



## Supplement

## Predicting EEG single trial responses with simultaneous fMRI and Relevance Vector Machine regression

Federico De Martino<sup>a,\*</sup>, Aline W. de Borst<sup>a,b</sup>, Giancarlo Valente<sup>a</sup>, Rainer Goebel<sup>a</sup>, Elia Formisano<sup>a</sup><sup>a</sup> Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands<sup>b</sup> Department of Neuroscience, University of Pisa, Pisa, Italy

## ARTICLE INFO

## Article history:

Received 29 November 2009

Revised 5 July 2010

Accepted 28 July 2010

Available online 4 August 2010

## ABSTRACT

The combination of electroencephalography (EEG) and functional Magnetic Resonance Imaging (fMRI) has been proposed as a tool to study brain dynamics with both high temporal and high spatial resolution. Integration through EEG–fMRI trial-by-trial coupling has been proposed as a method to combine the different data sets and achieve temporal expansion of the fMRI data (Eichele et al., 2005). To fully benefit of this type of analysis simultaneous EEG–fMRI acquisitions are necessary (Debener et al., 2006).

Here we address the issue of predicting the signal in one modality using information from the other modality. We use multivariate Relevance Vector Machine (RVM) regression to “learn” the relation between fMRI activation patterns and simultaneously acquired EEG responses in the context of a complex cognitive task entailing an auditory cue, visual mental imagery and a control visual target. We show that multivariate regression is a valuable approach for predicting evoked and induced oscillatory EEG responses from fMRI time series. Prediction of EEG from fMRI is largely influenced by the overall filtering effects of the hemodynamic response function. However, a detailed analysis of the auditory evoked responses shows that there is a small but significant contribution of single trial modulations that can be exploited for linking spatially-distributed patterns of fMRI activation to specific components of the simultaneously-recorded EEG signal.

© 2010 Elsevier Inc. All rights reserved.

## Introduction

The complementary nature of Electroencephalography (EEG; high temporal resolution and low spatial resolution) and functional Magnetic Resonance Imaging (fMRI; high spatial resolution and low temporal resolution) makes their combination appealing for investigating human brain dynamics (Dale and Halgren, 2001; Valdes-Sosa et al., 2009; Goebel and Esposito, in press; Ritter and Villringer, 2006; Debener et al., 2006; Logothetis, 2008).

Combining EEG and fMRI presents several challenges, which stem from the experimental and interpretational (Logothetis, 2008; Whittingstall and Logothetis, 2009; Sirotnin and Aniruddha, 2009) difficulty of linking measurements of electrophysiological potentials on the scalp (EEG) with local blood-oxygen-level-dependent (BOLD) signal changes as measured by fMRI, in most cases collected in different moments and conditions.

The development of simultaneous EEG–fMRI measurements offers several advantages over separate session recordings, despite the degraded EEG data quality. Simultaneous measurements guarantee identical sensory stimulation, perception and behaviour, and also

provide a unique way to study how ongoing fluctuations of brain signals can reflect anticipatory signals that interact with, and modify, the processing of sensory events (Debener et al., 2006).

In the last few years several techniques have been proposed to combine brain signals as measured by EEG and fMRI into one analytical framework. These methods range from the separate analysis of the data and subsequent juxtaposition of the results to truly integrated methods (Dale and Halgren, 2001; Valdes-Sosa et al., 2009). The combination of brain signals acquired with different imaging modalities relies on the assumption of common neuronal sources generating the different signals. In the case of EEG and fMRI this results in assuming that blood-oxygen-level changes depicted in the fMRI data and electrical activity at the level of the scalp as measured by EEG are generated from the same neuronal activity. Truly integrated analysis of simultaneous EEG–fMRI make this assumption explicit in the model while simple juxtaposition verifies this assumption post-hoc, when interpreting the results. The superiority of truly integrated analysis in the explanation of stimulus related effects with respect to the use of the single modalities has been recently investigated using an information theoretic approach (Ostwald et al., 2010).

Integrated analysis of multimodal data can be divided in two categories of methods: 1) fMRI-constrained EEG analysis (equivalent current dipoles estimates and continuous current dipoles estimates)

\* Corresponding author. Postbus 616, 6200 MD, Maastricht, The Netherlands. Fax: +31 433884125.

E-mail address: [f.demartino@maastrichtuniversity.nl](mailto:f.demartino@maastrichtuniversity.nl) (F. De Martino).

(e.g. Dale and Halgren, 2001; Goebel and Esposito, *in press*); 2) EEG-constrained fMRI analysis (fMRI correlates of EEG power modulations, trial-by-trial coupling) (e.g. Eichele et al., 2005).

In this article, we focus on the latter and use Relevance Vector Machine (RVM) regression to learn the relation between simultaneously collected EEG and fMRI responses at single trial level and derive a combined spatio-temporal representation of the data.

In most cases, a typical trial-by-trial analysis follows a massively univariate approach. First, the modulation of EEG single trial responses in both the time (ERP) and time-frequency (ERSP) domain are extracted from a single electrode (Eichele et al., 2005) or the weighted average multiple electrodes (e.g. by means of Independent Component Analysis of the EEG data; Debener et al., 2005). Second, these single trial responses are convolved with a canonical hemodynamic response function (HRF, Friston et al., 1998) in order to construct one or multiple EEG-based predictors. Third, these predictors are used in a conventional voxel-by-voxel statistical analysis of the fMRI data based on the General Linear Model (GLM). Applications, among others, to the processing of target-detection (oddball paradigm, Eichele et al., 2005; Goldman et al., 2009), performance monitoring (Debener et al., 2005), decision making (Mulert et al., 2008), working memory maintenance (Scheeringa et al., 2009) and simple auditory responses (Mayhew et al., 2010) have proved the relation between different spatial networks (fMRI) and the electroencephalographic response (both at the level of event related responses and event related spectral power). These ‘massively univariate’ approaches, however, do not consider the multivariate nature of the fMRI data and thus can be suboptimal in detecting relations between EEG signals and fMRI activation patterns.

As an alternative, the use of multivariate analysis of the fMRI data based e.g. on Independent Component Analysis (ICA) has been proposed (Eichele et al., 2008). In this approach, a parallel (multivariate) decomposition of the two data sets (*spatial* ICA in fMRI and *temporal* ICA in EEG) is performed. The temporal profiles of the fMRI spatial modes are correlated post-hoc with the EEG single trial responses. As the fMRI data are decomposed “blindly” (without making use of any a-priori hypothesis, e.g. the temporal profile of the EEG modulations), this approach leaves the problem of finding the unique (if present) fMRI network that explains most of the information in the trial-by-trial modulation present in the EEG data unresolved.

The use of *correlation* as the measure of the coupling between surface EEG and fMRI is reliable in explaining the relations between the available data sets. However, in order to estimate an fMRI pattern (i.e. multivariate model) associated with a specific EEG signal (thus not using ‘blind’ models of the fMRI data or EEG data), the use of a correlation-based approach is highly prone to overfitting (i.e. explaining EEG modulations with random multivariate noise in the fMRI). In order to avoid overfitting, the generalization to different data sets (i.e. *predictive* modelling) can be used instead of correlation to validate the estimated relation between fMRI patterns and EEG signals.

Recently, predictive models have been introduced to investigate the relation between multivariate fMRI BOLD signals and a continuous experimental variable (Formisano et al., 2008; Friston et al., 2008; Valente et al., *submitted for publication*). These methods are particularly suited to the analysis of fMRI data given the typical dimensionality of the problem (number of voxels  $\gg$  number of samples). Beyond the simple correlation, these methods allow, after a learning phase, the prediction of the experimental variable (i.e. stimulus, behaviour) by exploiting the multivariate information present in the data.

We have recently introduced the use of multivariate machine-learning based regression to the multi-modal analysis of the coupling between two brain signals. More specifically, we have shown the superiority of multivariate predictive models, compared to previously introduced univariate and multivariate methods, in *predicting*

ongoing EEG power oscillations from simultaneously acquired BOLD fMRI (De Martino et al., 2010).

Here we extend the use of multivariate predictive models to trial-by-trial coupling. We test the ability of this method to predict EEG single trial event related potentials (ERPs) or event related spectral power (ERSP) in a mental visual imagery task (de Borst et al., *submitted*), involving processing of an acoustic cue, imagining of a complex visual scene and a behavioral response to a control visual target.

We show that prediction of EEG from fMRI is largely influenced by the overall filtering effects of the hemodynamic response function, but also that patterns of fMRI responses allow predicting a small but significant portion of ERP and ERSP modulations.

## Methods

### General description of the approach

Raw EEG data are pre-processed in order to remove the artefacts induced by the Magnetic Resonance Imaging (MRI) environment (gradient artefact and ballistocardiogram (BCG) artefact; Allen et al., 2000; Niazy et al., 2005; Debener et al., 2007, see below). After epoching and removal of trials exhibiting clear artefactual responses, EEG data are decomposed using temporal Independent Component Analysis (tICA; Makeig et al., 1997). As for conventional trial-by-trial analysis, single trial responses (amplitude and latency) in both time (ERP) and time-frequency (ERSP) domain are then extracted from selected ICs (see below). Finally, an EEG-based predictor is obtained by convolution of single trial responses with a canonical HRF (Friston et al., 1998), and resampled at the temporal resolution of the fMRI data.

