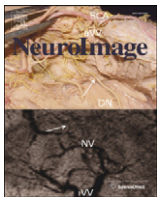
NeuroImage 54 (2011) 1662–1668

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg



Depression alters "top-down" visual attention: A dynamic causal modeling comparison between depressed and healthy subjects

Martin Desseilles [a,b,c,*], Sophie Schwartz [c], Thien Thanh Dang-Vu [a,d], Virginie Sterpenich [a,c], Marc Ansseau [e], Pierre Maquet [a,d], Christophe Phillips [a,f,*]

[a] Cyclotron Research Centre, University of Liège, Belgium
[b] Department of Psychiatry, University of Geneva, Switzerland
[c] Department of Neuroscience, University of Geneva, Switzerland
[d] Department of Neurology, University of Liège, Belgium
[e] Department of Psychiatry, University of Liège, Belgium
[f] Department of Electrical Engineering and Computer Science, University of Liège, Belgium

ARTICLE INFO

Article history:
Received 7 May 2010
Revised 27 July 2010
Accepted 23 August 2010
Available online 31 August 2010

ABSTRACT

Using functional magnetic resonance imaging (fMRI), we recently demonstrated that nonmedicated patients with a first episode of unipolar major depression (MDD) compared to matched controls exhibited an abnormal neural filtering of irrelevant visual information (Desseilles et al., 2009). During scanning, subjects performed a visual attention task imposing two different levels of attentional load at fixation (low or high), while task-irrelevant colored stimuli were presented in the periphery.

In the present study, we focused on the visuo-attentional system and used "Dynamic Causal Modeling" (DCM) on the same dataset to assess how attention influences a network of three dynamically-interconnected brain regions (visual areas V1 and V4, and intraparietal sulcus (P), differentially in MDD patients and healthy controls. Bayesian model selection (BMS) and model space partitioning (MSP) were used to determine the best model in each population. The best model for the controls revealed that the increase of parietal activity by high attention load was selectively associated with a negative modulation of P on V4, consistent with high attention reducing the processing of irrelevant colored peripheral stimuli. The best model accounting for the data from the MDD patients showed that both low and high attention levels exerted modulatory effects on P. The present results document abnormal effective connectivity across visuo-attentional networks in MDD, which likely contributes to deficient attentional filtering of information.

Introduction

Major depressive disorder (MDD) is characterized by cognitive and affective abnormalities (Chamberlain and Sahakian, 2006). Neuroimaging studies showed that these symptoms are mainly associated with impairments within fronto-limbic network involved in affective regulation and fronto-parietal networks involved in attention and executive functions (Mayberg et al., 1999). Lavie et al. (2004, 2005) proposed the 'Load Theory' of attention in which distractor rejection depends on the level and type of attentional load involved by the current task. This theory of attention and cognitive control implies reciprocal interactions for the processing of task-relevant and task-irrelevant information. On the one hand, increased perceptual demands (or load) related to the processing of task-relevant stimuli reduce the processing of task-irrelevant stimuli. On the other hand, irrelevant stimuli may also interfere with the processing task-relevant stimuli.

In a recent fMRI study, we scanned 14 drug-free MDD patients and 14 matched controls when they performed a visual attention task with two different levels of attentional load at fixation (low or high), while task-irrelevant colored stimuli were presented in the periphery (Desseilles et al., 2009). Healthy subjects showed increased responses in the color responsive area V4 during the easy, low-load condition compared to the difficult, high-load condition, consistent with the Load Theory (Lavie, 2005). By contrast, MDD patients showed decreased V4 activity during both the low and high load task (vs. a baseline, passive viewing condition). Using a psychophysiological interaction approach (PPI, (Friston et al., 1997), we also found enhanced functional connectivity in the context of low vs. high attentional load between intraparietal sulcus (P) and V4 selectively in controls but not in MDD patients (Desseilles et al., 2009).

While functional connectivity as assessed by PPI highlights correlations between activities in different regions ("physiological") as a

* Corresponding authors. Cyclotron Research Centre B30, University of Liège, Sart Tilman, 4000 Liège, Belgium. Fax: +32 4 366 29 46.
E-mail addresses: m.desseilles@ulg.ac.be (M. Desseilles), c.phillips@ulg.ac.be (C. Phillips).

1053-8119/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2010.08.061

function of experimental ("psychological") contexts, it does not provide any measurement of the directionality or causality for the observed correlations between brain regions. Here we sought to overcome these limitations by assessing effective connectivity with Dynamic Causal Modeling (DCM), which determines causal relationships across potentially distributed neural networks (Friston et al., 2003; Stephan et al., 2010).

In a DCM approach, a model is defined by a set of connected brain areas influenced by specific inputs. To date, few fMRI studies have used DCM to assess effective connectivity during cognitive or emotional task in MDD patients (Schlosser et al., 2008; Almeida et al., 2009). Overall studies assessing connectivity (either effective or not) mainly focused on fronto-cingulate (Anand et al., 2005; Schlosser et al., 2008; Vasic et al., 2009) or fronto-amygdala connectivity (Irwin et al., 2004; Johnstone et al., 2007; Almeida et al., 2009), and pointed to a failure to down-regulate rostral anterior cingulate cortex (ACC) and amygdala responses. Several studies showed deficient down-regulation within the default-mode network, including the ventro-medial prefrontal and the (subgenual) anterior cingulated cortex (Greicius et al., 2007; Sheline et al., 2009). How MDD affects effective connectivity across a distributed visuo-attentional network remains unclear.

