NeuroImage xxx (2017) 1–12

Contents lists available at ScienceDirect

NeuroImage

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

A multimodal encoding model applied to imaging decision-related neural cascades in the human brain

Jordan Muraskin [a], Truman R. Brown [b], Jennifer M. Walz [c], Tao Tu [a], Bryan Conroy [d], Robin I. Goldman [e], Paul Sajda [a,*]

[a] Department of Biomedical Engineering, Columbia University, New York, NY 10027, USA
[b] Center for Biomedical Imaging, Medical University of South Carolina, Charleston, SC 29425, USA
[c] Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA 94305, USA
[d] Philips Research, Cambridge, MA 02141, USA
[e] Center for Healthy Minds, University of Wisconsin-Madison, Madison, WI 53705, USA

ARTICLE INFO

Keywords:
EEG
fMRI
Simultaneous acquisition
Decision-making
Confidence

ABSTRACT

Perception and cognition in the brain are naturally characterized as spatiotemporal processes. Decision-making, for example, depends on coordinated patterns of neural activity cascading across the brain, running in time from stimulus to response and in space from primary sensory regions to the frontal lobe. Measuring this cascade is key to developing an understanding of brain function. Here we report on a novel methodology that employs multimodal imaging for inferring this cascade in humans at unprecedented spatiotemporal resolution. Specifically, we develop an encoding model to link simultaneously measured electroencephalography (EEG) and functional magnetic resonance imaging (fMRI) signals to infer high-resolution spatiotemporal brain dynamics during a perceptual decision. After demonstrating replication of results from the literature, we report previously unobserved sequential reactivation of a substantial fraction of the pre-response network whose magnitude correlates with a proxy for decision confidence. Our encoding model, which temporally tags BOLD activations using time localized EEG variability, identifies a coordinated and spatially distributed neural cascade that is associated with a perceptual decision. In general the methodology illuminates complex brain dynamics that would otherwise be unobservable using fMRI or EEG acquired separately.

1. Introduction

The detailed spatiotemporal brain dynamics that underlie human perception and cognition are difficult to measure. Invasive techniques with sufficient temporal or spatial resolution, such as depth electrodes or cortical arrays used with epilepsy patients, are only feasible in rare cases and, in addition, do not capture activity from the entire brain. In comparison, non-invasive measures such as electroencephalography (EEG) and magnetoencephalography (MEG) suffer from poor spatial resolution, and blood oxygen level dependent functional MRI (BOLD fMRI) from poor temporal resolution and indirect coupling to neural activity (Logothetis, 2008). In spite of this, EEG, MEG, and fMRI have been used individually to study functional activity in the human brain, although, by themselves these modalities provide a limited view of the underlying brain dynamics (Alexander et al., 2015).

Recently, methods enabling simultaneous acquisition of EEG and fMRI (EEG/fMRI) have led to varied analytic approaches aimed at integrating the electrophysiological and hemodynamic information acquired through the simultaneous measurements. Such approaches offer the potential to provide a more comprehensive picture of global brain dynamics, and will likely offer new insights into how the brain makes rapid decisions (Huster et al., 2012; Jorge et al., 2014). Some of the techniques that have been proposed for combining multi-modal brain signals have separately analyzed the EEG and fMRI data and subsequently juxtaposed the results (Plichta et al., 2013; Yuan et al., 2010), while others attempt for a truly integrated approach in order to fully exploit the joint information contained in the data sets (Dahne et al., 2015). In general, simultaneous EEG/fMRI and the associated analysis techniques have been used to identify neuronal sources of EEG trial-to-trial variability, linking them to cognitive processes such as attention (Warbrick et al., 2013) and inhibitory control (Baumeister et al., 2014).

Many previous studies have used known EEG markers (P1, N2, N170,

* Corresponding author.
E-mail addresses: jordan.muraskin@gmail.com (J. Muraskin), psajda@columbia.edu (P. Sajda).

http://dx.doi.org/10.1016/j.neuroimage.2017.06.059
Received 29 November 2016; Received in revised form 20 June 2017; Accepted 22 June 2017
Available online xxxx
1053-8119/

P300, α-rhythm) or data-driven approaches such as Independent Component Analysis (ICA) to combine EEG with fMRI data (Huster et al., 2012; Baumeister et al., 2014; De Martino et al., 2010; Jann et al., 2009; Jaspers-Fayer et al., 2012; Mayhew et al., 2013; Nguyen and Cunnington, 2014; Novitskiy et al., 2011; Omata et al., 2013; Warbrick et al., 2009). One promising approach has been to use supervised machine-learning techniques (e.g. classifiers) to find relevant projections of the EEG data, where single-trial variability of the electrophysiological response along these projections can be correlated in the fMRI space. Goldman et al. (2009), Walz et al. (2014), Muraskin et al. (2017), and Fouragnan et al. (2015) have demonstrated this technique on visual and auditory paradigms. This methodology has been shown to localize cortical regions that modulate with the task while preserving the temporal progression of task-relevant neural activity. However, these techniques treat neighboring temporal regions independently when fusing the EEG with fMRI. The approach we present in this paper considers the entire temporal progression of task-specific discriminant activity in the EEG and models how this temporal progression is encoded in the fMRI BOLD data. Using the full temporal progression allows us to improve temporal resolution and increase the quality of the fit in the encoding model and the accuracy in the subsequent decoding.

Encoding models have become an important machine learning tool for analysis of fMRI (Naselaris et al., 2011), in that they provide a generative model of the BOLD data (Fig. 1A). In most cases encoding models have been used to learn brain activity that encodes or represents features of a stimulus, such as visual orientation energy in an image/video (Hansen et al., 2007; Kay et al., 2008; Nishimoto et al., 2011), acoustic spectral power in sound/speech (Silbert et al., 2014), visual imagery during sleep (Horikawa et al., 2013) and even high-level concepts and semantics (Huth et al., 2016). In the method presented here, we employ an encoding model methodology to directly relate the simultaneously collected data from the two neuroimaging modalities. Instead of learning features derived from characteristics of the stimulus, they are derived from the trial-to-trial variability of the EEG components learned via a classifier (Fig. 1B). Specifically, we learn an encoding for the spatially precise fMRI data from the temporally precise trial-to-trial variability of the EEG – i.e. we learn a mapping that explains the BOLD data given the EEG variability across the time course of each trial. In contrast to previous simultaneous EEG/fMRI techniques that treat the EEG variability at different time points in the trial as independent, we are able to exploit the temporal structure embedded in the EEG by modeling the EEG variability as a time series to accurately predict time course activations in the spatial domain.