The use of predictive models requires both data sets (i.e. the pre-processed fMRI time series and EEG-based predictors) to be divided into training and testing sets. The training set is used to learn the coupling between fMRI and EEG at a trial-by-trial level using Relevance Vector Machines (RVM). A map depicting the contribution of each voxel to the learned coupling is obtained. The test data are used to assess the validity of the learned coupling on the basis of the *correlation* between EEG fluctuations from test data (after convolution with a canonical HRF) and prediction obtained from fMRI-based RVM (accuracy).

### Relevance Vector Machine Regression

Below we review briefly the concepts of RVM which are relevant to our application. For a complete description of RVM we refer to (Tipping, 2001; Formisano et al., 2008). Let us consider a data set  $D$  consisting of an fMRI time series, represented by an  $N \times V$  matrix  $\mathbf{X}$  ( $N$  being the number of volumes and  $V$  the number of voxels), and an EEG based predictor, represented by the  $N$  dimensional vector  $\mathbf{p}$ . Such dataset can be described as a collection of  $N$  pairs  $(\mathbf{x}_i, p_i)$ , where  $\mathbf{x}_i$  denotes a sample vector of dimension  $V$  (one volume of the fMRI time series) and  $p_i$  the corresponding one-dimensional label (the EEG trial-by-trial modulation of a specific component).

In machine learning for fMRI data analysis it is common choice to use a linear model due to a typically high feature (i.e. voxel) to sample (i.e. scan) ratio. A standard linear model has the following form:

$$p = y(\mathbf{x}, \mathbf{w}) + \varepsilon \quad (1)$$

where  $y(\mathbf{x}, \mathbf{w})$  is the deterministic input-output mapping part and  $\varepsilon$  accounts for the noise in the measurements. The deterministic mapping can be modelled as (Bishop, 2006):

$$y(\mathbf{x}, \mathbf{w}) = \mathbf{w}^T \phi(\tilde{\mathbf{x}}) \quad (2)$$

where  $\phi: \mathcal{R}^V \rightarrow \mathcal{R}^M$  is a mapping (i.e. kernel) from  $V$ -dimensional space of  $\mathbf{x}$  into an  $M$ -dimensional one and  $\mathbf{w}$  is an  $M$ -dimensional vector of parameters.

Without loss of generality, we refer to the first functional run ( $\mathbf{X}_1$ ) as the training data set and subsequent runs as test set. Suppose that both the time courses of the voxels in the fMRI data and the simultaneously recorded EEG signals have zero mean. The choice of a linear kernel results in the following model:

$$y(\mathbf{X}_1, \mathbf{w}) = \mathbf{X}_1 \mathbf{X}_1^T \mathbf{w} = \mathbf{K} \mathbf{w} \quad (3)$$

with  $\mathbf{w}(n_1 \times 1)$  being the model weights vector and  $\mathbf{K} = \mathbf{X}_1 \mathbf{X}_1^T (n_1 \times n_1)$  the linear kernel constructed from the starting training dataset  $\mathbf{X}_1 (n_1 \times V)$ .

The RVM training aims at finding an estimate of the posterior distribution of the weights  $\mathbf{w}$ . This posterior distribution can then be used to perform predictions on a new dataset (second functional run). Denoting with  $\tilde{\mathbf{w}}$  the estimated posterior mean, then the prediction on the second functional run  $\mathbf{X}_2 (n_2 \times V)$  can be expressed as:

$$\tilde{p}_2 = \mathbf{X}_2 \mathbf{X}_1^T \tilde{\mathbf{w}} \quad (4)$$

It is possible to express the prediction in terms of maps:

$$\tilde{p}_2 = \mathbf{X}_2 \tilde{\mathbf{M}} \quad (5)$$

with

$$\tilde{\mathbf{M}} = \mathbf{X}_1^T \tilde{\mathbf{w}} \quad (6)$$

where  $\tilde{\mathbf{M}} (V \times 1)$  can be interpreted as a map of relative contribution of the different voxels to the final prediction.

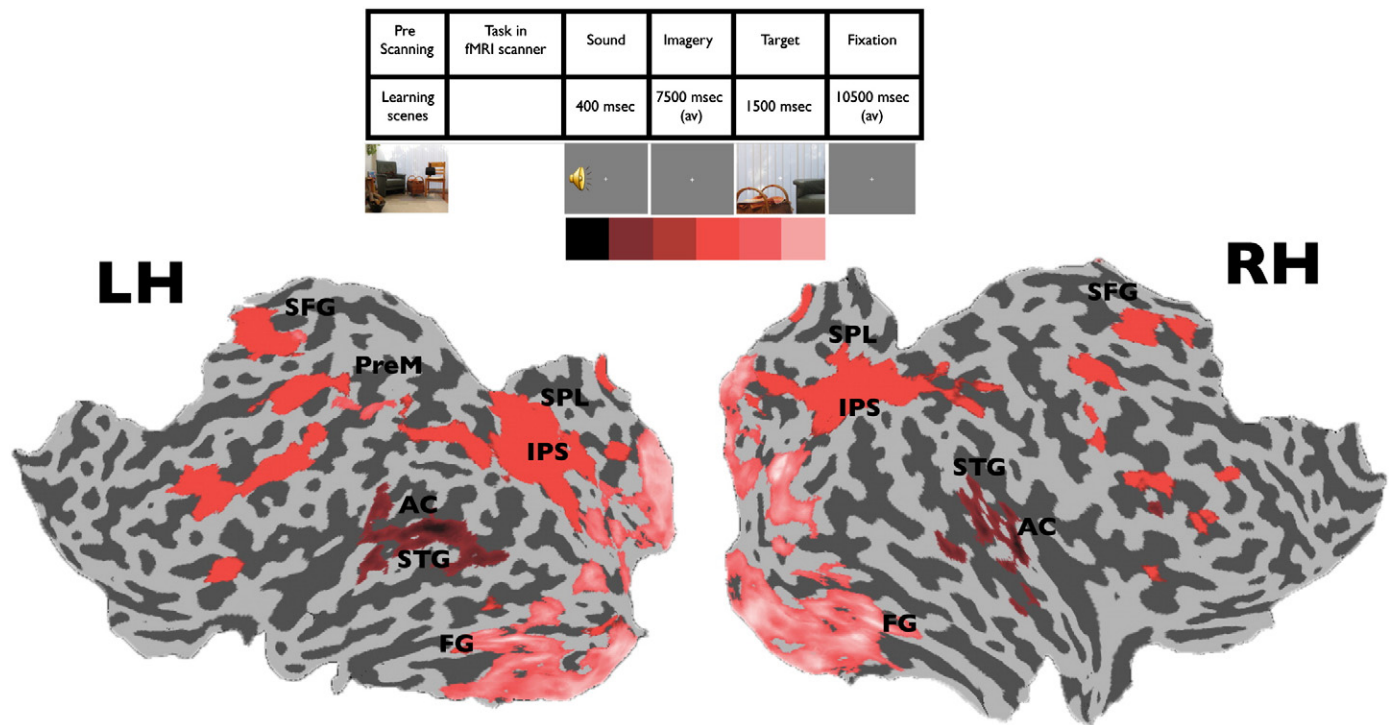
## Data acquisition

Simultaneous EEG-fMRI data were collected from three healthy subjects (all female with normal or corrected-to-normal vision, mean age of 25 years), three runs per subject, and 72 trials per run. Fig. 1 (top) shows the timeline of one exemplary trial. Each trial of the simultaneous EEG-fMRI experiment started with an auditory cue (one of three 400 ms tones [800, 1000 and 1200 Hz]) presented in silent gaps during the acquisition (see below). During the subsequent interval the participant imagined the scene, which they previously had learned to associate with the cue. After a variable delay (mean of 7500 ms) a mirrored or non-mirrored fragment of the scene was visually presented as a target. The participants decided, using a button-box, whether the fragment matched the imagined scene, or was mirrored. All trials ended with a fixation period (mean of 10500 ms) during which the participants performed no task.

Prior to the measurements, participants were trained on imagining three scenes, associated with three auditory cues. After training the participants received task instructions and were asked to vividly visualize the correct scene after each cue in the main experiment, while fixating at the fixation cross.

## EEG data

EEG data were recorded using a 64-channel high-input impedance amplifier system specifically designed to operate in the MRI environment (Brainproducts, Munich, Germany). The setup consisted of two 32-channel MR compatible EEG plus amplifiers powered by a rechargeable power unit. The amplifiers were placed directly behind the scanner bore inside the MR room, which allowed the use of short wires with a total length of about 1.2 m from recording electrodes to amplifier. Sintered Ag/AgCl ring electrodes with built-in 5 k $\Omega$  resistors were used. Data were recorded from 62 equidistant scalp sites



**Fig. 1.** (Top) General description of the mental visual imagery task including a training phase (pre-scanning session) and the trials performed during the scanning session. A typical trial includes the presentation of an auditory cue, an imagery period, the presentation of the visual target and the final fixation period. (Bottom) Spatio-temporal pattern of BOLD activation (fixed effect, deconvolution analysis) projected on the flattened representation of the mean cortex of the three subjects after cortex based alignment. The color in the map codes the time (TR) of the stick predictor whose beta weight contribute the most to the response (early responses in black late responses in white). Major anatomical landmarks are highlighted in both hemispheres (Auditory Cortex; Intra Parietal Sulcus; Superior Temporal Gyrus; Superior Parietal Lobule; Pre-Motor cortex; Superior Frontal Gyrus).



mounted in a cap system (EasyCap, Falk Minow Services, Herrsching, Germany). Continuous data were also recorded from one electrode placed below the left eye to monitor eye blinks and another electrode placed at the lower back for electrocardiogram (ECG) recording. All 64-channel data were referenced to the vertex. The data were recorded with a pass-band of 0.016–250 Hz and digitized with 5000 samples/s at 16-bit resolution, resulting in a dynamic range of 16.38 mV. The amplified signal was transmitted via fiber-optic cables to a recording PC placed outside the MR room. Electrode impedances were maintained below 20 k $\Omega$  before recordings.