Our goal was to better characterize the causal influence of top-down modulation of visual areas by parietal cortex, using a combination of DCM fitting, model space partitioning (MSP), Bayesian model selection and averaging (BMS and BMA), applied to each population (Penny et al., 2004; Stephan et al., 2009, 2010). We could thus, within a Bayesian framework, characterize the brain connectivity of the two groups of subjects, as well as their differences. We hypothesized that controls would display a connectivity network compatible with the Load Theory (Lavie et al., 2004) with a reduced drive of the activity in V4 in high attentional load mediated either by a direct action of P on V4 or by an indirect decrease of the gating from V1 to V4 (Schwartz et al., 2005; Stephan et al., 2008). We hypothesized that the connectivity of this visuo-attentional system would differ between MDD and controls in two distinct ways. First, we postulate that the low attention condition might modulate the observed network dynamics in MDD, but not in controls. Second, at the network level, the posterior estimates of the attention modulation should differ between MDD and controls (e.g. differential modulation of V4 activity by P). Importantly, we expect that MDD should not affect the global architecture of this visuo-attentional network.

Material and methods

Patients and controls

Details for the patients and controls recruitment can be found in (Desseilles et al., 2009). In summary, 14 subjects with a first lifetime episode of unipolar major depressive disorder, with no prior antidepressant or antipsychotic treatment, were included in the experimental protocol. Fourteen healthy controls were selected to match the MDD patients for gender, age, and socio-cultural background. None of them reported any severe medical problem, or any neurological or psychiatric history. After providing informed consent, they were also screened using the Structured Clinical Interview for Axis I DSM-IV Disorders (SCID-IV; (Spitzer et al., 1994) and the Hamilton Depression Rating Scale (HDRS, (Hamilton, 1960).

Experimental paradigm

Details for the experimental paradigm can be found in (Desseilles et al., 2009). During the main fMRI experiment, the participants performed a detection task on a rapid successive visual presentation (RSVP) of colored letters (one letter every 750 ms; 500 ms duration each, plus 250 ms blank) that was shown on a fixed central location at fixation. This RSVP stream consisted of T-shaped stimuli displayed with two possible orientations (upright or upside-down) and 8 different colors in a pseudorandom order. Blocks of 20-sec with Mondrian-like stimuli, referred later as "colored block" stimuli, formed by a $20°\times20°$ grid of colored rectangles, presented bilaterally at 6 degrees of visual angle from fixation (subtending $6°\times10°$), alternated with blocks without peripheral stimuli. All visual stimuli were generated using a MATLAB Toolbox, allowing visual presentation and response-recording with precise timing (Cogent, http://www.vislab.ucl.ac.uk/Cogent/).

Participants performed either a low load or a high load task, or were required to only fixate the central RSVP (40-sec period each) during scanning. The low load task required a key-press for any red T irrespective of its orientation. The high load task required a key-press for any upright yellow T or upside-down blue T. Baseline control condition required passive fixation of the letter stream but no key press. Low load, high load, and baseline passive viewing periods alternated in a pseudo-random order (randomized across participants) during one single continuous scanning session. Items that required a button-press response during both low and high attentional load appeared on average every 15 stimuli. In other words, only the task instructions distinguished the high load and low load conditions. The peripheral Mondrian-like stimuli were always irrelevant to the central task, and participants were instructed to ignore them.

fMRI data acquisition procedure

Data were acquired with a 3T head-only MR scanner (Allegra, Siemens, Erlangen) using a gradient echo EPI sequence (32 transverse slices with 30% gap, TR: 2130 ms, TE: 40 ms, FA: 90°, FOV: 220×220 mm, matrix size: $64\times64\times32$, voxel size: $3.4\times3.4\times3$ mm). Functional volumes ($n=255$) were acquired during one single continuous scanning run (sequential scheme of acquisition). The first three volumes were discarded to account for T1 saturation effects. A structural MR scan was acquired at the end of the experimental session (T1-weighted 3D MP-RAGE sequence, TR: 1960 ms, TE: 4.43 ms, TI: 1100 ms, FOV: 230×173 mm, matrix size: $256\times192\times176$, voxel size: $0.9\times0.9\times0.9$ mm). During scanning, eye movements were continuously monitored using an infrared eye tracking system (LRO5000, Applied Science Laboratories, Bedford, MA, sampling rate = 60 Hertz). Eye tracking data were used to ensure that all subjects included in the analyses maintained good central fixation during the whole scanning session.

Data used for DCM

The fMRI data of each subject were spatially preprocessed using the SPM2 toolbox (http://www.fil.ion.ucl.ac.uk/spm). The functional volumes were first realigned, then normalized to the MNI template space, using the EPI template provided with SPM toolbox, and finally smoothed with an $8\times8\times8$ mm Gaussian kernel. The structural MRI was registered to the fMR images and warped along into the MNI template space.