As an example of this method, we apply it to simultaneously acquired EEG/fMRI data collected while subjects performed a perceptual decision-making task. The EEG variability we exploit is derived from a classifier trained to predict the level of stimulus evidence on a trial-to-trial basis from the multichannel EEG data. This approach leverages the fact that in perceptual decision-making tasks, the level of stimulus evidence, as measured via EEG, persists across the trial (Banko et al., 2011; Philiastides and Sajda, 2006). We can discriminate this information in a time-localized way in the EEG and then temporally "tag" specific cortical areas by their trial-to-trial variability as they become involved and uninvolved in the decision process. We then proceed to discuss the

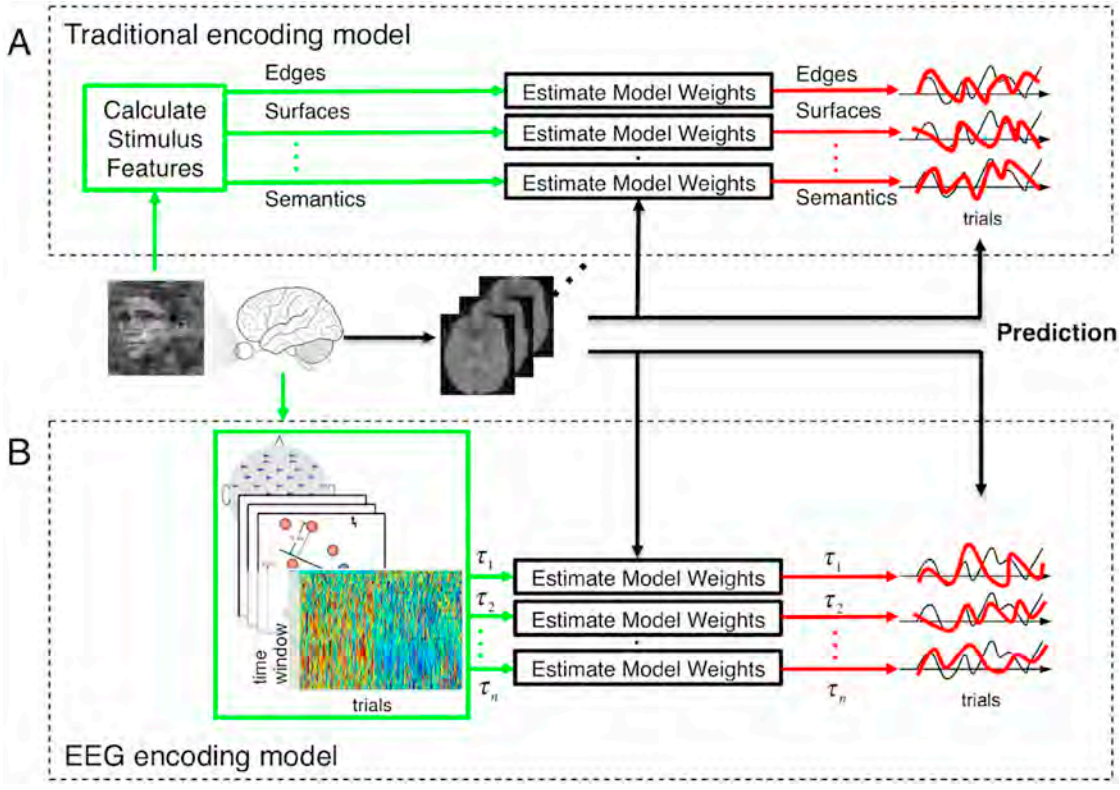

Fig. 1. Encoding models based on stimulus derived features versus EEG variability. **A**, A traditional encoding model used in fMRI analysis extracts a set of features from the stimulus that are potentially representative of low-level structure and high-level semantics (green box). Weights are learned to model how these stimulus features are encoded in the fMRI BOLD signal. The resulting encoding model is used to make predictions based on how well different voxels predict the features from novel stimuli. For example, one can create maps of the brain that are labeled based on the stimulus features that each voxel represents. **B**, The same encoding model concept applied to EEG variability (EEG encoding model). Instead of features being estimated from the stimulus, they are derived from EEG component trial-to-trial variability (as in Fig. 3a) with each temporal window representing a different feature (green box). In traditional fMRI encoding models, the stimulus features or concepts are derived through non-linear filters (Cukur et al., 2013; Hansen et al., 2007; Kay et al., 2008; Naselaris et al., 2011; Nishimoto et al., 2011; Stansbury et al., 2013), however in our model, the non-linearity step is produced by the subjects brain response to each stimuli which we are able to record with EEG. Weights are learned so as to model how the EEG variability at a given time window is encoded in the fMRI BOLD. As in the traditional encoding model, predictions on novel stimuli can be done to test the model and results can be used to construct a map, in this case a map of the brain that shows the timing of the EEG component variability that each voxels represents.

significance of the results of the new method relative to what can be inferred from non-simultaneously acquired EEG and fMRI.

2. Methods

2.1. Subjects

21 subjects (12 male, 9 female; age range 20–35 years) participated in the study. The Columbia University Institutional Review Board (IRB) approved all experiments and informed consent was obtained before the start of each experiment. All subjects had normal or corrected-to-normal vision.

2.2. Stimuli

We used a set of 30 face (from the Max Planck Institute face database), 30 car, and 30 house (obtained from the web) gray scale images (image size 512 × 512 pixels, 8 bits/pixel). The images all had identical magnitude spectra (average magnitude spectrum of all images in the database) and their corresponding phase spectra were manipulated using the weighted mean phase (WMP) (Dakin et al., 2002) technique to generate a set of images characterized by their percent phase coherence. The stimulus evidence (high or low) for each trial was systematically varied by modifying the salience of the image via randomization of image phase at either 35% (low) or 50% (high) coherence.

2.3. Experimental task

The stimuli were used in an event-related three choice reaction time task. On each trial an image of either a face, car, or house was presented for 50 ms and subjects were instructed to respond with the category of the image by pressing one of three buttons on an MR compatible button controller. Stimuli were presented to subjects using E-Prime software (Psychology Software Tools) and a VisuaStim Digital System (Resonance Technology) with 600 × 800 goggle display. Images subtended 11°x 8° of visual angle. Over four runs, a total of 720 trials were acquired (240 of each category with 120 high coherence trials) with a random inter-trial interval (ITI) sampled uniformly between 2 and 4 s. Each run lasted for 560 s.

