., 2013). More plausible models also exist for fMRI data, incorporating multiple states per region ("two-state DCM" — Marreiros et al., 2008), as well as non-linear gating effects ("nonlinear DCM" — Stephan et al., 2008). The added temporal resolution afforded by EEG/MEG data

permits DCM to employ a more rigorous biophysical model (see Box 1), allowing for inferences to be drawn on state-dependant changes in presynaptic current, postsynaptic conductance, or even properties of specific ion channels (Moran et al., 2011a,b). The application of DCM at these microscopic scales could be used to explain human physiological behavioural and cognitive processes, or even pathophysiological processes in terms of synaptic mechanisms within specific regions of the brain, *in vivo*.

You've made it to the end of a DCM paper

Congratulations. If you managed to get here without skipping the middle bit, you should now have the tools (or at least a crib-sheet — see Supplementary Table 1) with which to dissect most DCM papers. Should you wish to probe the statistical and biophysical underpinnings of DCM further, as well as explore the limitations and DCM-friendly experimental design, a recent review critically evaluates these matters (Daunizeau et al., 2011a,b), and a number of electrophysiological validation studies have been published in recent years (Daunizeau et al., 2012a,b; David et al., 2008). For a comparison with methods that rest on similar principles, see Friston et al. (2013). In addition, a glossary of commonly used terms are included (see Supplementary Table 1), and teaching material (videos and example datasets) is available online (http://www.fil.ion.ucl.ac.uk/spm/course/). Importantly, although the theory behind these analyses is complicated, implementation is not; DCM analyses can be executed through the standard SPM graphical user interface. More practical guides to employing these analyses in patient and control populations have been published
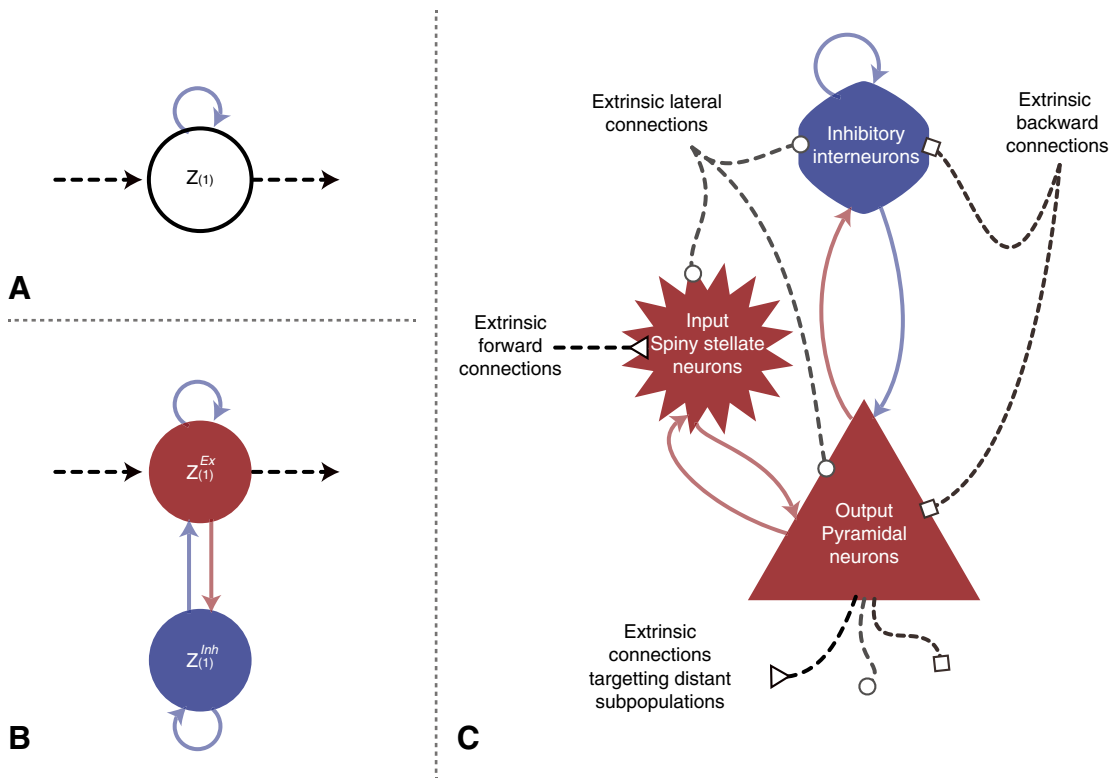
Fig. 4. The relative complexities of modelling fMRI and EEG/MEG/LFP data. (A) The basic fMRI model, with a single intrinsic connection, and non-specific extrinsic afferents and efferents. (B) The two-state fMRI model containing four intrinsic connections linking the excitatory and inhibitory subpopulation. (C) The convolution based ERP and LFP model — an example of a biophysical model employed in EEG/MEG/LFP data (adapted from Moran et al., 2009) — is relatively more complex. Firstly, three subpopulations are modelled with five intrinsic connections. Additionally, extrinsic afferents are categorised as either forward connections arriving at the input population, backward connections arriving at both the output and interneuron populations, or lateral connections arriving at all three populations. All extrinsic efferents project from the output pyramidal population. Red populations signify glutamatergic cell types, blue populations signify GABAergic cell types. Dashed lines represent extrinsic connections, solid lines represent intrinsic connections.