Based on the results from our previous study using the same paradigm and from research showing attentional modulation of networks dynamics (Friston and Buchel, 2000; Friston et al., 2003), we considered 3 main "regions of interest" (ROI) for our DCM analysis: the right early visual region (V1) and the right color responsive area (V4), both likely to process the visual aspects of the colored block stimuli, and the right intraparietal sulcus (P) linked to top-down attentional modulation of visual areas (Corbetta and Shulman, 2002). We chose regions in the right hemisphere because they showed robust activation in all our subjects (at the fixed effect level) and because attentional effects in V4 were also stronger on the right for both populations. For each subject, the time series of activity for each area was extracted as the first eigen-component of the time series of

the voxels within a ROI. These time series were adjusted for subject's movements using the movement parameters as regressors, and the button-pressing responses (corresponding to the detection of targets) convolved with the standard hemodynamic response. Using the individual statistical map displayed over the subject's structural MRI, each ROI was defined as a sphere of 10 mm radius, centered individually on the local activation maximum closest to each peak of interest from the group analysis (coordinates in MNI space: V1 [12, −86, −8] mm ; V4 [22, −74, −8] mm; P [30, −50, 50] mm).

Dynamic causal modeling

Dynamic causal modeling (DCM) is a mathematical framework used to characterize the effective connectivity at the neuronal level, i.e. the causal influence that one brain region exerts over another brain region (Friston et al., 2003; Stephan et al., 2010). DCM for fMRI is an input-state-output deterministic model of the neuronal activity across a network of brain regions (Friston et al., 2003). Inputs consist in direct stimuli, such as sensory stimulations, or contextual conditions provided by behavioral task instructions, history of the stimuli (e.g. in memory studies), pharmacological manipulations, etc. These inputs are exactly the same as those used to build design matrices in conventional fMRI analyses and are the "causes" used in the dynamical model. The brain areas considered are directionally connected such that they form a dynamic causal network. The direct inputs excite some brain areas and the induced activity propagates through the network, while the contextual inputs modulate the directional connections of the network. Finally, the neuronal activity from each brain area is translated into BOLD signal, i.e. an observable output, via a hemodynamic model (Buxton et al., 1998; Friston et al., 2000; Stephan et al., 2007a,b). Once the architecture of the model is specified, the coupling and hemodynamic parameters of the model can be estimated from the known inputs and recorded BOLD signal. This estimation proceeds with a Bayesian scheme (Friston, 2002; Friston et al., 2003, 2007), with shrinkage priors on the coupling parameters and empirical priors for the hemodynamic parameters, and provides a posteriori estimates of the model parameters.

Within a DCM framework, multiple hypotheses on brain functioning are expressed by different connectivity models. Several competing models are therefore fitted to the same data then compared across subjects (Penny et al., 2004; Stephan et al., 2009). Families of models, i.e. separate subsets of models sharing a specific feature, can compared and models within a chosen family can be averaged (Penny et al., 2010). This family comparison can answer a question such as: "Is this model feature necessary to explain the observed data, irrespective of the other characteristics of the model?" Finally an optimal model for a population can be identified within a family of models and the model parameters can be averaged (via Bayesian model averaging, BMA) across subjects. This BMA provides the full posterior distribution of the parameters of the model averaged across the subset of models and group of subjects.

Equivalently to standard multi-subject fMRI analyses, family and model selection (within a group) can proceed in a "fixed effect" (FFX) or "random effect" (RFX) way (Stephan et al., 2010). Given the nature of the brain mechanism modelled here (basic level visuo-attentional process) and the homogeneity of the subjects in each group, a FFX approach is preferred here both at the family and model inference level. Indeed with a FFX scheme, one assumes that every subject (in a group) uses the same model architecture, an assumption commonly accepted (Kumar et al., 2007; Stephan et al., 2007a,b; Acs and Greenlee, 2008; Rowe et al., 2010) and warranted when studying a basic physiological mechanism that is unlikely to vary across subjects (Stephan et al., 2010). The results of the RFX analysis, both at family and model level, are available in the supplementary material section for comparison.

The two groups of subjects, controls and MDD, can be compared at three different levels (family, model and parameter level). Using the same set of models and families of models for both groups, the posterior probability of each family and model will therefore be estimated for each group and the output directly compared. For the parameter level inferences, we chose to stay within the same Bayesian framework and to build on the two previous inference levels: "family," then "model," then "parameter." Group-BMA provides the full posterior probability of each parameter and one can check the posterior probability of any parameter to be significantly bigger (or smaller) than zero (Stephan et al., 2010). The parameters common to both group models (thus having the same interpretation) can also be directly compared: the posterior distribution of the difference between the group parameters was estimated from the posterior distribution of the parameters from each group.

We used the latest version of the SPM8 toolbox (http://www.fil.ion.ucl.ac.uk/spm) for all the DCM estimations and comparisons.

Attentional models

Among the almost infinite number of theoretically possible bilinear and non-linear models which can be built with 3 brain areas and 2 types of inputs, we consider here only a sub-set of 10 models, with plausible physiological connections, which allow us to test our hypothesis (Stephan et al., 2010). The selection of these models is based on the paradigm used, neuroanatomy and the literature (Friston and Buchel, 2000; Friston et al., 2003). The endogenous network connecting the 3 areas (V1, V4 and P) is the same for all the models and consists in a cascade of forward and backward connections between V1 and V4, and between V4 and P (Fig. 1). The central relevant stimuli formed a continuous stream of stimulation, generating steady-state activity, and thus are not explicitly modeled. Our goal is rather to model the variation of neuronal activity, induced by the peripheral stimuli and the top-down attentional demand of the discrimination task, around this steady-state.