2.4. fMRI acquisition

Blood-oxygenation-level-dependent (BOLD) T2*-weighted functional images were acquired on a 3T Philips Achieva scanner using a gradient-echo echo-planar imaging (EPI) pulse sequence with the following parameters: Repetition time (TR) 2000 ms, echo time (TE) 25 ms, flip angle 90, slice thickness 3 mm, interslice gap 1 mm, in-plane resolution 3 × 3 mm, 27 slices per volume, 280 volumes. For all of the participants, we also acquired a standard T1-weighted structural MRI scan (SPGR, resolution 1 × 1 × 1 mm).

2.5. EEG acquisition

We simultaneously and continuously recorded EEG using a custom-built MR-compatible EEG system (Sajda et al., 2007, 2010), with differential amplifiers and bipolar EEG montage. The caps were configured with 36 Ag/AgCl electrodes including left and right mastoids, arranged as 43 bipolar pairs. Further details are given in the Supplementary Material.

2.6. Functional image pre-processing

Image preprocessing was performed with FSL (www.fmrib.ox.ac.uk/fsl/). Functional images were spatially realigned to the middle image in the times series (motion-correction), corrected for slice time acquisition, spatially smoothed with a 6 mm FWHM Gaussian kernel, and high-pass filtered (100 s). The structural images were segmented (into grey matter, white matter and cerebrospinal fluid), bias corrected and spatially normalized to the MNI template using FAST (Zhang et al., 2001). Functional images were registered into MNI space using boundary based registration (BBR) (Greve and Fischl, 2009).

2.7. EEG data preprocessing

We performed standard EEG preprocessing offline using MATLAB (MathWorks) with the following digital Butterworth filters: 0.5 Hz high-pass to remove direct current drift, 60 and 120 Hz notches to remove electrical line noise and its first harmonic, and 100 Hz low-pass to remove high-frequency artifacts not associated with neurophysiological processes. These filters were applied together in the form of a zero-phase finite impulse response filter to avoid distortions caused by phase delays. We extracted stimulus-locked 1500 ms epochs (−500 ms:1000 ms) and subtracted the mean baseline (−200 ms:0 ms) from the rest of the epoch. Through visual inspection, we discarded trials containing motion and/or blink artifacts, evidenced by sudden high-amplitude deflections.

2.8. Estimating EEG components and trial-to-trial variability

We used logistic regression to associate each trial with the level of stimulus evidence represented in the EEG. We considered high stimulus and low stimulus evidence trials irrespective of behavioral accuracy. Regularized logistic regression (r-LR) was used as a classifier to find an optimal projection for discriminating between high and low stimulus evidence trials over a specific temporal window. A sweep of the regularization parameters was implemented using FaSTGLZ (Conroy et al., 2013). This approach has been previously applied to identify neural components underlying rapid perceptual decision-making (Goldman et al., 2009; Walz et al., 2014; Muraskin et al., 2015; Parra et al., 2005; Philiastides et al., 2014; Sherwin et al., 2012; Sherwin and Sajda, 2013; Walz et al., 2013).

Specifically, we defined 50 ms duration training windows centered at time, τ, ranging from stimulus onset to 800 ms following the stimulus in 25 ms steps. We used r-LR to estimate a spatial weighting, on all the N EEG channels, that discriminated high stimulus evidence from low stimulus evidence trials. This identifies an $N\times1$ vector, w_τ, which defines a hyperplane that maximally discriminates between each class (e.g., high vs. low stimulus evidence trials) in the N dimensional space of electrodes. Using this w_τ we calculated the distance each trial is from the decision boundary

$$y_\tau = w_\tau^T E_\tau \,. \tag{1}$$

In Eqn. (1), E_τ is an $N\times p$ vector (N sensors per time window, τ, by p trials). We varied the center of the window (τ) across the trial in 25 ms time-steps. For each subject, this produced a matrix Y where the rows corresponded to trials and the columns to training windows, i.e. Y incorporates the time-window localized trial-to-trial variability in the EEG that we subsequently use to tag the BOLD data. We quantified the performance of the linear discriminator by the area under the receiver operator characteristic (ROC) curve, referred to here as AUC, using a leave-one-out procedure. We used the ROC AUC metric to characterize the discrimination performance at each τ.

2.9. Traditional fMRI analysis

We first ran a traditional general linear model (GLM) fMRI analysis in FSL, using event-related (high and low stimulus evidence) and response time (RT) variability regressors. The two event-related regressors were unit amplitude boxcar functions matching onset and offset of high and low stimulus evidence trials. For each stimulus evidence level, we modeled out the response time variability by constructing regressors with each trials boxcar amplitude matching the z-scored RT within each stimulus evidence level. The onset and offset of the boxcar function of the

RT variability regressor matches that of the corresponding stimulus regressor. Activated regions that passed a family-wise error (FWE) (Nichols and Hayasaka, 2003) corrected cluster threshold of $p < 0.01$ at a z-score threshold of 2.57 were considered significant. Further details can be found in the Supplemental Material.

2.10. fMRI deconvolution

Associating fMRI data to each trial is challenging for two main reasons: (a) the temporal dynamics of the hemodynamic response function (HRF) evolve over a longer time-scale than the mean ITI of the event-related design, resulting in overlapping responses between adjacent trials; and (b) the ITI is random for each trial so that the fMRI data was not acquired at a common lag relative to stimulus onset. To overcome these issues, we employed the 'least squares - separate' (LS-S) deconvolution (Mumford et al., 2012) method to estimate the voxel activations for each trial. We extracted 58697 voxels from a common gray matter group mask at 3 × 3 × 3 mm spatial resolution that excluded white matter and CSF and assembled the resulting voxel activations into rows of the data matrix F.

2.11. Single subject encoding model

All encoding model analyses were performed in MATLAB. To relate the EEG data with the fMRI, we devised a subject-wise spatiotemporal decomposition using singular value decomposition (SVD). Let F be an $m \times p$ matrix denoting m-voxels and p-trials that is the deconvolved high and low stimulus evidence fMRI data for each trial. Let Y be the $r \times p$ matrix denoting r-windows (33) EEG time windows and response time (RT) and p-trials. Specifically for each trial, the first row of Y represents the RTs while subsequent rows are the y values at each EEG window time. Let W be an $m \times r$ matrix that is the multivariate weights on Y that solve for F with some error ε

$$F = WY + \varepsilon \,. \tag{2}$$

Normally, if we solve for W using the least squares approach, we obtain:

$$W = (FY^T)(YY^T)^{-1} \,. \tag{3}$$

However, each time point might be highly correlated with its neighbors, which reduces the stability of the least-squares regression. We use SVD to reduce the feature space and improve our estimation of W (the weights on each window).