The EEG data processing was carried out using EEGLAB (Delorme and Makeig, 2004). Functional MRI gradient and ballistocardiogram (BCG) artefacts were removed using the EEGLAB plug-in FMRIB 1.21 (Niazy et al., 2005), as developed by the Centre for the Functional MRI of the Brain (Oxford, UK). For each subject, temporal Independent Component Analysis was used on the continuous data to extract 62 components after concatenating the data from the three runs. The Independent Components (ICs) were visually reviewed in order to discard components reflecting residual BCG or gradient artefacts (Debener et al., 2007).

After removing channels with a high noise level, the continuous data were epoched around the auditory cue (–1000 to 7500 ms) and the visual target (–1000 to 1000 ms). Epochs exhibiting clear artefacts (eye blinks or muscle artefacts) were removed prior to subsequent processing steps. Independent component analysis was used on the concatenated epoched data sets to find components of interest associated to the presentation of the auditory cue, the imagery processing and the presentation of the visual target. The use of ICA on all available data preserves the single trial variability and allows for a more accurate estimation of the components (i.e. better de-noising). Furthermore the unsupervised nature of ICA does not bias the subsequent use of predictive models as no labels are used in the extraction of the components. In this study, relevant independent components (ICs) were manually selected based on the inspection of topography, ERP, ERSP, consistency across subject and correspondence to results obtained in a separate EEG study conducted outside the MR scanner (de Borst et al., submitted for publication). Time frequency decomposition (Morlet wavelets; window size = 1119 ms) was applied to the selected ICs time courses in order to extract the power modulations in the band [2 – 28 Hz].

In order to estimate the single trial event related responses (ERP), first the mean in a baseline period [–200 to 0 ms] was removed trial-by-trial in the selected IC time course, second the maximum (minimum) value and its latency were extracted from temporal windows (50 ms) selected around the peaks of the average ERP response (e.g. auditory N1, visual P1). This procedure resulted in one single point per single trial as representative of the selected peak response. Note that the single trial latency estimation as adopted here (limited to a 50 ms window around the mean ERP peak) likely leads to the same results if compared to the conventional strategy in EEG-fMRI studies of attributing to each trial a latency equal to the mean latency of the ERP peak.

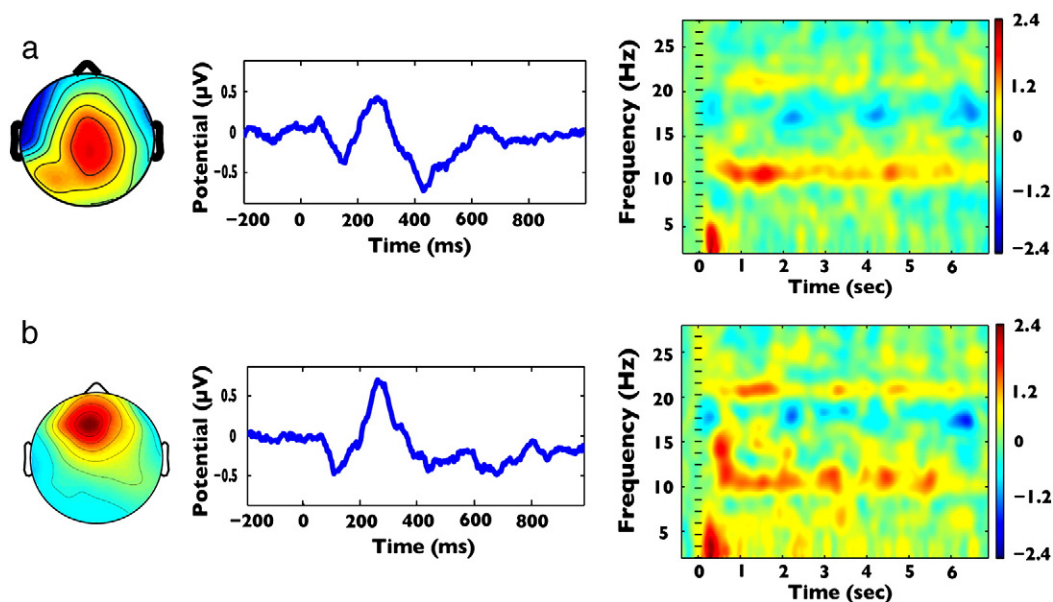
ERSP single trials were characterized by a continuous response over extend time windows. First the mean in a baseline period [–200 to 0 ms] was removed trial-by-trial for each frequency, second the average across a selected frequency window was considered as an estimate of the time-varying power modulation over a selected time interval (i.e. in this case the estimation procedure resulted in a signal with several points for each trial which was subsequently convolved with the hemodynamic response function). Frequency and time windows were chosen based on the average event related spectral power (ERSP) of the selected ICs. All single trial measures were estimated using custom software implemented in Matlab (The MathWorks, Inc.).

Single trial information from both the ERP and the ERSP was subsequently convolved with a standard hemodynamic response function (HRF, Friston et al., 1998) and re-sampled to the fMRI sampling rate (0.5 Hz, see below) to obtain the final predictors. Prior to the regression procedure all predictors were z-normalized.

#### fMRI data

Functional magnetic resonance time series were acquired in a 3 T system (Siemens Allegra, Erlangen, Germany). Functional runs consisted of 22 axial slices obtained with a T2\*-weighted gradient echo, EPI sequence (TR 2 s; TA 1.4 s; FOV224 $\times$ 224; matrix size64 $\times$ 64, voxel size = 3.5 $\times$ 3.5 $\times$ 4 mm). Anatomical images were obtained using a high resolution (1 $\times$ 1 $\times$ 1 mm) T1-weighted sequence.

After standard pre-processing steps (slice-scan-time correction; head motion correction; temporal filtering; resampling in three-dimensional standard, Talairach, space) data sets of the three functional runs of each subject were used for the multivariate regression analysis



**Fig. 2.** Independent Components from the analysis of the EEG data epoched with respect to the auditory cue [–1000 to 7500 ms]. a) Group topography, event related potential (ERP) and event related spectral power (ERSP) of the IC representing a typical auditory response. b) Group topography, ERP and ERSP of the IC associated with imagery processing.

using Relevance Vector Machines (RVM). Training was performed using the data of two runs (e.g. run 1 and 2) and predictions were evaluated on the data from the left out run (i.e. leave-one-run-out strategy).

Voxel pre-selection was based on conventional univariate analysis (GLM with predictors accounting for the auditory cue, delay period and visual cue) performed only on the training data (i.e. separately for each cross-validation step). The most significant ( $N=5000$ ) voxels according to an F-statistics were used for the subsequent RVM training procedure.

#### Evaluation of single trial modulations in multivariate predictions

Considering the presence of the hemodynamic response, a high value of cross-correlation between predicted and actual response may simply reflect a good prediction of the overall response but not necessarily of the single-trial EEG fluctuations (see Results). To evaluate the contribution of the actual single trial EEG modulations to our predictions we employed three different procedures.

First, we modelled *post hoc* (using a GLM analysis) the time courses of the RVM-based predictions in terms of “overall BOLD response” (modelled as a standard fMRI predictor with constant amplitude and duration equal to the auditory cue) and “EEG modulations” (obtained from the EEG-based predictor, informed by both amplitude and latency of test-data and orthogonalized with respect to the overall BOLD response). To evaluate the contribution of considering EEG information during learning in explaining EEG modulations in test data, we compared the beta coefficients for the second predictor for RVM-based predictions obtained with and without EEG information (see Fig. 5). Comparison was based on a two-way ANOVA for repeated measures with Subject (s1, s2, s3) and Training (EEG-informed, no-EEG) as factors. Significance values in Fig. 5 ( $p<0.05$ ) refers to the factor “Training”; in all cases, the factor “Subject” and the interaction term were not significant.

Second, we orthogonalized both fMRI time series (voxel-by-voxel) and EEG-based predictors with respect to the overall BOLD response to the experimental protocol. Training and testing was then performed on the orthogonalized time series. This voxel-by-voxel orthogonalization is necessary as methods of multivariate regression are currently limited to “learning” of single predictors and treat multiple predictors sequentially/separately.

Third, we extracted single trial values from our prediction time courses and evaluated the coupling of predicted and real EEG modulations on a trial-by-trial basis (Debener et al., 2005). This procedure largely reduces the contribution of the overall hemodynamic response to the obtained correlation values. Furthermore, the significance of correlations is not inflated by the intrinsic autocorrelation of time courses.