In all 10 models, blocks of irrelevant color stimuli presented bilaterally enter the network in V1 as a direct visual input while the models differ between each other only with respect to the level of attentional load (high and low) and/or the modulation of the connections by the attentional load. These models naturally split into 2 subsets or "Families" of 5 models each: including either both low and high attention contextual inputs (Family A comprising models A.1 to A.5), or only the high attention contextual input (Family B, comprising models B.1 to B.5, Fig. 1). By modeling low and high attention level separately, with models A.1-5, the relative effect of attention level is therefore not fixed, and the models B.1-5 simply assume that the low attention context does not affect the network dynamics.

In models A.1 and B.1 (respectively A.2 and B.2), the attention input simply modulates the strength of the V1-to-V4 forward (respectively, P-to-V4 backward) connection. With these models, the attention only modulates how the activity propagates through the network. Models A.3 and B.3 are similar to models A.1 and B.1 respectively, except for the additional direct influence of attention on activity in P. In models A.4 and B.4, the attentional effect directly drives the activity in P, which in turn modulates the V1-to-V4 forward connection. The modulating effect of attention is thus carried out indirectly via area P. Finally model A.5 (respectively, B.5) is a combination of models A.1 and A.4 (respectively, B.1 and B.4). The effect of attention is thus twofold: direct on the V1-to-V4 forward connection and indirect via area P.

The 10 DCM models were fitted with the data from each of the 14 controls and 14 MDD subjects. This provided the model log-evidence and posterior parameter estimates for each of the 280 ($2\times14\times10$) model fits.

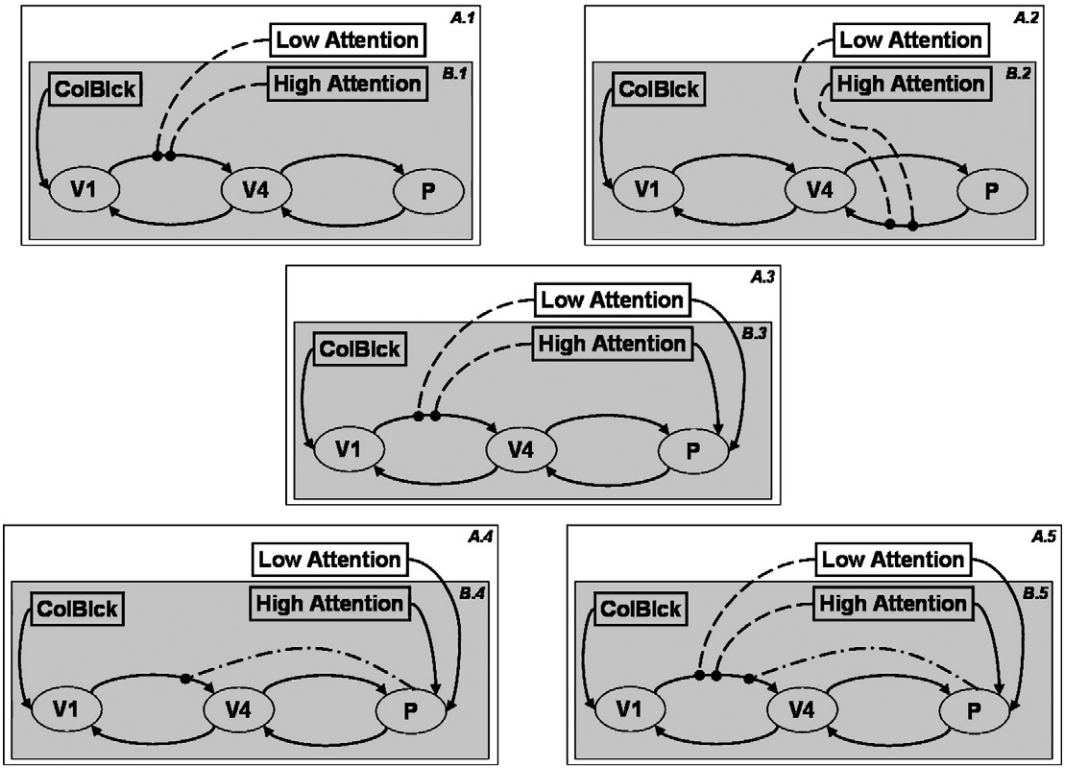


Fig. 1. Attentional models considered: Family A, consisting of models A.1 to A.5, are a subset of models which include both attention loads, high and low, contextual inputs. Another subset of models, Family B, models B.1 to B.5 (shaded box), is similar to Family A but only includes the high attentional contextual input, i.e. low attention is not taken into account. The network of interconnected regions includes visual area V1, colour responsive area V4, and superior parietal cortex (P). Plain arrows indicate the direct driving of an area by an external input or another area. Dashed lines represent the gating influence of one context, here low or high attention, over the connection between two different areas, here V1-to-V4 or P-to-V4. Similarly dashed-dotted lines represent the gating influence of one area, here P, over the connection between two different areas. For dashed and dashed-dotted lines, the end dot indicates the location of the modulatory effect by the input or area.

Results

Details for the clinical characteristics and for the standard fMRI results can be found in Desseilles et al. (2009).