(Litvak et al., 2011; Rowe, 2010; Seghier et al., 2010; Stephan et al., 2010).

Ultimately, research is only interesting to those who understand it. For ideas and findings to permeate the neuroscience and clinical communities, it is essential that complicated methods do not become unfathomable. As time progresses, the method is slowly evolving to answer new questions. While advances herald new findings, they have provided additional complexity to DCM. However, these 'bells and whistles' aside, the fundamental aim remains true. DCMs seek to model how your observed data were generated from immeasurable neural activity.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2013.07.008.

Conflict of interest

The authors declare no conflicts of interests.

References

Abutalebi, J., et al., 2009. Bilingual aphasia and language control: a follow-up fMRI and intrinsic connectivity study. Brain Lang. 109 (2–3), 141–156 (Available at: http://www.ncbi.nlm.nih.gov/pubmed/19427522 [Accessed May 18, 2012]).

Anon., 2000. How experts communicate. Nat. Neurosci. 3 (2), 97 (Available at: http://www.ncbi.nlm.nih.gov/pubmed/10649559 [Accessed June 10, 2012]).

Bastos, A.M.M., et al., 2012. Canonical microcircuits for predictive coding. Neuron 76 (4), 695–711 (Available at: http://linkinghub.elsevier.com/retrieve/pii/S0896627312009592 [Accessed November 21, 2012]).

Biswal, B., et al., 1995. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. Magn. Reson. Med. 34 (4), 537–541 (Available at: http://www.ncbi.nlm.nih.gov/pubmed/8524021 [Accessed July 13, 2012]).

Buxton, R.B., Wong, E.C., Frank, L.R., 1998. Dynamics of blood flow and oxygenation changes during brain activation: the balloon model. Magn. Reson. Med. 39 (6), 855–864 (Available at: http://www.ncbi.nlm.nih.gov/pubmed/9621908 [Accessed September 1, 2011]).

Damoiseaux, J.S., et al., 2006. Consistent resting-state networks across healthy subjects. Proc. Natl. Acad. Sci. U. S. A. 103 (37), 13848–13853 (Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1564249&tool=pmcentrez&rendertype=abstract [Accessed March 12, 2013]).

Daunizeau, J., David, O., Stephan, K.E., 2011a. Dynamic causal modelling: a critical review of the biophysical and statistical foundations. NeuroImage 58 (2), 312–322 (Available at: http://www.ncbi.nlm.nih.gov/pubmed/19961941 [Accessed March 3, 2013]).

Daunizeau, Jean, et al., 2011b. Optimizing experimental design for comparing models of brain function. PLoS Comput. Biol. 7 (11), e1002280 (Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3219623&tool=pmcentrez&rendertype=abstract [Accessed May 31, 2013]).

Daunizeau, J., Lemieux, L., et al., 2012a. An electrophysiological validation of stochastic DCM for fMRI. Front. Comput. Neurosci. 6 (January), 103 (Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3548242&tool=pmcentrez&rendertype=abstract [Accessed March 12, 2013]).

Daunizeau, J., Stephan, K.E., Friston, K.J., 2012b. Stochastic dynamic causal modelling of fMRI data: should we care about neural noise? NeuroImage 62 (1), 464–481 (Available at: http://www.ncbi.nlm.nih.gov/pubmed/22579726 [Accessed October 11, 2012]).