We include a regressor for RT since it is a strong predictor for evidence type. Higher stimulus evidence leads to faster decisions than low evidence trials (see Supplementary Material Fig. S1). The faster RTs will manifest itself in EEG variability. We wanted to find neural correlates of the decision process independent of the response time/motor execution. For that reason, we included the RT variability in the SVD to remove any variability in the EEG that correlates with motor execution. If RT was not added to the SVD, the encoding model weights might be contaminated with differences in motor execution processes instead of the decision process (see Walz et al., 2013, Fig. 2).

For a leave-one-out cross validation, we hold out a single trial from the EEG Y matrix and the corresponding volume from the fMRI data F and train on the remaining trials. We repeated this for all trials. For each leave one out fold we calculated,

$$Y^{Train} = U\Sigma V^T \,, \tag{4}$$

where U is an $r \times r$ orthonormal matrix, Σ is a $r \times p$ diagonal matrix and V is a $p \times p$ orthonormal matrix. After SVD on Y^{Train}, we reduced the feature dimensions on Y^{Train} to retain 75% of the variance by only keeping v components. To do this, we selected the first v rows of Σ and zeroed the other rows. This results in $\tilde{\Sigma}$ which is our reduced subspace matrix. We recalculate our least squares solution where we have replaced Y by its reduced form $U\tilde{\Sigma}V^T$ in equation (3):

$$\widehat{W} = \left(F^{Train} V \tilde{\Sigma}^T\right)(\Sigma\Sigma^T)^{-1} U^T \,. \tag{5}$$

For each leave-one-out fold, we first calculated the SVD of the training set. We then calculated the number of components to keep and then solve for $\widehat{W}$, the weight estimate per fold. To test, we applied the weights to the left-out test data Y^{Test} to estimate the encoded fMRI data $\widehat{F}$ for the encoding part:

$$\widehat{F} = \widehat{W} Y^{Test}. \tag{6}$$

We constructed the $m \times r$ weight matrix $\overline{W}$ by taking the average of all the trained $\widehat{W}$ matrices. For the decoding model we use the left out test data F^{Test} and estimate the EEG variability $\widehat{Y}$:

$$\widehat{Y} = \widehat{W}^T F^{Test} \left(\widehat{W}^T \widehat{W}\right)^{+} \,. \tag{7}$$

Here, $\widehat{W}^T \widehat{W}$ is not invertible, and instead use the pseudo-inverse (denoted by $+$).

Given $\widehat{F}$, an $m \times p$ matrix with m voxels by p trials, for each voxel j, we calculated the correlation over trials of $\widehat{F}_j$ with F_j^{Test}, resulting in the matrices R^{fMRI} (Pearson Correlation Map) and P^{fMRI} (p-value map of the Pearson Correlation) that are $m \times 1$. The P^{fMRI} was then converted to a z-score map. Thus, we have a map for each subject indicating that subjects encoding model significance at each voxel. To test which time windows were significant in the decoding model, we also calculated, R_τ^{EEG} , the correlation between $\widehat{Y_\tau}$ and Y_τ. For each subject, R_τ^{EEG} indicates how well the decoding model predicts EEG variability, at each time window, from a single fMRI volume.

2.12. Group level spatiotemporal analysis

We evaluated both the encoding model and the resulting decoding model at the group level. First, to identify group level voxels where our encoding model predictions were consistent across subjects, each subject's p-value maps for the leave-one-out correlation (R^{fMRI}) were converted into their respective z-values, then a one-sample t-test testing for the mean difference from 0 was calculated for each voxel and voxel-wise significance was calculated using threshold-free cluster enhancement (TFCE) using a non-parametric randomization procedure implemented in FSL (Smith and Nichols, 2009). TFCE is a multiple comparison correction technique in which raw voxel-wise t-statistics are adjusted based on their t-scores and their neighboring spatial clusters, producing corrected whole brain volumes. It generally allows for better sensitivity than other cluster thresholding methods. In our analysis we used the default TFCE parameters provided by FSL ($E = 2$ and $H = 0.5$). Voxels were considered significant if they passed a false discovery rate (FDR) threshold of $p < 0.01$.

These significant voxels were then used as a mask to temporally localize weight values in $\widehat{W}$ by computing the voxels that were consistent in their direction (positive:high stimulus evidence or negative:low stimulus evidence) and timing (τ window) across subjects. Here, the goal was to develop a statistical parametric map of $\widehat{W}$ that identifies significant results in both space and time, thereby creating a spatiotemporal map of a perceptual decision.

First, a one-sample t-test across the subjects $\widehat{W}$ was computed for all voxels. This generates a statistical parametric map of t-scores in space (voxels) and time (windows). To correct for multiple comparisons, we implemented a spatiotemporal TFCE (stTFCE) in both space (neighboring voxels) and time (neighboring time windows - response time window not

included) and computed significance through a randomization procedure. For each of the 33 temporal windows, we randomly altered the sign of the voxel weights for each subject. We then computed the *t*-scores (across subjects) at each voxel for all time windows. The spatiotemporal *t*-score map was then transformed into the stTFCE map. This procedure was repeated 5000 times to generate the empirical null distribution of the stTFCE map. The final statistical parametric map was calculated by comparing the true stTFCE value with the distribution of permuted values for each voxel. Again, voxels were considered significant if they passed FDR correction at $p < 0.05$ (high stimulus evidence: FDR-Corrected $p < 0.0004$, low stimulus evidence: FDR-Corrected $p < 0.0002$). Note that now our number of multiple comparisons was the number of voxels in the FDR-mask (20256) times the number of time windows (33). We analyzed the response time separately with a standard TFCE randomization procedure implemented in FSL.

We also tested the quality of the resulting decoding model at the group level. To construct group level statistics for the decoding model, we first analyzed the R_{τ}^{EEG} vectors across all subjects to identify time windows when the model was consistent across subjects. The correlation vectors were converted into their z-values using the Fisher-Z transform, and for each time window (τ), were used to compute a z-test across all subjects. These group level results were then corrected for multiple comparisons using $p < 0.05$ FDR.