Behavioral results

We found slower reaction-times, reduced hit rates and more false alarms during the high compared to the low load condition, indicating that the instructions successfully modulated task difficulty in both groups. Eye-tracking data demonstrated that both populations fixated the central stream of letters equally well.

Family-level inference

The numerical results for both control and MDD groups family partitioning are summarized in Table 1. For the control group, Family B models in which only high attention affects the network dynamics is more likely than Family A, models including both low and high attention, regardless of any other aspect of the models (posterior probability, $p = 1.0$ versus $p = 3.964 \times 10^{-17}$). By contrast, for the MDD group, Family A is significantly more likely ($p = 1.0$) than Family B ($p = 4.322 \times 10^{-42}$). This constitutes a significant proof, at the group level, that the low attention context is equivalent to the baseline passive viewing context for the controls, and that both low and high attention context should be included in the optimal model for the MDD patients. For the sake of completion, the results for an RFX approach, which are congruent with the FFX approach, are available in Supplementary materials.

Model-level inference and model averaging

The numerical results for both control and MDD groups model selection are summarized in Table 2. For the control group, we have significant evidence in favor of model B.4 (posterior probability $p = 1.0$), versus any other model ($p = 3.159 \times 10^{-11}$). Similarly, the optimal model for the MDD patients is model A.4 ($p = 1.0$, alternative hypothesis $p = 0.716 \times 10^{-11}$). The only difference between models A.4 and B.4 is the inclusion (or not) of the 'Low Attention' contextual influence on the parietal area. Interestingly, the architecture of these models (see Fig. 2) is similar to that found previously in a healthy subject but with a different paradigm (Stephan et al., 2008). For the sake of completion, the results for an RFX approach, which are congruent with the FFX approach, are available in Supplementary materials.

Using BMA, the full posterior distribution of the optimal model parameters within each group was calculated. The maximum *a posteriori* (MAP) values are shown in Fig. 2a and b for the control and MDD group

Table 1
Estimated family posterior probabilities for each group.

GROUP	Family A	Family B
Controls	3.9640×10^{-17}	1.0
MDD	1.0	4.3219×10^{-42}

Table 2
Model selection in "winning family" for each group: model index, posterior probability of selected model and probability of alternative hypothesis.

GROUP	Model index	Posterior probability	Alternative probability
Controls	B.4	1.0	0.3159×10^{-12}
MDD	A.4	1.0	0.0716×10^{-12}

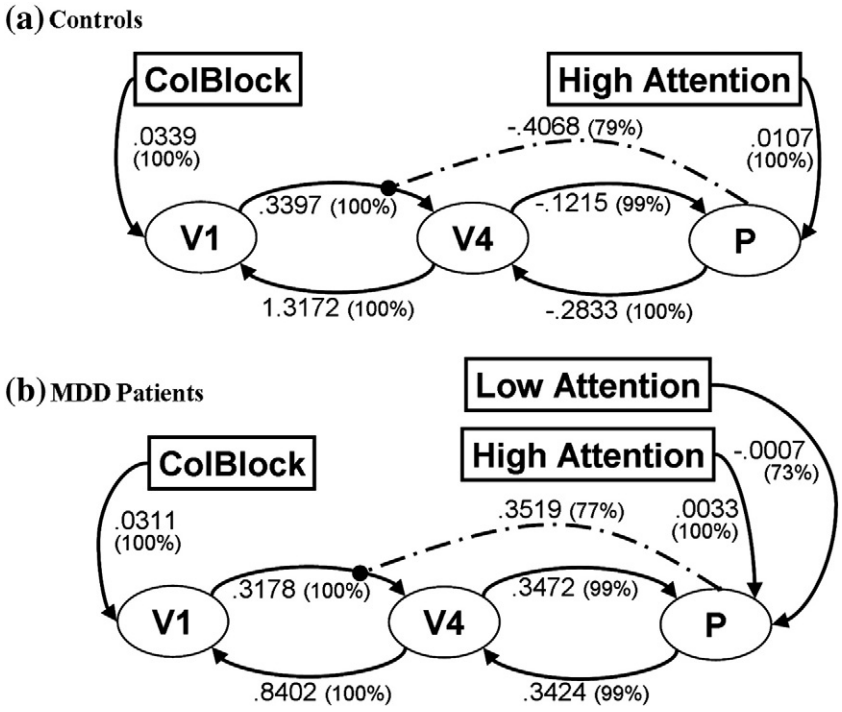


Fig. 2. Models selected: (a) B.4 for the healthy controls group, and (b) A.4 for the MDD patients group. The values indicated correspond to the maximum a posteriori estimates of the parameters after the group-BMA. The values in brackets are the confidence that these MAP estimates are different from zero.

respectively. The posterior confidence that these values were different from zero was calculated and are indicated in parenthesis.

Parameter-level inference

Table 3 summarizes the connection parameter estimates from each group obtained after BMA, as well as the difference (MDD minus control) between parameters common in both group optimal model. The value in parenthesis is the posterior probability of the parameter to be bigger (respectively smaller) than zero for a positive (respectively negative) MAP value. All the tests were performed in a Bayesian framework in line with the previous steps: parameter estimation, family and model selection. With the control group, the high attentional load effect positively modulates parietal cortex which in turn weakens V4 activity either directly, P-to-V4 connection, or indirectly, P-to-(V1-to-V4 gating). For the MDD group, most parameters are significantly different from the values observed in the control group. Nevertheless at the lowest level of the networks, V1 response to the colored stimuli and its projection onto V4 are almost identical in both groups.