David, Olivier, Harrison, Lee, Friston, Karl J., 2005. Modelling event-related responses in the brain. NeuroImage 25 (3), 756–770 (Available at: http://www.ncbi.nlm.nih.gov/pubmed/15808977 [Accessed March 5, 2013]).

David, Olivier, et al., 2006. Dynamic causal modeling of evoked responses in EEG and MEG. NeuroImage 30 (4), 1255–1272 (Available at: http://www.ncbi.nlm.nih.gov/pubmed/16473023 [Accessed March 12, 2013]).

David, Olivier, et al., 2008. Identifying neural drivers with functional MRI: an electrophysiological validation. PLoS Biol. 6 (12), 2683–2697 (Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2605917&tool=pmcentrez&rendertype=abstract [Accessed July 15, 2011]).

Do, C.B., Batzoglou, S., 2008. What is the expectation maximization algorithm? Nat. Biotechnol. 26 (8), 897–899 (Available at: http://www.ncbi.nlm.nih.gov/pubmed/18688245).

Douglas, R.J., Martin, K.a.C., Whitteridge, D., 1989. A canonical microcircuit for neocortex. Neural Comput. 1 (4), 480–488 (Available at: http://www.mitpressjournals.org/doi/abs/10.1162/neco.1989.1.4.480).

Felleman, D.J., Van Essen, D.C., 1991. Distributed hierarchical processing in the primate cerebral cortex. Cereb. Cortex 1 (1), 1–47 (Available at: http://www.ncbi.nlm.nih.gov/pubmed/1822724 [Accessed March 27, 2013]).

Friston, Karl J., 1994. Functional and effective connectivity in neuroimaging: a synthesis. Hum. Brain Mapp. 2 (1–2), 56–78 (Available at: http://doi.wiley.com/10.1002/hbm.460020107).

Friston, K.J., 2002. Bayesian estimation of dynamical systems: an application to fMRI. NeuroImage 16 (2), 513–530 (Available at: http://www.ncbi.nlm.nih.gov/pubmed/12030834 [Accessed November 16, 2012]).

Friston, K., 2009. Causal modelling and brain connectivity in functional magnetic resonance imaging. PLoS Biol. 7 (2), e33 (Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2642881&tool=pmcentrez&rendertype=abstract [Accessed March 21, 2012]).

Friston, Karl J., 2011. Functional and effective connectivity: a review. Brain Connectivity 1 (1), 13–36 (Available at: http://www.liebertonline.com/doi/abs/10.1089/brain.2011.0008 [Accessed March 13, 2012]).

Friston, K.J., et al., 2000. Nonlinear responses in fMRI: the Balloon model, Volterra kernels, and other hemodynamics. NeuroImage 12 (4), 466–477 (Available at: http://www.ncbi.nlm.nih.gov/pubmed/10988040 [Accessed November 10, 2012]).

Friston, K.J., Harrison, L., Penny, W., 2003. Dynamic causal modelling. NeuroImage 19 (4), 1273–1302 (Available at: http://linkinghub.elsevier.com/retrieve/pii/S1053811903002027 [Accessed March 1, 2012]).

Friston, K., Moran, R., Seth, A.K., 2013. Analysing connectivity with Granger causality and dynamic causal modelling. Curr. Opin. Neurobiol. 23 (2), 172–178 (Available at: http://www.ncbi.nlm.nih.gov/pubmed/23265964 [Accessed March 4, 2013]).

Goulden, N., et al., 2012. Sample size estimation for comparing parameters using dynamic causal modeling. Brain Connectivity 2 (2), 80–90 (Available at: http://www.ncbi.nlm.nih.gov/pubmed/22559836 [Accessed June 17, 2013]).

Greicius, M.D., et al., 2007. Resting-state functional connectivity in major depression: abnormally increased contributions from subgenual cingulate cortex and thalamus. Biol. Psychiatry 62 (5), 429–437 (Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2001244&tool=pmcentrez&rendertype=abstract [Accessed March 4, 2013]).

Hacker, C.D., et al., 2012. Resting state functional connectivity of the striatum in Parkinson's disease. Brain 135 (Pt 12), 3699–3711 (Available at: http://www.ncbi.nlm.nih.gov/pubmed/23195207 [Accessed March 3, 2013]).