2.13. Drift diffusion model (DDM) and confidence proxy

To analyze and better interpret our novel finding of a reactivation in the neural cascade (see Results), we used a DDM to estimate a proxy of confidence in the decision for each trial. The DDM models decision-making in two-choice tasks. Here, we treated the decision (correct vs. incorrect) as our two choices. A drift-process accumulates evidence over time until it crosses one of two boundaries (upper or lower) and initiates the corresponding response (Ratcliff and McKoon, 2008). The speed with which the accumulation process approaches one of the two boundaries (boundary represented as parameter a) is called drift-rate (v) and represents the relative evidence for or against a particular response.

We used Hierarchical Bayesian estimation of the Drift-Diffusion Model in Python (HDDM) to calculate the drift rate (v), decision boundary (a) and non-decision time T_{non} for each subject (Wiecki et al., 2013). Specifically, we modeled high and low stimulus evidence response time data separately. This was to ensure our confidence proxies were consistent within trial types. We included the response time and whether the subject got the trial correct. HDDM obtains a sequence of samples (i.e., a Markov chain Monte Carlo; MCMC) from the posterior of each parameter in the DDM. In our model, we generated 5000 samples from the posteriors, the first 1000 (burn-in) samples were discarded, and the remaining samples were thinned by 5%.

Recently, Philiastides et al. (2014) showed that for conditions in which the boundary does not change, a proxy for decision confidence for each trial (i) can be computed by calculating $1/\sqrt{RT_i - T_{non}}$, where T_{non} is a subject dependent non-decision time fit by the model. Thus, after modeling the DDM process, each trial's (i) confidence proxy (CP) for each subject (j) was computed by $CP_{i,j} = 1/\sqrt{RT_i - T_{non,j}}$ and then z-scored across trials where $T_{non,j}$ was varied for high or low stimulus evidence trials, separately, so that CP was a measure of relative trial confidence within difficulty levels.

Using the encoding model weights ($\overline{W}$), we are able to estimate the trial-to-trial activation amplitude at different time windows. To estimate the reactivation amplitudes, we took the mean across time of $Y_{j,i}^{R} = W_{j,PostACC}^{T} F_{j,i}$ for each subject (j) and trial (i). Here, $W_{PostACC}$ is the weight matrix from the encoding model restricted to voxels that were significant in the spatiotemporal group results from the 600–800 ms windows. By applying the weights restricted to these spatial areas and temporal windows to each trial, we are able to estimate a trial-to-trial value for reactivation. Note that because the reactivation is stronger in the more difficult trials, the significant voxels in W are negative and so more negative y's indicate stronger reactivation while more positive y's indicate weaker reactivation.

$Y_{j,i}^{R}$ was quintiled for high and low stimulus evidence and the average confidence proxy was calculated within each quintile. A linear mixed effects model (Bates et al., 2015) was used to test if the slope of confidences across quintile grouping, $Y_{j,i}^{R}$, was significantly different from 0 while including stimulus evidence as a condition. Separate similar analyses with non-replay windows (175–250 ms) and testing for behavioral accuracy were also performed. To test the contribution of each cluster to the correlation with confidence, we implemented feature elimination procedure (Martino et al., 2008; Sherwin et al., 2015), where our features were clusters of significant voxels (> 48 voxels) during the 600–800 ms time window. This procedure removed clusters from the reactivation network before calculating trial-to-trial reactivation similar to the procedure in Sherwin et al. (2015). We then calculated the percent change in slope (reactivation × confidence proxy) when the cluster was removed compared to the total network ((new-original)/original). This procedure ranks cluster importance by sorting which clusters, when removed, had the strongest negative effect on slope height. For example, a cluster that does not contribute to the reactivation × confidence proxy slope when removed would not reduce the slope height, however, a cluster that has a large contribution to the reactivation × confidence proxy slope would reduce the slope height when removed. While such an analysis did not provide insight into the sensitivity of the interactions between clusters, it does provide insight into each cluster on its own.

3. Results

As a demonstration of how our method for encoding time-localized EEG variability in simultaneously acquired BOLD data can be used to infer neural cascades underlying perception and cognition, we collected EEG/fMRI data from subjects as they performed a three choice reaction time task discriminating between faces, cars, and houses (Fig. 2A). Subjects were instructed to discriminate the object class after briefly viewing an image corrupted by varying levels of noise (Fig. 2B) and respond by pressing one of three buttons. Overall, subjects responded with accuracies of 94 ± 5% and 58 ± 12% and with response times of 634 ± 82 ms and 770 ± 99 ms for high and low stimulus evidence trials, respectively (Fig. 2C and D). Subject accuracies and response times across stimulus types (faces, cars, houses) for low stimulus evidence trials were similar; however, for high stimulus evidence trials subject accuracies were higher and response times were shorter for faces than for cars or houses (see Supplementary Material Fig. S1).

3.1. GLM based analysis of BOLD fMRI shows superposition of cortical areas correlated with stimulus evidence

We first used a traditional general linear model (GLM) analysis of the fMRI (see Methods) to demonstrate how a difference in stimulus evidence is reflected in the BOLD data (Fig. 2E, Supplementary Material Table S1). These results become a baseline case we will compare, relative to the additional information that we can obtain using our encoding model for fusing the BOLD and EEG data. Brain regions showing greater BOLD activation to high vs. low stimulus evidence trials included areas associated with early visual perception and the default mode network, such as fusiform gyrus, parahippocampal gyrus, lateral occipital cortex, superior frontal gyrus, and posterior cingulate cortex. Regions with greater BOLD activation to low vs. high stimulus evidence trials included areas in the executive control and difficulty networks, such as dorsal lateral prefrontal cortex, anterior cingulate cortex, intraparietal sulcus, and insula. Important to note is that these GLM results, from simultaneous EEG and fMRI acquisitions, reproduced previous results in the literature for fMRI collected alone (Heekeren et al., 2004) (see Supplementary Material Fig. S2A). The encoding model results we present below are compared to