Another approach would have been to extract the MAP parameter values from each subject individually and proceed with a classic frequentist test: *t*-test or ANOVA, within or between groups (Schlosser et al., 2008). The results of such an approach are actually very similar to those obtained with a Bayesian approach (see Supplementary Material).

Table 3
Maximum *a posteriori* connectivity values from the Bayesian model averaging (BMA) for the optimal model in each group, and *maximum a posteriori* difference of connectivity parameters between groups. The posterior probability for each parameter (or difference) to be bigger (or smaller for negative values) than zero is indicated in parentheses.

Connection names	Control group, BMA MAP estimate of parameter, and posterior *p*-value	MDD group, BMA MAP estimate of parameter, and posterior *p*-value	MAP of parameter difference (MDD-CO), and posterior *p*-value
V1-to-V4	0.3397 (1.0)*	0.3178 (1.0)*	−0.0215 (0.79)
V4-to-V1	1.3172 (1.0)*	0.8402 (1.0)*	−0.4757 (0.98)*
V4-to-P	−0.1215 (0.99)*	0.3472 (0.99)*	0.4685 (1.0)*
P-to-V4	−0.2833 (1.0)*	0.3424 (0.99)*	0.6274 (1.0)*
ColBlock-to-V1	0.0339 (1.0)*	0.0311 (1.0)*	−0.0029 (0.86)
LowAtt-to-P	–	−0.0007 (0.73)	–
HighAtt-to-P	0.0107 (1.0)*	0.0033 (1.0)*	−0.0074 (1.0)*
P-to-(V1-to-V4)	−0.4068 (0.79)	0.3519 (0.77)	0.7594 (0.87)

* value (parameter estimate or between group parameter difference) significantly different from zero, based on the posterior probability ($p>0.95$).

Discussion

Consistent with the Load Theory of attention and cognitive control (Lavie et al., 2004; Lavie, 2005), high perceptual demands of a task at fixation reduce visual responses to peripheral task-irrelevant stimuli (Schwartz et al., 2005). In the present study, we used a DCM approach to show that abnormal effects of attention on the processing of irrelevant stimuli in MDD is mediated by modifications in effective connectivity within a distributed visuo-attentional network (Desseilles et al., 2009). Specifically, DCM allowed us to explore how inputs and task contexts may interact to shape the dynamics of effective connectivity across brain regions, differentially in healthy and MDD populations. Our results suggest that the description of the propagation of complex influences such as attention can be useful to better understand functional changes underlying the pathophysiology of depression.

In the present DCM study we tested in 14 MDD patients (compared to 14 healthy controls) whether changing attentional demands of a central task would interfere with parietal top-down control on visual cortex activity elicited by task-irrelevant colored visual stimuli presented in the periphery. As in a previous DCM study of attention to motion (Friston et al., 2003; Stephan et al., 2008), we tested several DCM models encompassing visual and attentional regions. Here we selected the following regions in the right hemisphere: primary visual area V1, color responsive area V4, and intraparietal sulcus (P).

In healthy controls, Family level inference showed that the low attentional load condition did not influence task-irrelevant colored

stimuli processing independently of the actual underlying model architecture. In a next step, model selection showed that model B.4 (see Fig. 1) was optimal for the healthy controls. In this model, increase of attention demands in high attentional load condition activated parietal cortex which in turn diminished V4 activity both directly (parietal-to-V4 connection) and indirectly (parietal to V1-to-V4 gating).

These results are congruent with an increased filtering of task-irrelevant colored stimuli when attentional demand for the central task (high attentional load condition) increases (Lavie, 1995, 2005). Similarly, in our recent study (Desseilles et al., 2009), the fMRI signal in V4 from the controls was maintained during low attentional load condition but decreased during the high load condition, while functional connectivity measured by PPI increased in the context of low (vs. high) attentional load between right intraparietal sulcus and V4. Unfortunately, given the correlational nature of PPI, it was not possible to interpret the causality of activity modulation between these two regions (top-down P to V4 or bottom-up V4 to P). Our DCM analysis revealed that in low attentional load condition, activity in P is driven by V4 (model B.4), which is congruent with our PPI analysis. In high attentional load condition, activity in P is driven by attention which in turn weakens V4 activity, both directly and indirectly. This suggests a correlation between P and V4 in low attentional load condition with a predominance of bottom-up effects in V4 and P; and a decorrelation between P and V4 in high attention load condition with a predominance of top-down control from P exerted on V4, which may reflect control processes deployed to avoid automatic orienting of attention on peripheral stimuli. Thus, our DCM results not only confirm our PPI results but go well beyond these previous results by revealing the exact pattern of effective connectivity in healthy subjects underlying attention-mediated suppression of information in visual regions.