Heeger, D.J., Ress, D., 2002. What does fMRI tell us about neuronal activity? Nat. Rev. Neurosci. 3 (2), 142–151 (Available at: http://www.ncbi.nlm.nih.gov/pubmed/11836522 [Accessed July 21, 2011]).

Jansen, B.H., Rit, V.G., 1995. Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. Biol. Cybern. 73 (4), 357–366 (Available at: http://www.ncbi.nlm.nih.gov/pubmed/7578475 [Accessed May 1, 2013]).

Jech, R., et al., 2013. Levodopa increases functional connectivity in the cerebellum and brainstem in Parkinson's disease. Brain 1–2 (Available at: http://www.ncbi.nlm.nih.gov/pubmed/23370091 [Accessed June 16, 2013]).

Kahan, J., et al., 2012. Therapeutic Subthalamic Nucleus Deep Brain Stimulation Reverses Cortico-Thalamic Coupling during Voluntary Movements in Parkinson's Disease. In: Baron, J.-C. (Ed.), PLoS ONE, 7. 12, p. e50270 (Available at: http://dx.plos.org/10.1371/journal.pone.0050270 [Accessed January 2, 2013]).

Li, B., Daunizeau, J., Stephan, K.E., Penny, W., Hu, D., Friston, K., 2011. Generalised filtering and stochastic DCM for fMRI. NeuroImage 58 (2), 442–457. http://dx.doi.org/10.1016/j.neuroimage.2011.01.085.

Litvak, V., et al., 2011. EEG and MEG data analysis in SPM8. Comput. Intell. Neurosci. 852961 (Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3061292&tool=pmcentrez&rendertype=abstract [Accessed March 2, 2013]).

Logothetis, N.K., 2008. What we can do and what we cannot do with fMRI. Nature 453 (7197), 869–878 (Available at: http://www.ncbi.nlm.nih.gov/pubmed/18548064 [Accessed February 29, 2012]).

Lowe, M.J., Mock, B.J., Sorenson, J.A., 1998. Functional connectivity in single and multislice echoplanar imaging using resting-state fluctuations. NeuroImage 7 (2), 119–132 (Available at: http://dx.doi.org/10.1006/nimg.1997.0315 [Accessed February 6, 2013]).

Marreiros, a C., Kiebel, S.J., Friston, K.J., 2008. Dynamic causal modelling for fMRI: a two-state model. NeuroImage 39 (1), 269–278 (Available at: http://www.ncbi.nlm.nih.gov/pubmed/17936017 [Accessed March 6, 2012]).

Marreiros, A.C., et al., 2012. Basal ganglia-cortical interactions in Parkinsonian patients. NeuroImage 301–310 (null(null), Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3573233&tool=pmcentrez&rendertype=abstract [Accessed November 21, 2012]).

Moran, R.J., et al., 2007. A neural mass model of spectral responses in electrophysiology. NeuroImage 37 (3), 706–720 (Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2644418&tool=pmcentrez&rendertype=abstract [Accessed March 12, 2013]).

Moran, R.J., et al., 2009. Dynamic causal models of steady-state responses. NeuroImage 44 (3), 796–811 (Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2644453&tool=pmcentrez&rendertype=abstract [Accessed March 12, 2013]).

Moran, Rosalyn J., et al., 2011a. An in vivo assay of synaptic function mediating human cognition. Curr. Biol. 21 (15), 1320–1325 (Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3153654&tool=pmcentrez&rendertype=abstract [Accessed March 9, 2013]).

Moran, Rosalyn J., et al., 2011b. Alterations in brain connectivity underlying beta oscillations in parkinsonism. PLoS Comput. Biol. 7 (8), e1002124 (Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3154892&tool=pmcentrez&rendertype=abstract [Accessed August 14, 2011]).

Moran, R., Pinotsis, D.A., Friston, K., 2013. Neural masses and fields in dynamic causal modeling. Frontiers in Computational Neuroscience 7. http://dx.doi.org/10.3389/fncom.2013.00057 (Available at: http://www.frontiersin.org/Computational_Neuroscience/10.3389/fncom.2013.00057/full [Accessed July 17, 2013]).

Mori, S., Zhang, J., 2006. Principles of diffusion tensor imaging and its applications to basic neuroscience research. Neuron 51 (5), 527–539 (Available at: http://www.ncbi.nlm.nih.gov/pubmed/16950152 [Accessed February 28, 2013]).