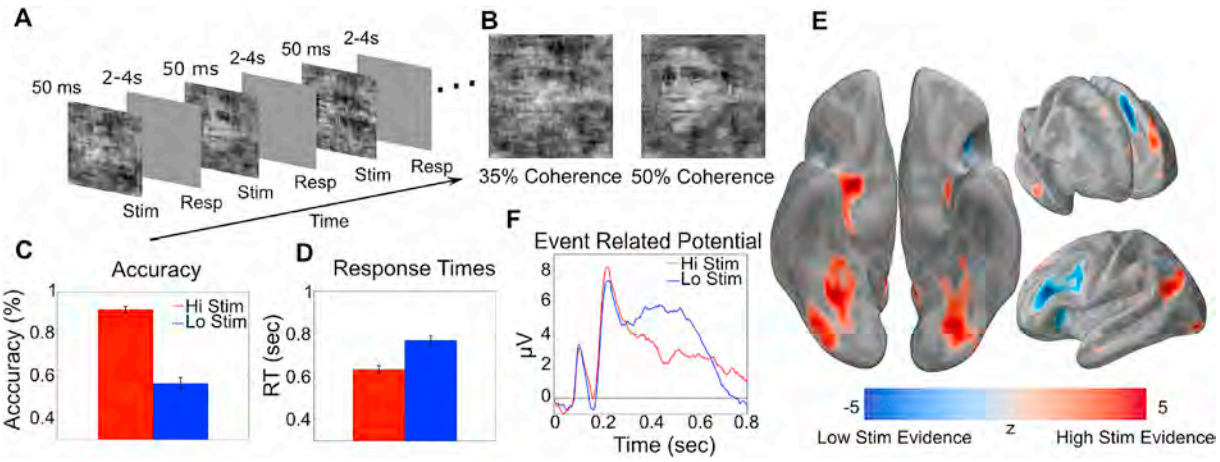

Fig. 2. Paradigm and traditional EEG and fMRI results. A, Three choice reaction time task where stimulus evidence for each category is modulated by varying the percent phase coherence in the images. **B**, Example of face images with high stimulus evidence (high coherence: 50%) and low stimulus evidence (low coherence: 35%). **C**, Behavioral performance shows significant differences, as a function of stimulus evidence, in accuracy ($p < 10^{-12}$, degrees of freedom (dof) = 20, paired t-test) and **D**, response time ($p < 10^{-8}$, dof = 20 paired t-test) across the group. **E**, fMRI analysis showing cortical areas correlated with high (red) vs. low (blue) stimulus evidence across the entire trial ($z > 2.57$ with $p < 0.01$ Family-Wise Error cluster corrected). **F**, Grand average stimulus-locked event related potentials (ERPs) for electrode Pz show that differences in stimulus evidence span the time from stimulus to response.

these results in terms of how the temporal superposition of activations can be separated in time.

3.2. *Extracting temporally localized EEG signatures of stimulus evidence variability*

The traditional fMRI results showed multiple brain regions correlated with the stimulus evidence of the trial; however, this traditional approach does not enable one to infer the relative timing of these fMRI activations. In the trial averaged event related potentials, we see that the difference in stimulus evidence (high vs. low stimulus evidence) spans the trial (see Fig. 2F). To infer timing at a scale of tens of milliseconds, we exploit this difference and use linear classification (Parra et al., 2005; Sajda et al., 2009) of the multichannel EEG to estimate EEG components selective to the trial-to-trial variability in the stimulus evidence. Importantly these components were estimated at 25 ms steps, locked to either the onset of the stimulus or response, so as to encode the BOLD activity at specified time points in the trial. The basic idea is illustrated in Fig. 3,

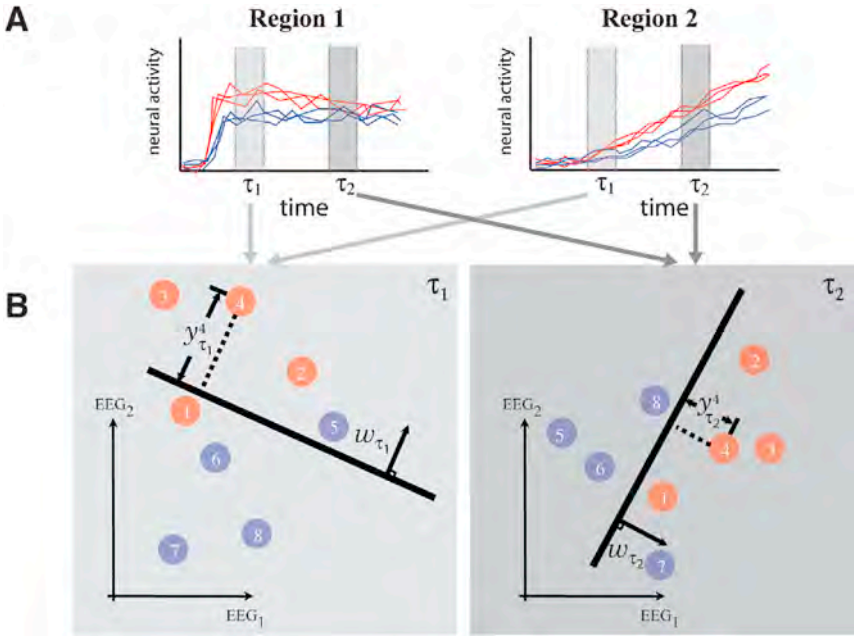

Fig. 3. Temporally precise trial-to-trial EEG variability tags brain regions during decision-making. A, Illustration of how trial-to-trial variability of neural activity in spatially distinct cortical areas can be used to tag brain regions. In this hypothetical example, Region 1 is involved in sensory encoding while Region 2 integrates sensory evidence to form a decision (in NHP literature, Region 1 might represent MT, and Region 2 LIP). Neural activity across the trial is shown for two stimulus types, one with high sensory evidence for the choice (red curves) and one with low sensory evidence (blue curves). Also shown are two temporal windows (τ_1 and τ_2) that represent different times during the trial. **B**, Linear classifiers are trained to separate trials based on the two levels of stimulus evidence at specific temporal windows. Shown are classifiers (parameterized by weight vectors w_1 and w_2) for two temporal windows (τ_1 and τ_2) with respect to two EEG sensors (for simplicity only two dimensions of the full $N = 43$ sensor space are shown. Though the component hyperplane is optimal for the full 43 dimensions, when projected to a line in two dimensions for illustration, it may appear that the separation is sub-optimal). This yields an EEG discriminant component for each temporal window. Variability along these components serves as a unique feature vector for temporally tagging the BOLD data, e.g. variability along an EEG component trained with data from τ_1 tags BOLD voxels with time τ_1 while variability along an EEG component trained with data from τ_2 tags them with τ_2.

where hypothetical neural activity is shown for two different regions that are constituents of the perceptual decision-making network. Averaging over trials would clearly reveal a difference in the mean neural activity between high and low stimulus evidence (as we see in Fig. 2F). However, the two regions contribute differentially to the network, with one region encoding the stimulus evidence (Region 1) and the other integrating it over time (Region 2); both are sensitive to the level of stimulus evidence, though varyingly so at different times in the trial. By taking advantage of this sensitivity to the stimulus evidence, we can learn EEG discriminant components, i.e. spatial filters, that best classify trials at different time windows given the neural data. We used the trial-to-trial variability along these component directions as features to uniquely tag fMRI voxels with the specific time window of the component. This tagging is done by building an encoding model of the features, given the BOLD signal, details of which are described in the following section.