## Results

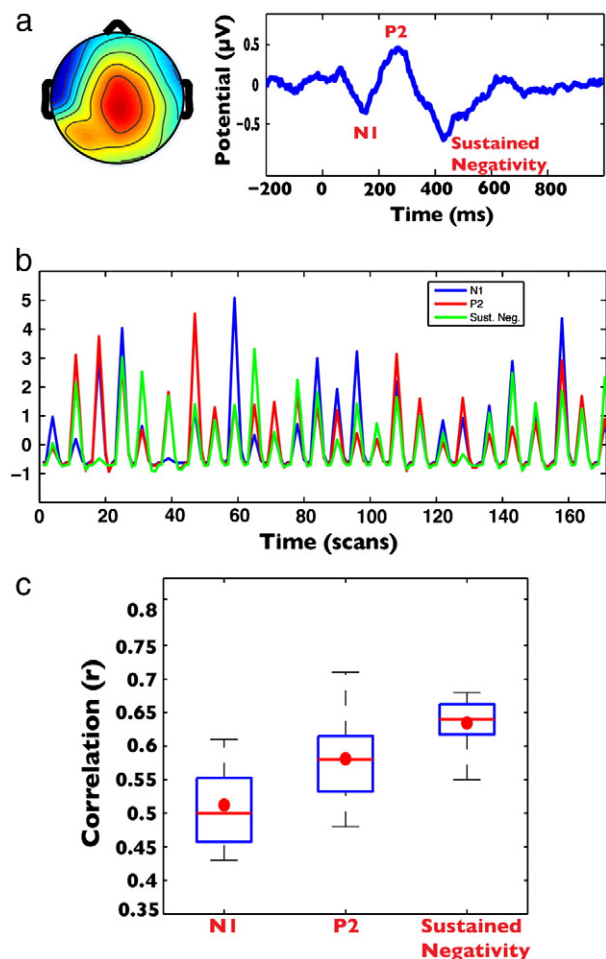
#### fMRI data analysis

Fig. 1 shows the spatio-temporal pattern of BOLD activation, based on a fixed effect deconvolution analysis (F map,  $p=5 \cdot 10^{-15}$ , Bonferroni correction performed based on the number of voxels), projected on the flattened representation of the mean cortex of the three subjects after cortex-based alignment (Goebel et al., 2006). The color in the map codes for the time of the stick predictor for which the beta weights contribute the most to the response (e.g. a dark red color corresponds to the first TR after sound onset). After the presentation of the auditory cue, early activation was detected in the auditory cortex (AC) and superior temporal gyrus (STG) (black and dark red color). In the subsequent mental imagery period, activation extended to a large fronto-parietal network including superior parietal lobule (SPL) and intra parietal sulcus (IPS) together with the mesial superior frontal gyrus (SFG) (extending from the supplementary motor area to

the anterior cingulate cortex) (PreM) were activated during the imagery period (red color). Finally, a large network of visual areas among which the fusiform gyrus (FG) was associated with the presentation of the target picture (light-pink to white). These results are consistent with data acquired in a separate non-simultaneous fMRI study (de Borst et al., submitted for publication) and with previous studies of visual mental imagery employing a time-resolved fMRI approach (Formisano et al., 2002).

#### EEG data analysis

Results obtained on the EEG data epoched with respect to the presentation of the auditory cue [−1000 – 7500 ms] are summarized in Fig. 2. After pre-processing and de-noising, ICs reflecting clear event related potentials and power modulations could be detected (Fig. 2). The most consistent of these ICs included: a) a component with a typical ERP pattern and central topography reflecting the response to the auditory instruction (Fig. 2a) and b) a component with a fronto-central topography and prominent ERSP power modulations in the 9 – 12 Hz and 19 – 23 Hz frequency bands (Fig. 2b). In a separate EEG-only study, this latter component was found to be associated with the imagery



**Fig. 3.** Results obtained from the prediction of the prominent ERP peaks of the early auditory IC. a) Group topography and event related potential (ERP) of the IC representing a typical auditory response (the three prominent peaks are highlighted). b) Example of predictors obtained for the three peaks (N1, P2, Sustained Negativity) in one exemplary subject and run (first 180 TRs). c) Group accuracy (correlation) in the prediction of the three selected peaks. Median (red line) lower and upper quartiles (blue box) and mean (red dot) accuracy across subjects are plotted for each of the three peaks. The mean accuracy across subjects was computed after applying a Fisher  $r$  to  $z$  transform to the single subject accuracies and then transforming back to  $r$  the mean score (red dot).

processing and related to the behavioural response of the subjects (i.e. showing a difference in the ERSP between “fast” and “slow” responding trials). Considering consistency, physiological plausibility and functional interest, these two ICs were considered as signal of interest for the fMRI-based predictions.

#### Prediction of auditory ERPs

To understand the relation between fMRI and EEG responses following the auditory instruction, RVM regression of fMRI data was used to predict the prominent ERP peaks of the early auditory IC (Fig. 3a). The predictors (after HRF convolution and resampling to the fMRI sampling frequency) obtained for the three separate peaks (N1, P2 and sustained negativity, see e.g. Fig. 3b) were used for RVM training in three separate analyses. Cross-correlation between predicted and actual response was high and significant in all cases (N1,  $cc=0.51$ ,  $p=5.4 \cdot 10^{-32}$ ; P2,  $cc=0.58$ ,  $p=4.1 \cdot 10^{-42}$ ; Sustained Negativity,  $cc=0.63$ ,  $p=1.9 \cdot 10^{-51}$ ,  $p$ -values computed using the number of points in each predicted run [ $n=450$ ], see Fig. 3c and Table 1 for single run predictions).

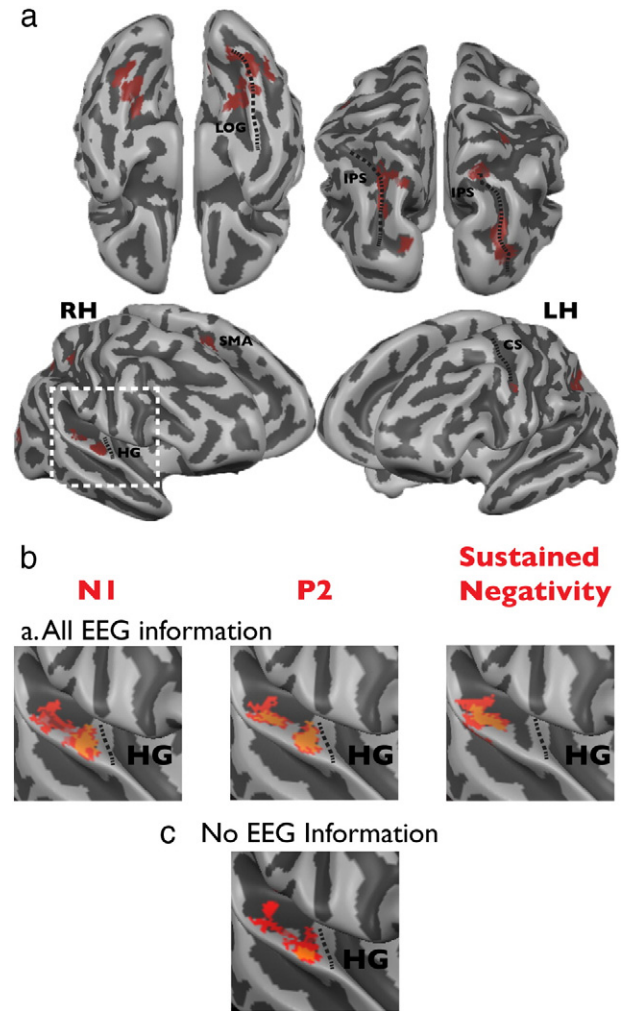
High correlation values were obtained also when learning was based on non-EEG informed predictors (constant amplitude and latency defined by the duration of the auditory cue, see Table 1 for single subject/run and mean accuracies). This indicates that high values of correlation between predicted and actual response and their significance (computed on the total number of 450 points) simply reflect a good prediction of the hemodynamic response to the sound presentation but not necessarily of the single-trial EEG fluctuations.

Fig. 4a shows the “group” map, resulting from the training of the RVM on the three different ERP based predictors (N1, P2 and sustained negativity). The map represents the union of the maps obtained for the three different peaks. To obtain these maps, single subject maps were sampled on the cortex, cortically realigned, thresholded (to select the 40% most relevant vertices for the prediction) and binarized. A vertex in the group map is colour coded if present in at least two of the three single subject maps. The map includes regions in the auditory cortex (mostly right hemisphere), bilateral intraparietal sulcus (IPS), supplementary motor area (SMA) and lateral occipital gyrus (LOG). Fig. 4b shows a zoomed-in portion of the predictive patterns in the right auditory cortex associated with the three selected peaks. For comparison, the map obtained after non-EEG informed training (i.e. one predictor with constant amplitude and duration equal to the auditory cue is used to predict all three peaks) is shown in Fig. 4c.

**Table 1**

Comparison between EEG-informed and uniform learning. Accuracies are computed as the correlation between fMRI-based predictions and real EEG signals after convolution with a canonical HRF response. Mean accuracies were computed by transforming individual  $r$ -values into  $z$ -values and re-transforming the resulting mean  $z$  into an  $r$ -value. Statistical significance is computed using  $n$  equal to the minimum number of points present in each run ( $n=450$ ).