For the MDD patients, Family level inference showed that both low and high attentional load conditions influenced the processing of the task-irrelevant colored stimuli independently of the actual underlying model architecture. In a second step, model selection identified model A.4 (see Fig. 1) as the best model for the MDD data. These results indicate that major depression can profoundly modify the dynamics of the attentional network involved in this simple visual task. In model A.4, both low and high attentional loads directly drive P, which in turn modulates positively V4 either directly (P to V4 connection) or indirectly (V1-to-V4 gating). It may seem paradoxical that, in the MDD group, the drive of P by the "Low Attention" condition is not significantly different from zero, even though the inclusion of this connection is significant at the family and model level inference. In fact Bayesian estimation and averaging account for the covariance between the parameters. A strong or consistent covariance between a pair of parameters across subjects could lead to less reliable parameter estimates when looked at individually, leading to a reduced sensitivity (Rowe et al., 2010). Thus the estimated posterior of a single parameter could well be not significantly different from zero on average but the connection modelled still necessary from a model selection point of view.

In Desseilles et al. (2009), we found a decreased fMRI signal in V4 in MDD patients during low attentional load condition (vs. baseline, passive viewing condition) indicating an abnormal filtering of task-irrelevant stimuli in that condition as compared to healthy controls, but with no further decrease during high load. PPI analyses did not reveal any modification of functional connectivity between right intraparietal sulcus and V4 as a function of attentional load. However, the patients clearly showed a behavioral effect of attention (slower reaction times and more errors during high compared to low load) and activity level in parietal regions were also strongly modulated by task-load. Our DCM study clarifies this seemingly paradoxical result by revealing indirect effects of attention on the V1-to V4 forward connection (indirect positive drive from P) and on V4 (direct positive drive from P) that increased with attentional load. Our results demonstrate that the same dysfunctional connectivity network operates for low and high attentional loads, with counterbalanced effects within the model as the connections weights increased with higher attentional demand. Our DCM findings suggest that the network dynamics accounting for the effects of attention is abnormal not only for the high load condition but also during the low load condition. The confrontation of the results for the control and MDD groups also suggests that the predictions from the Load Theory about top-down control are valid in healthy (decrease of V4 activity by increased attentional load) but not in pathological condition (increase of V4 activity by increased attentional load driven either directly or indirectly via P). Decreased fMRI signal in V4 in MDD patients during low attentional load condition does not correspond to a better filtering of task-irrelevant stimuli but to a counterproductive recruitment of their attentional resources within a dysfunctional network dynamics.

One possible limitation of the present findings is that they are based on relatively small sample of subjects (14 subjects in each group). This is due to the very strict inclusion criteria of MDD patients (first episode, drug-free, no other psychiatric comorbidity) and very careful selection of the corresponding healthy matched controls. In addition, we restricted our analysis to the visuo-attentional system as was done in previous DCM studies in which attentional effects on visual responses were also explored (e.g. (Friston et al., 2003). Yet, in our previous classical fMRI analysis, we found some differential effects of load in medial prefrontal regions selectively in the patients. Whether this region also exerts or receives modulatory influences from the visuo-attentional network still need to be assessed in future studies.

Conclusion

In the present study, we show that attentional demands influence the dynamics of activity within a visuo-attentional network, but in a distinct manner in depressed patients and healthy controls. Moreover, these functional changes in MDD were observed when the patients were exposed to simple non-emotional visual processing. These results do not only improve our understanding of the cerebral mechanisms underlying cognitive-perceptual deficits in depression but may also have useful therapeutic implications. Indeed, because MDD patients present a counterproductive top-down recruitment of their attentional resources, an attentional training aiming at increasing the control of attentional resources could be proposed. By decreasing the disordered self-focused ruminations (Deyo et al., 2009), metacognitive therapy such as mindfulness-based strategies (Bondolfi et al., 2009) are likely to substantially improve top-down attentional recruitment. A promising extension of the present study might be to more systematically investigate to what extent such therapeutic interventions promote a normalization of regional brain activity or of network connectivity (or of both) in MDD.

Acknowledgments

Funded by the FRS-FNRS grant 3.4516.05, the research funds of the University of Liège (Belgium), Fondation Médicale Reine Elisabeth (FMRE), and "Interuniversity Attraction Poles Programme – Belgian State – Belgian Science Policy", the Horlait-Dapsens Foundation, the Belgian American Educational Fundation, Wallonie-Bruxelles International, and the Jean and Madeleine Vachoux Foundation. MD, TD, VS, CP and PM are supported by the FRS-FNRS. S.S. is supported by the Swiss National Science Foundation.

Appendix A. Supplementary data

Supplementary data to this article can be found online at doi:10.1016/j.neuroimage.2010.08.061.

References

Acs, F., Greenlee, M.W., 2008. Connectivity modulation of early visual processing areas during covert and overt tracking tasks. Neuroimage 41 (2), 380–388.

Almeida, J.R., Versace, A., et al., 2009. Abnormal amygdala-prefrontal effective connectivity to happy faces differentiates bipolar from major depression. Biol. Psychiatry 66 (5), 451–459.

Anand, A., Li, Y., et al., 2005. Activity and connectivity of brain mood regulating circuit in depression: a functional magnetic resonance study. Biol. Psychiatry 57 (10), 1079–1088.

Bondolfi, G., Jermann, F., et al., 2009. Depression relapse prophylaxis with Mindfulness-Based Cognitive Therapy: replication and extension in the Swiss health care system. J. Affect. Disord. 122 (3), 224–231.

Buxton, R.B., Wong, E.C., et al., 1998. Dynamics of blood flow and oxygenation changes during brain activation: the balloon model. Magn. Reson. Med. 39 (6), 855–864.