We constructed EEG components by learning linear classifiers at 25 ms steps, starting from stimulus onset to 50 ms past the average low stimulus evidence response time. We chose a time step of 25 ms due to an empirical analysis showing a half width of 50 ms in the temporal autocorrelation of the EEG data, though in principle this methodology allows for temporal resolution up to the EEG sampling rate. Each classifier was associated with a set of discriminant values, which can be represented as a vector y_τ; each element of the vector is the distance of a given trial to the discrimination boundary for the classifier at time step τ (Fig. 3). This distance can be interpreted as a measure of the EEG classifier's estimate of the level of stimulus evidence for that trial (Goldman et al., 2009; Walz

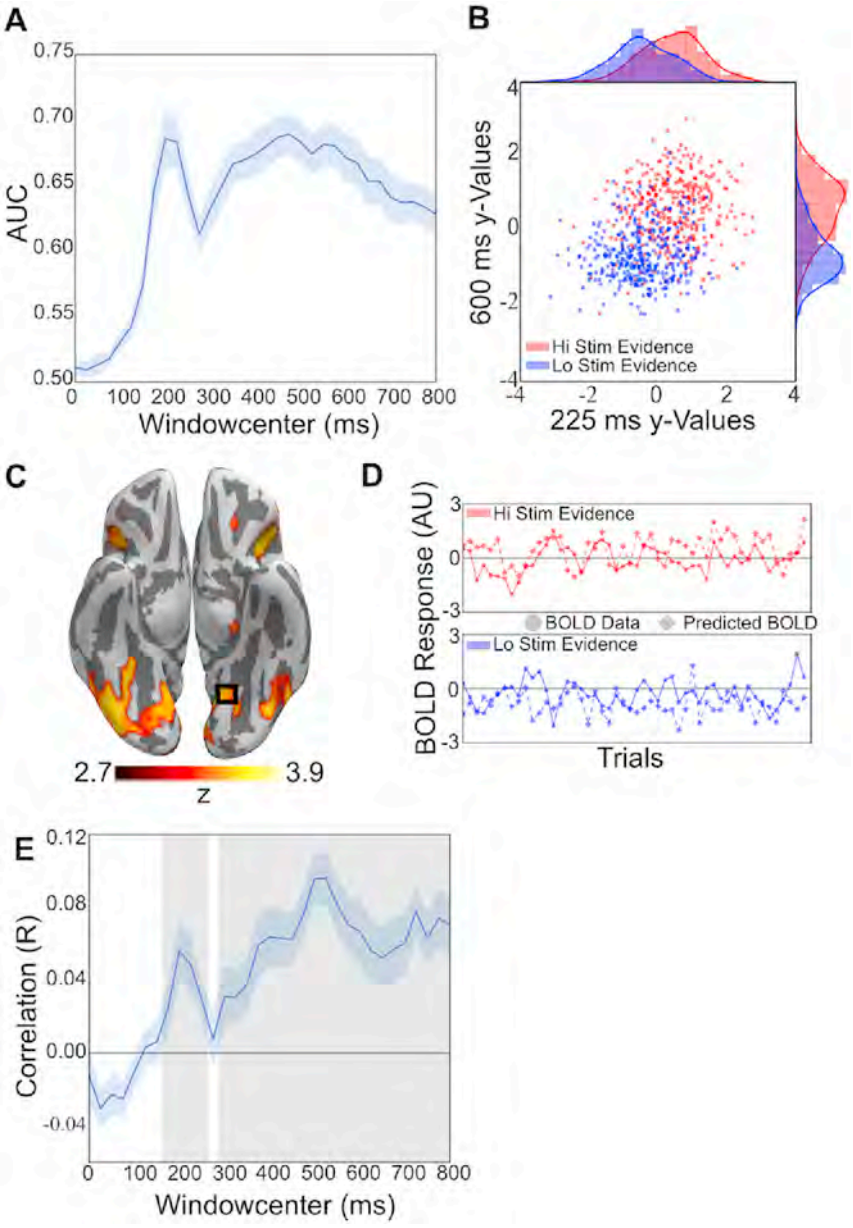

Fig. 4. EEG discrimination and encoding model results. A, Group average area under the receiver operating curve (AUC) for the sliding window logistic regression EEG discrimination analysis, comparing high versus low stimulus evidence trials; standard error across subjects is shown with shading. **B**, A single subject's discriminating y-value distributions for high (red) and low stimulus evidence (blue) trials for two EEG time points (225 ms and 600 ms). **C**, Significant fMRI voxels resulting from the group level analysis for the encoding model ($p < 0.01$ TFCE-False Discovery Rate (FDR) corrected). Activity is seen encompassing early visual processing regions, attention networks, and the task positive network. **D**, A random subset of 100 (50 for each stimulus evidence condition) from 700 total trials of the actual (circle) and predicted (diamond) BOLD responses from the encoding model, for an example subject at a single voxel (MNI X/Y/Z mm: −27/-54/-15, r = 0.206, dof = 709, $p < 10^{-6}$). High and low stimulus evidence trials are shown separately for clarity. **E**, The averaged correlation of the predicted y-values with the true y-values across the trial duration. Blue shading represents the standard error across subjects. Grey shading indicates significant time windows ($p < 0.05$ FDR-corrected).

et al., 2014; Muraskin et al., 2015; Parra et al., 2005; Sherwin et al., 2012; Sajda et al., 2009).

Results of the EEG analysis show discriminating information for stimulus evidence spanning the trial (see Fig. 4A), beginning roughly 175 ms post-stimulus to past the average response times. A dip occurs around 300 ms, indicating stimulus evidence is less discriminative at this time and serves to demarcate early and late cognitive processes. The early process corresponded to the time of the D220 ERP component, which has been shown to modulate with the degree of task difficulty, whether via stimulus noise or task demands (Philiastides and Sajda, 2006). The later and more prolonged component is likely related to more complex cognitive and motor preparatory processes that differ between high and low stimulus evidence trials. Importantly, although the early and late EEG components were both discriminative, we found their trial-to-trial variability to be uncorrelated (Fig. 4B and see Supplementary Material Fig. S3E), indicating that while the discriminating information (level of stimulus evidence) persists across the trial, it couples differently to processes across time.