		All EEG Info			No EEG Info		
		Run1	Run2	Run3	Run1	Run2	Run3
S1	N1	0.46	0.54	0.61	0.51	0.55	0.60
	P2	0.59	0.63	0.71	0.59	0.62	0.62
	Sust. Neg.	0.62	0.64	0.66	0.60	0.67	0.70
S2	N1	0.43	0.50	0.52	0.54	0.55	0.59
	P2	0.55	0.58	0.61	0.54	0.55	0.53
	Sust. Neg.	0.61	0.64	0.67	0.66	0.60	0.62
S3	N1	0.45	0.49	0.59	0.50	0.55	0.58
	P2	0.48	0.56	0.48	0.48	0.55	0.63
	Sust. Neg.	0.55	0.63	0.68	0.59	0.60	0.75
Mean	N1	0.51			0.55		
		$p=5.4 \cdot 10^{-32}$			$p=3 \cdot 10^{-37}$		
	P2	0.58			0.57		
		$p=4.1 \cdot 10^{-42}$			$p=5.2 \cdot 10^{-41}$		
	Sust. Neg.	0.63			0.65		
		$p=1.9 \cdot 10^{-51}$			$p=4.1 \cdot 10^{-56}$		

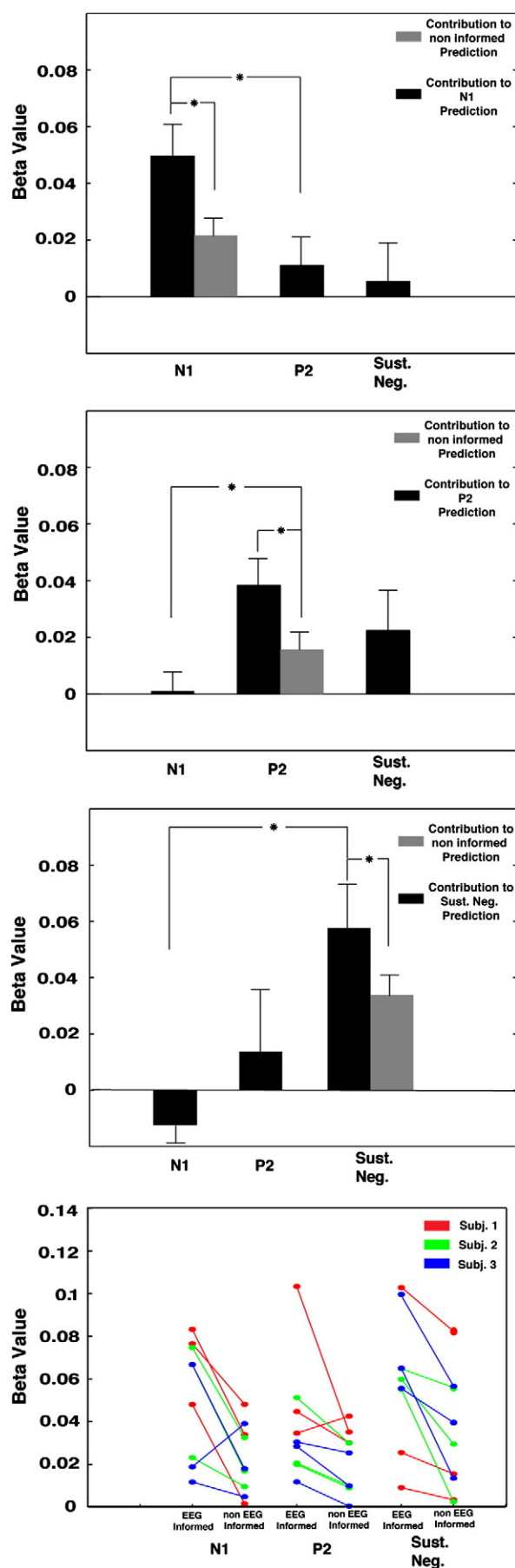


**Fig. 4.** Group maps obtained from EEG-informed and non EEG-informed training. a) Union of the group map (after cortex based alignment) obtained from training based on the single trial modulations of the N1, P2 and sustained negativity in each subject. Single subject maps were sampled on the cortex, thresholded (in order to select the 40% most relevant vertices for the prediction) and binarized. A vertex in the group map is colour coded if present in at least two of the three single subject maps. Major anatomical landmarks are highlighted (Heschl's Gyrus, Central Sulcus, Intra Parietal Sulcus, Supplementary Motor Area and Lateral Occipital Gyrus). b) Zoomed-in portion (right auditory cortex, group map) of the maps obtained using three different sources of ERP information during training (N1, P2 and sustained negativity). The maps are displayed after thresholding single subject maps to select the best 60% (red) and 40% (yellow) of the voxels. HG = Heschl's gyrus. c) Zoomed-in portion (right auditory cortex, group map) of the map obtained using non EEG-informed training (one predictor with constant amplitude and duration equal to the auditory cue is used to predict all three peaks). The maps are displayed after thresholding the single subjects to select the best 60% (red) and 40% (yellow) of the voxels. HG = Heschl's gyrus.

Although the various maps are very similar, the map for the prediction of auditory electrical activity in the N1 and P2 time windows includes both a region in the lateral aspect of the Heschl's Gyrus and a more posterior region on the planum temporale; prediction maps of later electrical activity (Sustained Negativity) only includes the region in planum temporale. Considering the small number of subjects and the empirical thresholding procedure employed, however, these maps are to be considered as a qualitative description of the informative patterns and the statistical validity of the observed difference needs to be addressed in further studies with a larger number of subjects.

Fig. 5 shows the results of a first analysis we conducted to examine the contribution of single trial EEG modulations (see Methods). With a post hoc GLM, we compared the beta values associated to the EEG





modulations obtained with the EEG-based learning predictions and non-EEG based learning predictions for the same data (two-way ANOVA for repeated measures with “Subject” and “Training” as factors). Each prediction (N1, P2, sustained negativity) is illustrated in a different panel. It can be seen that the beta values associated to the EEG information are significantly ( $p < 0.05$ ) higher for the prediction obtained by the EEG informed training (black bars) compared to non-EEG-informed training (gray bars). Furthermore, the beta values are higher (in most cases significantly at  $p < 0.05$ , see Fig. 5 for detail) for the targeted EEG component. For example, when N1 information is used during training (top panel), the beta value associated to the N1 variations in test data (orthogonalized with respect to the protocol) is significantly higher than the beta obtained when analyzing the non-informed training prediction are higher than betas obtained for P2 (significantly at  $p < 0.03$ ) and Sustained Negativity (although not significantly  $p = 0.051$ ). Similar results are found for P2 (middle panel) and sustained negativity (lower panel). The last panel of Fig. 5 shows the single subject/run beta values associated to the three relevant peaks (N1, P2, Sustained Negativity) obtained with EEG informed and non EEG-informed predictions.

Fig. 6 shows the results of a second analysis we conducted to examine the contribution of single trial EEG modulations (see Methods). After the standard hemodynamic response to the stimulus was removed from the fMRI time-courses and EEG-based predictors, the correlations between predicted and actual responses in test data were much smaller, but still significant for P2 and Sustained Negativity (Fig. 6b, N1,  $cc = 0.07$ ,  $p = 0.07$ ; P2,  $cc = 0.08$ ,  $p = 0.03$ ; Sustained Negativity,  $cc = 0.11$ ,  $p = 0.007$ ,  $p$ -values computed using the minimum number of points in each prediction [ $n = 450$ ]; see also prediction accuracies at single-subject/run level in Table 2). The maps obtained from training using single trial information of the three auditory ERP peaks (N1, P2, Sustained Negativity) were similar to those obtained with non-orthogonalized data (Fig. 6).

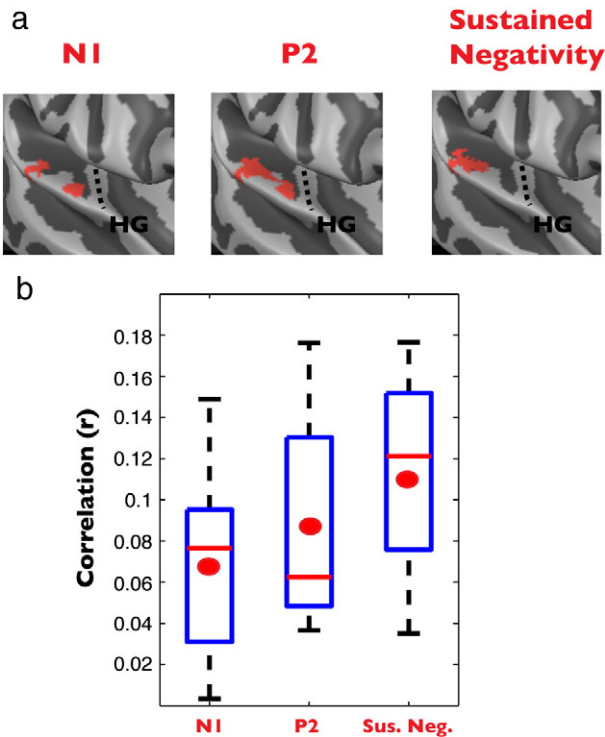
Finally, the predictions obtained from EEG-informed training on both non-orthogonalized (Fig. 3 and Table 1) and orthogonalized data (Fig. 6 and Table 2) were evaluated in terms of trial-by-trial correlations. Estimates of single trial EEG modulations were obtained averaging the signal values at three TRs after each stimulus onset for both the RVM-based predicted response and the real EEG modulations. The results of this analysis are reported in Table 3 ( $p$ -values were computed using the minimum number of trials in each prediction [ $n = 68$ ]). This table also reports the correlations obtained in case of non-EEG informed training. It can be seen that – although variable across runs and subjects – correlation values obtained with EEG informed training (original and orthogonal data) are generally higher than those obtained with non-EEG informed training.

Overall, these three different analyses of the auditory event-related responses indicate that considering the information of single-trial modulations in training data provides a small but significant improvement in predicting modulations of corresponding EEG responses in test data.