Penny, W.D., et al., 2004. Comparing dynamic causal models. NeuroImage 22 (3), 1157–1172 (Available at: http://www.ncbi.nlm.nih.gov/pubmed/15219588 [Accessed July 18, 2012]).

Penny, Will D., et al., 2010. Comparing families of dynamic causal models. PLoS Comput. Biol. 6 (3), e1000709 (Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2837394&tool=pmcentrez&rendertype=abstract [Accessed March 6, 2013]).

Rehme, A.K., et al., 2011. Dynamic causal modeling of cortical activity from the acute to the chronic stage after stroke. NeuroImage 55 (3), 1147–1158 (Available at: http://www.ncbi.nlm.nih.gov/pubmed/21238594 [Accessed July 24, 2012]).

Rowe, James B., 2010. Connectivity analysis is essential to understand neurological disorders. Front. Syst. Neurosci. 4 (September), 1–13 (Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2953412&tool=pmcentrez&rendertype=abstract [Accessed October 9, 2012]).

Rowe, J.B., et al., 2010. Dynamic causal modelling of effective connectivity from fMRI: are results reproducible and sensitive to Parkinson's disease and its treatment? NeuroImage 52 (3), 1015–1026 (Available at: http://www.sciencedirect.com/science/article/pii/S105381190901369X [Accessed August 25, 2011]).

Seghier, M.L., et al., 2010. Identifying abnormal connectivity in patients using dynamic causal modeling of FMRI responses. Front. Syst. Neurosci. 4 (Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2936900&tool=pmcentrez&rendertype=abstract [Accessed June 23, 2011]).

Smith, Stephen M., 2012. The future of FMRI connectivity. NeuroImage 62 (2), 1257–1266 (Available at: http://www.ncbi.nlm.nih.gov/pubmed/22248579 [Accessed October 26, 2012]).

Sonty, S.P., et al., 2007. Altered effective connectivity within the language network in primary progressive aphasia. Neuroscience J. Neurosci. Off. J. Soc. Neurosci. 27 (6), 1334–1345 (Available at: http://www.ncbi.nlm.nih.gov/pubmed/17287508 [Accessed June 10, 2013]).

Stephan, Klaas Enno, et al., 2008. Nonlinear dynamic causal models for fMRI. NeuroImage 42 (2), 649–662 (Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2636907&tool=pmcentrez&rendertype=abstract [Accessed November 14, 2012]).

Stephan, Klaas Enno, Penny, Will D., et al., 2009a. Bayesian model selection for group studies. NeuroImage 46 (4), 1004–1017 (Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2703732&tool=pmcentrez&rendertype=abstract [Accessed July 22, 2012]).

Stephan, Klaas Enno, Tittgemeyer, M., et al., 2009b. Tractography-based priors for dynamic causal models. NeuroImage 47 (4), 1628–1638 (Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2728433&tool=pmcentrez&rendertype=abstract [Accessed March 8, 2013]).

Stephan, K.E., et al., 2010. Ten simple rules for dynamic causal modeling. NeuroImage 49 (4), 3099–3109 (Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2825373&tool=pmcentrez&rendertype=abstract [Accessed June 15, 2011]).

Urner, M., et al., 2013. Early visual learning induces long-lasting connectivity changes during rest in the human brain. NeuroImage 77, 148–156 (Available at: http://www.ncbi.nlm.nih.gov/pubmed/23558105 [Accessed April 6, 2013]).

Vossel, S., et al., 2012. Deconstructing the architecture of dorsal and ventral attention systems with dynamic causal modeling. J. Neurosci. 32 (31), 10637–10648. http://dx.doi.org/10.1523/JNEUROSCI.0414-12.2012 (Available at: [Accessed August 2, 2012]).

Wu, T., et al., 2009. Changes of functional connectivity of the motor network in the resting state in Parkinson's disease. Neurosci. Lett. 460 (1), 6–10 (Available at: http://www.ncbi.nlm.nih.gov/pubmed/19463891 [Accessed July 26, 2011]).

Zhang, D., Raichle, M.E., 2010. Disease and the brain's dark energy. Nat. Rev. Neurol. 6 (1), 15–28 Available at: http://www.ncbi.nlm.nih.gov/pubmed/20057496 [Accessed February 28, 2013].