Chamberlain, S.R., Sahakian, B.J., 2006. The neuropsychology of mood disorders. Curr. Psychiatry Rep. 8 (6), 458–463.

Corbetta, M., Shulman, G.L., 2002. Control of goal-directed and stimulus-driven attention in the brain. Nat. Rev. Neurosci. 3 (3), 201–215.

Desseilles, M., Balteau, E., et al., 2009. Abnormal neural filtering of irrelevant visual information in depression. J. Neurosci. 29 (5), 1395–1403.

Deyo, M., Wilson, K.A., et al., 2009. Mindfulness and rumination: does mindfulness training lead to reductions in the ruminative thinking associated with depression? Explore (NY) 5 (5), 265–271.

Friston, K.J., 2002. Bayesian estimation of dynamical systems: an application to fMRI. Neuroimage 16 (2), 513–530.

Friston, K.J., Buchel, C., 2000. Attentional modulation of effective connectivity from V2 to V5/MT in humans. Proc. Natl. Acad. Sci. USA 97 (13), 7591–7596.

Friston, K.J., Buechel, C., et al., 1997. Psychophysiological and modulatory interactions in neuroimaging. Neuroimage 6 (3), 218–229.

Friston, K.J., Mechelli, A., et al., 2000. Nonlinear responses in fMRI: the Balloon model, Volterra kernels, and other hemodynamics. Neuroimage 12 (4), 466–477.

Friston, K.J., Harrison, L., et al., 2003. Dynamic causal modelling. Neuroimage 19 (4), 1273–1302.

Friston, K., Mattout, J., et al., 2007. Variational free energy and the Laplace approximation. Neuroimage 34 (1), 220–234.

Greicius, M.D., Flores, B.H., 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.

Hamilton, M., 1960. A rating scale for depression. J. Neurol. Neurosurg. Psychiatry 23, 56–62.

Irwin, W., Anderle, M.J., et al., 2004. Amygdalar interhemispheric functional connectivity differs between the non-depressed and depressed human brain. Neuroimage 21 (2), 674–686.

Johnstone, T., van Reekum, C.M., et al., 2007. Failure to regulate: counterproductive recruitment of top-down prefrontal-subcortical circuitry in major depression. J. Neurosci. 27 (33), 8877–8884.

Kumar, S., Stephan, K.E., et al., 2007. Hierarchical processing of auditory objects in humans. PLoS Comput. Biol. 3 (6), e100.

Lavie, N., 1995. Perceptual load as a necessary condition for selective attention. J. Exp. Psychol. Hum. Percept. Perform. 21 (3), 451–468.

Lavie, N., 2005. Distracted and confused?: selective attention under load. Trends Cogn. Sci. 9 (2), 75–82.

Lavie, N., Hirst, A., et al., 2004. Load theory of selective attention and cognitive control. J. Exp. Psychol. Gen. 133 (3), 339–354.

Mayberg, H.S., Liotti, M., et al., 1999. Reciprocal limbic-cortical function and negative mood: converging PET findings in depression and normal sadness. Am. J. Psychiatry 156 (5), 675–682.

Penny, W.D., Stephan, K.E., et al., 2004. Comparing dynamic causal models. Neuroimage 22 (3), 1157–1172.

Penny, W.D., Stephan, K.E., et al., 2010. Comparing families of dynamic causal models. PLoS Comput. Biol. 6 (3), e1000709.

Rowe, J.B., Hughes, L.E., 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.

Schlosser, R.G., Wagner, G., et al., 2008. Fronto-cingulate effective connectivity in major depression: a study with fMRI and dynamic causal modeling. Neuroimage 43 (3), 645–655.

Schwartz, S., Vuilleumier, P., et al., 2005. Attentional load and sensory competition in human vision: modulation of fMRI responses by load at fixation during task-irrelevant stimulation in the peripheral visual field. Cereb. Cortex 15 (6), 770–786.

Sheline, Y.I., Barch, D.M., et al., 2009. The default mode network and self-referential processes in depression. Proc. Natl. Acad. Sci. USA 106 (6), 1942–1947.

Spitzer, M.B., Gibbon, R.L., et al., 1994. Structured Clinical Interview for Axis I DSM-IV Disorders: Patient Edition (SCID-I/P, Version 2.0). New York, Biometrics Research Department, New York State Psychiatric Institute.

Stephan, K.E., Marshall, J.C., et al., 2007a. Interhemispheric integration of visual processing during task-driven lateralization. J. Neurosci. 27 (13), 3512–3522.

Stephan, K.E., Weiskopf, N., et al., 2007b. Comparing hemodynamic models with DCM. Neuroimage 38 (3), 387–401.

Stephan, K.E., Kasper, L., et al., 2008. Nonlinear dynamic causal models for fMRI. Neuroimage 42 (2), 649–662.

Stephan, K.E., Penny, W.D., et al., 2009. Bayesian model selection for group studies. Neuroimage 46 (4), 1004–1017.

Stephan, K.E., Penny, W.D., et al., 2010. Ten simple rules for dynamic causal modeling. Neuroimage 49 (4), 3099–3109.

Vasic, N., Walter, H., et al., 2009. Aberrant functional connectivity of dorsolateral prefrontal and cingulate networks in patients with major depression during working memory processing. Psychol. Med. 39 (6), 977–987.