#### Prediction of mental imagery related oscillations

Fig. 7 shows the results obtained from the prediction of two time-frequency windows ([9 – 12 Hz], [19 – 23 Hz], highlighted in Fig. 7a) in

**Fig. 5.** Statistical analysis performed on the prediction time course obtained using N1 (first panel), P2 (second panel) and Sustained Negativity (third panel) information for training. Black bars represent the beta values associated with the orthogonalized single-trial EEG predictor in the case of EEG-informed training. Gray bars represent the beta values associated with the same predictor in the case of uninformed learning. Significant ( $p < 0.05$ ) differences are highlighted (asterisk). Bottom panel shows beta values associated to the relevant EEG modulations obtained from EEG-informed and non EEG-informed predictions for each individual run/subjects.



**Fig. 6.** Results obtained from the prediction of the prominent ERP peaks of the early auditory IC after decorrelating the fMRI signals and predictors with respect to the experimental protocol. a) Zoomed-in portion (right auditory cortex, group map) of the maps obtained using three different sources of ERP information during training (N1, P2 and sustained negativity). The maps are displayed after thresholding the single subjects to select the best 40% of the voxels. HG = Heschl's gyrus. a) Group accuracy (correlation) in the prediction of the three selected peaks. Median (red line) lower and upper quartiles (blue box) and mean (red dot) accuracy across subjects are plotted for each of the three peaks. The mean accuracy across subjects was computed after applying a Fisher  $r$  to  $z$  transform to the single subject accuracies and then transforming back to  $r$  the mean score (red dot).

the ERSF of the IC associated with mental visual imagery (7a). Fig. 7b shows the group map (after cortex based alignment, thresholding as in Fig. 3d) obtained from the training of the RVM in each subject. The map highlights areas in bilateral intra parietal sulcus (IPS), supplementary motor area (SMA) and lateral occipital gyrus (LOG). Group accuracy (computed as in Fig. 3c) in the prediction of the two selected windows is presented in Fig. 7c. Significance for the predictions was computed based on the mean  $z$  value and the minimum time samples in each run ( $n = 450$ ). Obtained accuracy values are lower - compared to the ERP results [ $9 - 12$  Hz]  $cc = 0.21$ ,  $p = 3.7 \cdot 10^{-6}$ ; [ $19 - 23$  Hz]  $cc = 0.19$ ,  $p = 2.1 \cdot 10^{-5}$ ).

#### Prediction of visual ERPs

Fig. 8 summarizes the results obtained on the EEG data epoched with respect to the visual target [ $-1000 - 1000$  ms]. The top panel shows the selected independent component (group topography and ERP) characterized by a posterior distribution typical of early visual processing. The middle panel shows the group map (after cortex based alignment, thresholding as in Fig. 3c) obtained from the training of the RVM in each subject to the prediction of the P1 peak. The map highlights areas in the calcarine sulcus (CalS) and lateral occipital gyrus (LOG). Group accuracy is reported in the lower panel ( $cc = 0.75$ ;  $p = 4.9 \cdot 10^{-86}$ , computed based on the mean  $z$  value and the minimum points in each run [ $n = 450$ ]) was obtained. While highly significant these prediction accuracies are largely determined by the overall hemodynamic response to the visual stimulus.

**Table 2**

Results of EEG-informed learning with fMRI time series orthogonalized with respect to the overall auditory response. Accuracies are computed as the correlation of the EEG modulations predicted by the fMRI and the real modulations of each run convolved with a canonical HRF response and orthogonalized with respect to the stimulation protocol. Mean accuracies were computed by transforming individual  $r$ -values into  $z$ -values and re-transforming the resulting mean  $z$  into an  $r$ -value. Statistical significance on the single runs and on the mean was computed using  $n$  equal to the minimum number of points present in each run ( $n = 450$ ).

		All EEG Info		
		Run1	Run2	Run3
S1	N1	0.10 $p = 0.01$	0.04 $p = 0.16$	0.08 $p = 0.04$
	P2	0.13 $p = 0.002$	0.06 $p = 0.09$	0.17 $p = 8.5 \cdot 10^{-5}$
	Sust. Neg.	0.16 $p = 1.7 \cdot 10^{-4}$	0.09 $p = 0.01$	0.12 $p = 0.003$
S2	N1	0.09 $p = 0.02$	0.01 $p = 0.3$	0.07 $p = 0.05$
	P2	0.04 $p = 0.16$	0.04 $p = 0.16$	0.03 $p = 0.2$
	Sust. Neg.	0.07 $p = 0.05$	0.07 $p = 0.05$	0.03 $p = 0.2$
S3	N1	0.003 $p = 0.4$	0.14 $p = 7.6 \cdot 10^{-4}$	0.03 $p = 0.2$
	P2	0.08 $p = 0.04$	0.12 $p = 0.003$	0.05 $p = 0.11$
	Sust. Neg.	0.17 $p = 8.5 \cdot 10^{-5}$	0.13 $p = 0.002$	0.14 $p = 7.6 \cdot 10^{-4}$
Mean	N1	0.06 $p = 0.07$		
	P2	0.08 $p = 0.03$		
	Sust. Neg.	0.11 $p = 0.007$		

#### Discussion

Combining fMRI with EEG allows studying brain dynamics with high temporal and spatial resolution. In particular, the use of EEG modulations (event related responses and oscillations), on a trial-by-trial basis has been shown to increase the sensitivity of fMRI analysis in a variety of tasks. This technique belongs to the broader spectrum of multi-modal integration methods that are referred to as “integration through prediction” (Kilner et al., 2005; Ostwald et al., 2010). Such integration methods have been implemented by means of the correlation between single brain locations (univariate analysis) or global patterns (multivariate analysis) and the EEG signal (represented by a single channel or a weighted average of the channels).

Here we have introduced the use of multivariate regression and predictive modelling to the analysis of trial-by-trial coupling in simultaneous EEG-fMRI measurements. Beyond the simple correlation between the two signals, we have shown that it is possible to predict - in a new data set - the event related (ERP and ERSF) modulations of electrical activity as measured by surface EEG from the multivariate patterns measured in fMRI.

Among multivariate techniques for fMRI, multivariate regression is an appealing tool for investigating the relevant question of coupling between simultaneous EEG/fMRI responses. Given the massive multivariate nature of the dataset, the use of “regularized” models such as Relevance Vector Machines (RVM) is necessary in order to avoid “overfitting” of the data. These methods expressly aim at optimizing prediction abilities to unseen data sets, by estimating smooth models, which may reduce the fit of the training set. This is in contraposition with classical univariate models that aim at reducing the error of fit in the available data. Maps obtained using RVM show the brain network that (multivariately) is most relevant in generalizing the learned coupling.

To examine the ability of multivariate regression in predicting trial-by-trial modulations of the EEG signal we have used data from a



**Table 3**

Comparison between EEG informed learning with original and orthogonalized fMRI time series and non-EEG informed training. Accuracies are computed as the correlation of predicted (estimated as the mean over 3 TRs after each stimulus onset) and real single trial EEG modulations. Mean accuracies were computed by transforming individual r-values into z-values and re-transforming the resulting mean z into an r-value. Statistical significance is computed using n equal to the minimum number of trials present in each run (n = 68).

		All EEG Info (non orth. data)			All EEG Info (orth. data)			No EEG Info		
		Run1	Run2	Run3	Run1	Run2	Run3	Run1	Run2	Run3
S1	N1	0.23	0.21	0.24	0.16	0.21	0.19	0.02	0.14	0.004
		$p = 0.02$	$p = 0.03$	$p = 0.02$	$p = 0.07$	$p = 0.03$	$p = 0.05$	$p = 0.4$	$p = 0.1$	$p = 0.5$
	P2	0.17	0.12	0.13	0.24	0.11	0.12	0.01	0.05	0.11
		$p = 0.07$	$p = 0.14$	$p = 0.12$	$p = 0.02$	$p = 0.16$	$p = 0.14$	$p = 0.4$	$p = 0.3$	$p = 0.16$
	Sust. Neg.	0.30	0.11	0.14	0.22	0.17	0.09	0.11	0.09	0.04
		$p = 0.005$	$p = 0.16$	$p = 0.1$	$p = 0.02$	$p = 0.07$	$p = 0.2$	$p = 0.16$	$p = 0.2$	$p = 0.4$
S2	N1	0.19	0.15	0.12	0.17	0.16	0.16	0.01	0.05	0.006
		$p = 0.05$	$p = 0.1$	$p = 0.14$	$p = 0.07$	$p = 0.07$	$p = 0.07$	$p = 0.4$	$p = 0.33$	$p = 0.5$
	P2	0.09	0.17	0.05	0.12	0.15	0.14	0.09	0.13	0.21
		$p = 0.2$	$p = 0.07$	$p = 0.33$	$p = 0.14$	$p = 0.1$	$p = 0.1$	$p = 0.2$	$p = 0.12$	$p = 0.03$
	Sust. Neg.	0.17	0.19	0.33	0.22	0.12	0.27	0.01	0.09	0.09
		$p = 0.07$	$p = 0.05$	$p = 0.001$	$p = 0.02$	$p = 0.14$	$p = 0.009$	$p = 0.4$		$p = 0.2$
S3	N1	0.08	0.13	0.18	0.14	0.15	0.19	0.0006	0.01	0.01
		$p = 0.24$	$p = 0.12$	$p = 0.06$	$p = 0.1$	$p = 0.1$	$p = 0.05$	$p = 0.5$	$p = 0.4$	$p = 0.4$
	P2	0.19	0.35	0.09	0.33	0.11	0.09	0.01	0.01	0.0003
		$p = 0.05$	$p = 0.001$	$p = 0.2$	$p = 0.001$	$p = 0.16$	$p = 0.2$	$p = 0.4$	$p = 0.4$	$p = 0.5$
	Sust. Neg.	0.14	0.09	0.24	0.13	0.18	0.18	0.01	0.14	0.07
		$p = 0.1$	$p = 0.2$	$p = 0.02$	$p = 0.13$	$p = 0.06$	$p = 0.06$	$p = 0.4$	$p = 0.1$	$p = 0.2$
Mean	N1	0.17			0.17			0.03		
		$p = 0.07$			$p = 0.07$			$p = 0.3$		
	P2	0.15			0.16			0.07		
		$p = 0.1$			$p = 0.07$			$p = 0.2$		
	Sust. Neg.	0.19			0.18			0.07		
		$p = 0.05$			$p = 0.06$			$p = 0.2$		

visual imagery task. The nature of the task allowed us to analyse ERP responses associated with auditory (cue) and visual (target) processing as well as the induced ERSP modulations of a component previously shown to be associated with the performance of the subjects in the task. We have recently investigated the same task in a study in which fMRI and EEG data were collected separately (de Borst et al., submitted for publication). Here we focused our analysis to those EEG components that replicated the results obtained outside the MR environment (Fig. 2).

Our results show that RVM is able to capture the modulations of the ERP peaks associated with the processing of the auditory cue (N1, P2 and Sustained Negativity). RVM-based predictions lead to a high correlation value between predicted and actual test data. However, because of the filtering effects of the hemodynamic response, such high values are largely determined by the overall variations of the fMRI signal with the auditory stimulus (the on-off type of response to the sound), which makes it difficult evaluating the actual contribution of EEG modulations. Comparing the results obtained with EEG informed training to results obtained from non-EEG informed training (Table 1) supports this conclusion.

In previous trial-by-trial EEG-fMRI coupling studies the specific contribution of EEG information was assessed in two ways: 1) introducing two predictors in the univariate statistical model (one accounting for the overall “evoked” response and the second accounting for the trial-by-trial variability decorrelated from the first) (Eichele et al., 2005); 2) extracting single trial responses from the fMRI time courses and measuring the trial-by-trial coupling with the EEG modulations (Debener et al., 2005).

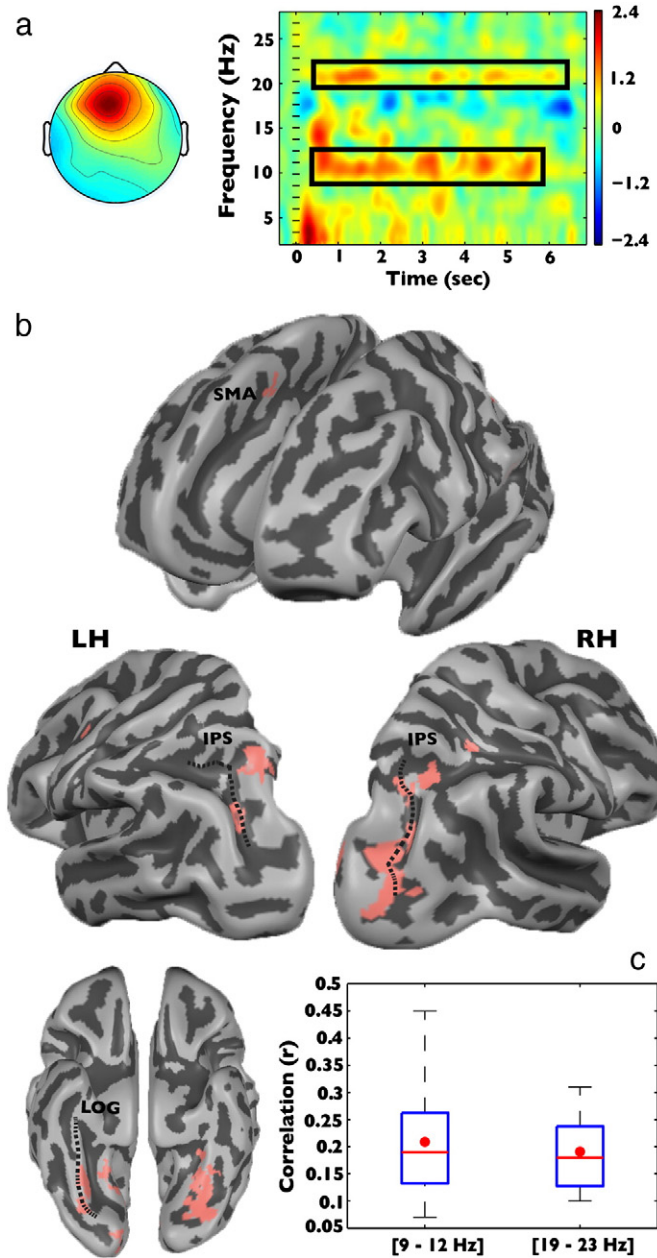
Here we attempted to address this same issue with three different analyses. First, we analyzed the predicted modulations with a statistical model accounting for the overall “evoked” response and the trial-by-trial variability. We compared (two-way ANOVA for repeated measures) the beta values associated to the EEG modulations in predictions obtained from EEG informed training and non-EEG informed training (Fig. 5). We show that predictions obtained from an EEG-informed learning can better explain EEG modulations in

test data than predictions obtained without using EEG information the training. Furthermore the predictions are selective to the specific targeted information, which shows that - even without removing the main response - the RVM training is able to pick up a portion of the trial-by-trial variations of selected peaks.

Methods of multivariate regression are currently limited to “learning” of single predictors and treat them sequentially/separately. Therefore, in our second analysis we decorrelated both the fMRI time series and the EEG modulations with respect to the experimental protocol prior to the training. Despite being lower, accuracies in predicting the test data are still significant (for the P2 and Sustained Negativities, Fig. 6 and Table 2). The maps obtained from this analysis highlight similar regions in the auditory cortex to the maps obtained from learning on the non-orthogonalized data (Figs. 3 and 6). Extension of multivariate regression to “learning” multiple predictors at the same time may lead to further improvements in our results as it would allow fully multivariate modeling of both the overall response and the single-trial variability simultaneously.

In a third analysis we evaluated the accuracies of the predictions obtained by EEG informed (original and orthogonalized data) and non-EEG informed training in terms of trial-by-trial coupling (Table 3). As expected this procedure reduced the correlation values obtained from the original data as it accounted for the effect of the hemodynamic response to the overall protocol. Conversely, the predictions obtained from the orthogonalized data were higher (but not more significant) when computed on the single trials instead of the whole time course. Predictions obtained by non-EEG informed training were not significant and in line with the results obtained in the previous analysis. The variability of the results at single run/subject level indicates that - although promising - this approach needs further improvement and optimization to allow an accurate trial-by-trial prediction.

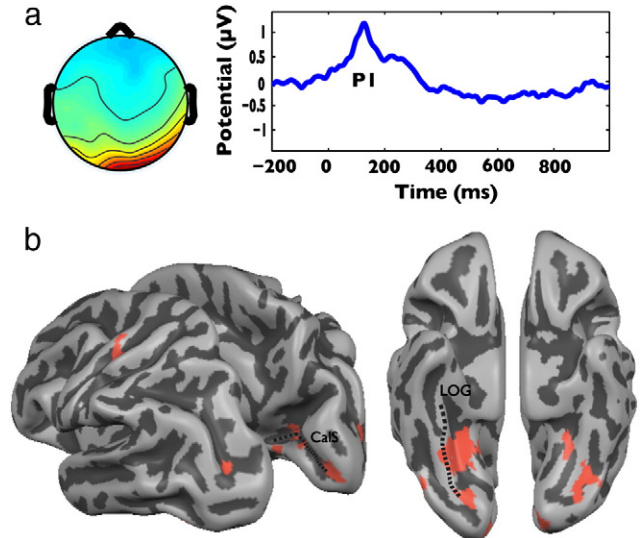
Maps obtained from the RVM-training (Fig. 4) highlight the involvement of a broad network comprising the auditory cortex and superior temporal gyrus (mainly on the right), intra parietal sulcus and supplementary motor area. The maps obtained from the prediction of the three peaks of the auditory ERPs suggest a temporal difference



**Fig. 7.** Results obtained from the prediction of the single trial power modulation of the “imagery” IC. a) Group topography and event related spectral power (ERSP) of the IC representing imagery processing. Time-frequency windows showing a prominent synchronization are highlighted. b) Group map (after cortex based alignment) obtained from the training of the RVM on the two selected time-frequency windows ([9 – 12 Hz; 19 – 23 Hz]) in each subject. Maps are thresholded as in Fig. 3. Major anatomical landmarks are highlighted (Intra Parietal Sulcus, Supplementary Motor Area and Lateral Occipital Gyrus). c) Group accuracy in the prediction of the power modulations of the two selected time-frequency windows. Median (red line) lower and upper quartiles (blue box) and mean (red dot) accuracy across subjects are plotted for each of the time-frequency windows.

between areas of the auditory cortex closer to the Heschl's gyrus and more posterior regions. While physiologically plausible a definitive statistical conclusion on these observations is limited by the small sample of subjects.

The EEG results highlighted the contribution of a frontal EEG component to the process of forming the mental visual image (Fig. 2). In particular, EEG measurements outside the MR scanner (de Borst et al., submitted for publication) showed that oscillations in the low beta and theta ranges of this frontal component are associated with the



**Fig. 8.** Results obtained from the prediction of the prominent ERP peaks of the early visual IC. a) Group topography and event related potential (ERP) of the IC characterized by an early visual event related response (P1 highlighted). b) Group map (after cortex based alignment) obtained from the training of the RVM on the P1 response in each subject. The map is thresholded as in Fig. 3. Major anatomical landmarks are highlighted (Calcarine Sulcus and Lateral Occipital Gyrus).

behavioral performances of the subjects in the imagery task. When predicting such ERSF modulations from simultaneously acquired fMRI data (Fig. 7), the accuracy was lower (but significant) if compared to the ERP-based predictions. It has to be noted that as the predictors for the imagery period were obtained from a temporal window encompassing almost the whole trial length, these predictors were less affected by the problem of a high degree of correlation to the experimental protocol. This may be a contributing factor to the lower prediction accuracies. Furthermore, degraded quality of the single trial oscillations and a non adequate selection for the HRF used for the convolution (fixed to same parameters used for the ERP modulations) may be other relevant factors. Nonetheless the maps obtained from the training of the RVM (Fig. 7) point towards the hypothesized coupling between the frontal EEG component and the supplementary motor area (de Borst et al., submitted for publication) within a network of areas including the bilateral intra parietal sulcus and ventral visual areas.

To show that our method can produce physiologically plausible results we also predicted the ERP responses associated with the presentation of the visual target (Fig. 8). While highly significant these results are still contaminated by the overall hemodynamic response. Similar analyses as the one carried on for the auditory ERPs are needed to evaluate the contribution to the prediction of the single trial ERP modulations.

## Conclusions

We have shown that multivariate regression is a valuable tool to study the relation between single trial modulations as measured by EEG and fMRI. While previously introduced methods for measuring trial-by-trial coupling are based on the simple correlation between simultaneously acquired signals, we suggest that the use of machine learning approaches as RVM allows evaluating the real capability of predicting one modality (EEG) from another modality (fMRI).

## References

- Allen, P.J., Josephs, O., Turner, R., 2000. A method for removing imaging artifact from continuous EEG recorded during functional MRI. *Neuroimage* 12 (2), 230–239.
- Bishop, C.M., 2006. *Pattern Recognition and Machine Learning*. Springer.
- Dale, A.M., Halgren, E., 2001. Spatiotemporal mapping of brain activity by integration of multiple imaging modalities. *Current opinion in neurobiology* 11, 202–208.

- Debener, S., Ullsperger, M., Siegel, M., Fiehler, K., von Cramon, D.Y., Engel, A.K., 2005. Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. *Journal of Neuroscience* 25 (50), 11730–11737.
- Debener, S., Ullsperger, M., Siegel, M., Engel, A.K., 2006. Single-Trial EEG-fMRI reveals the dynamics of cognitive function. *Trends in Cognitive Sciences* 10 (12), 558–563.
- Debener, S., Strobel, A., Sorger, B., Peters, J., Kranczioch, C., Engel, A.K., Goebel, R., 2007. Improved quality of auditory event-related potentials recorded simultaneously with 3-T fMRI: removal of the ballistocardiogram artefact. *Neuroimage* 34 (2), 587–597.
- de Borst, A.W., Sack, A.T., Jansma, B.M., Esposito, F., De Martino, F., Valente, G., Roebroeck, A., di Salle, F., Goebel, R., Formisano, E. Working title: Spatial topography (fMRI) and spectro-temporal brain dynamics (EEG) during mental imagery of scenes. submitted for publication.
- Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134 (1), 9–21.
- De Martino, F., Valente, G., de Borst, A.W., Esposito, F., Roebroeck, A., Goebel, R., Formisano, E., 2010. Multimodal imaging: an evaluation of univariate and multivariate methods for simultaneous EEG/fMRI. *Magnetic Resonance Imaging* 28 (8), 1104–1112.
- Eichele, T., Specht, K., Moosmann, M., Jongsma, M.L.A., Quiroga, R.Q., Nordby, H., Hugdahl, K., 2005. Assessing the spatiotemporal evolution of neuronal activation with single-trial event-related potentials and functional MRI. *Proc. Natl. Acad. Sci.* 102 (49), 17798–17803.
- Eichele, T., Calhoun, V.D., Moosmann, M., Specht, K., Jongsma, M.L.A., Quiroga, R.Q., Nordby, H., Hugdahl, K., 2008. Unmixing concurrent EEG-fMRI with parallel independent component analysis. *Int. J. Psychophysiol.* 67 (3), 222–234.
- Formisano, E., Linden, D.E.J., Di Salle, F., Trojano, L., Esposito, F., Sack, A.T., Grossi, D., Zanella, F.E., Goebel, R., 2002. Tracking the Mind's Image in the Brain I. *Neuron* 35 (1), 185–194.
- Formisano, E., De Martino, F., Valente, G., 2008. Multivariate analysis of fMRI time series: classification and regression of brain responses using Machine Learning. *Magnetic Resonance Imaging* 26 (7), 921–934.
- Friston, K.J., Fletcher, P., Josephs, O., Holmes, A., Rugg, M.D., Turner, R., 1998. Event-Related fMRI: Characterizing Differential Responses. *Neuroimage* 7 (1), 30–40.
- Friston, K., Chu, C., Mourão-Miranda, J., Hulme, O., Rees, G., Penny, W., Ashburner, J., 2008. Bayesian decoding of brain images. *Neuroimage* 39 (1), 181–205.
- Goebel, R., Esposito, F., Formisano, E., 2006. Analysis of functional image analysis contest (FIAC) data with brainvoyager QX: From single-subject to cortically aligned group general linear model analysis and self-organizing group independent component analysis. *Human Brain Mapping* 25 (5), 392–401.
- Goebel, R., Esposito, F., in press. The Added Value of EEG-fMRI in Imaging Neuroscience. Chapter in: Mulert, C., Lemieux, L. *EEG-fMRI Physiology, Technique and Application*. Springer. doi:10.1007/978-3-540-87919-0.
- Goldman, R.I., Wei, C., Philiastides, M.G., Gerson, A.D., Friedman, D., Brown, T.R., Sajda, P., 2009. Single-trial discrimination for integrating simultaneous EEG and fMRI: identifying cortical areas contributing to trial-to-trial variability in the auditory oddball task. *Neuroimage* 47, 136–147.
- Kilner, J.M., Mattout, J., Henson, R., 2005. Hemodynamic correlates of EEG: A heuristic. *Neuroimage* 28, 280–286.
- Logothetis, N.K., 2008. What we can do and what we cannot do with fMRI. *Nature* 453, 869–878.
- Makeig, S., Jung, T.P., Bell, A.J., Ghahremani, D., Sejnowski, T.J., 1997. Blind separation of auditory event-related brain responses into independent components. *Proc. Natl. Acad. Sci.* 94 (20), 10979–10984.
- Mayhew, S.D., Dirckx, S.G., Niazy, R.K., Iannetti, G.D., Wise, R.G., 2010. EEG signatures of auditory activity correlate with simultaneously recorded fMRI responses in humans. *Neuroimage* 49, 849–864.
- Mulert, C., Seifert, C., Leicht, G., Kirsch, V., Ertl, M., Karch, S., Moosmann, M., Lutz, J., Moller, H.-J., Hegerl, U., Pogarell, O., Jager, L., 2008. Single-trial coupling of EEG and fMRI reveals the involvement of early anterior cingulate cortex in effortful decision making. *Neuroimage* 42, 158–168.
- Niazy, R.K., Beckmann, C.F., Iannetti, G.D., Brady, J.M., Smith, S.M., 2005. Removal of fMRI environment artefacts from EEG data using optimal basis sets. *Neuroimage* 28 (3), 720–737.
- Ostwald, D., Porcaro, C., Bagshaw, A.P., 2010. An information theoretic approach to EEG-fMRI integration of visually evoked responses. *Neuroimage* 49 (1), 498–516.
- Ritter, P., Villringer, A., 2006. Simultaneous EEG-fMRI. *Neuroscience and behavioral reviews* 30, 823–838.
- Scheeringa, R., Petersson, K.M., Oostenweld, R., Norris, D.G., 2009. Trial-by-trial coupling between EEG and BOLD identifies networks related to alpha and theta EEG power increases during working memory maintenance. *Neuroimage* 44, 1224–1238.
- Sirotin, Y.B., Aniruddha, D., 2009. Anticipatory hemodynamic signals in sensory cortex not predicted by local neuronal activity. *Nature* 457, 475–479.
- Tipping, M.E., 2001. Sparse Bayesian learning and the relevance vector machine. *Journal of Machine Learning Research* 1, 211–244.
- Valdes-Sosa, P.A., Sanchez-Bornot, J.M., Sotero, J.C., Iturria-Medina, Y., Aleman-Gomez, Y., Bosh-Bayard, J., Carbonell, F., Ozaki, T., 2009. Model Driven EEG/fMRI Fusion of Brain Images. *Human Brain Mapping* 30, 2701–2721.
- Valente, G., De Martino, F., Esposito, F., Goebel, R., Formisano, E., submitted for publication. Predicting subject-driven actions and sensory experience in a virtual world with Relevance Vector Machine Regression of fMRI data. *Neuroimage*. doi:10.1016/j.neuroimage.2010.09.062.
- Whittingstall, K., Logothetis, N., 2009. Frequency-band coupling in surface EEG reflects spiking activity in monkey visual cortex. *Neuron* 64, 281–289.