3.3. Using an encoding model to link fMRI activations with temporally distinct EEG trial-to-trial variability

After extracting the trial-to-trial variability from the EEG discriminant components, feature vectors y_τ are collected across time steps, τ, along with a response time vector to construct a matrix Y. This matrix is the temporally precise representation of the trial-to-trial EEG variability that reflects high vs. low stimulus evidence. An encoding model (see Methods) is then fit to predict the trial-to-trial variability of the BOLD response for each fMRI voxel. The model can be seen as "taging" each voxel with a "time," or set of times, when it encodes the variability in the given EEG discriminant component(s).

The group level result of the encoding model is shown in Fig. 4C (see also Supplementary Material Fig. S2B) where significant voxels are shown in yellow. Fig. 4D shows the trial-to-trial variability of BOLD signal at a specific voxel, comparing it to the variability predicted by the encoding model. Additional validity of the encoding model and single subject results are presented in the Supplementary Material Fig. S4A/B. The encoding model was also evaluated as a decoding model (see Methods) with the BOLD activity used to predict the trial-to-trial variability in the EEG for unseen data–i.e. data on which the encoding model was not trained. Fig. 4E shows these results, expressed as the correlation between the measured and predicted EEG trial-to-trial variability across the 800 ms epoch. The shape of the curve is highly consistent with that observed for the EEG data itself (comparing Fig. 4A and E) (additional analysis of the fidelity of the model is provided in the Supplementary Material Fig. S3).

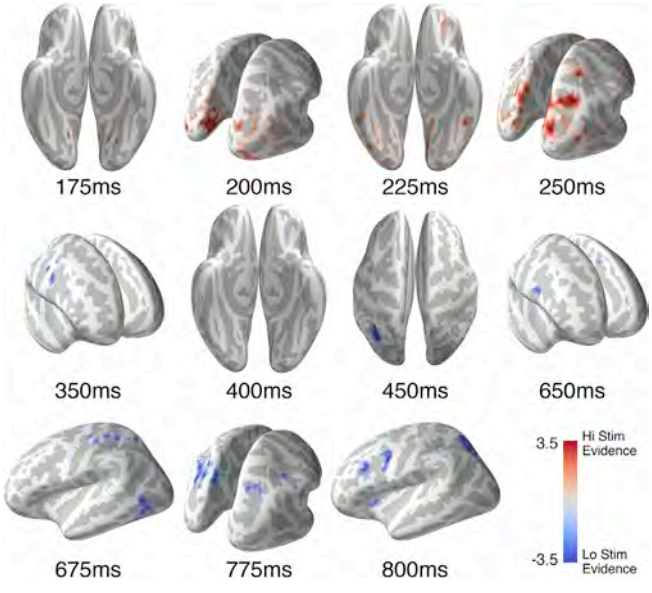

Fig. 5. Group-level encoding model weights results show a cascade of neural activation. Subset of thresholded ($p < 0.05$ FDR-Corrected, $k = 10$) group level statistical parametric maps created by the stTFCE randomization procedure on the encoding model weight matrices. Shown is the progression of spatial activity across the trial. Activation can be seen early in the trial in the occipital regions while progressing more anteriorly later in the trial to executive control areas. Activations in red indicate areas where high stimulus evidence trials had larger activations than low stimulus evidence trials, with blue the inverse relationship.

Given the encoding model, we construct the progression of the BOLD activity across time by identifying voxel weights that are consistent across subjects in space and time (see Methods). Fig. 5 shows these results for a group level analysis. The weights from the encoding model can be interpreted similarly to the β-weights in a traditional GLM contrast. Here, red voxels indicate significant regions where high stimulus evidence variability positively covary between EEG discrimination y-values and BOLD activation and where blue voxels indicate a negative covariation between more negative EEG discrimination y-values (stronger low evidence discrimination) and BOLD activation. We observe a progression of activity (see also Supplementary Material Movie S1), at 25 ms resolution, which proceeds simultaneously down the dorsal and ventral streams of visual processing for the first 250 ms. After that the cascade becomes more complex with activation in the IPS at 425 ms and 750 ms (see Fig. 6A), reactivation of the SPL at 675 ms and activation of ACC at 600 ms. The reactivation pattern is particularly significant since it would not be observable via a traditional fMRI general linear model (GLM) analysis (results of which were presented in Fig. 2E), since it would integrate over time and thus superimpose these activities. For example, the changing sign of the middle temporal gyrus (MT) encoding weights in Fig. 6A manifested as no activity in the MT for the traditional fMRI GLM analysis of this same BOLD data–i.e. the change in sign canceled the effective correlation in the GLM. The areas of activation we find are internally consistent with the traditional GLM results (see Fig. S2E/F) and with previous reports in the literature for human subjects (Heekeren et al., 2004; Philiastides and Sajda, 2007); however, here we are able to link activations across time in a way that was previously only possible with invasive techniques.

3.4. Cortical reactivation correlates with decision confidence

We identified a reactivation pattern in the spatiotemporal dynamics which was not observable using non-simultaneously acquired EEG or fMRI. Further analysis of the spatiotemporal dynamics (see Fig. 6B) shows that the reactivation pattern in the network occurs after decision-monitoring areas become engaged (i.e. after activation of ACC).

Spontaneous reactivation, or "replay," of neural activity in the human brain has been observed and is believed to be important for memory consolidation (Deuker et al., 2013) and more recently has been hypothesized to play a role in perceptual decision-making by enabling the formation of decision confidence (Fleming et al., 2012). To test the hypothesis that the reactivation activity we see may be related to decision confidence, we used a hierarchical drift diffusion model (DDM) (Ratcliff and McKoon, 2008; Wiecki et al., 2013) to fit the behavioral data for high and low stimulus evidence conditions (see Methods). Specifically, our model enables us to define a proxy for decision confidence based on the DDM fits to the behavior (Philiastides et al., 2014; Kiani et al., 2014).

Correlating the reactivation level to this confidence proxy shows a strong and significant monotonic relationship between confidence proxy and the level of reactivation (high stimulus evidence-slope $= 0.037 \pm 0.008$, $t = 4.657$, dof $= 206$, $p = 3.2 \times 10^{-6}$; low stimulus evidence-slope $= 0.062 \pm 0.008$, $t = 7.754$, dof $= 206$, $p = 8.88 \times 10^{-15}$), with low stimulus evidence trials reactivated more strongly than high stimulus evidence trials (difference in slopes $= -0.025 \pm 0.011$, $t = 2.189$,

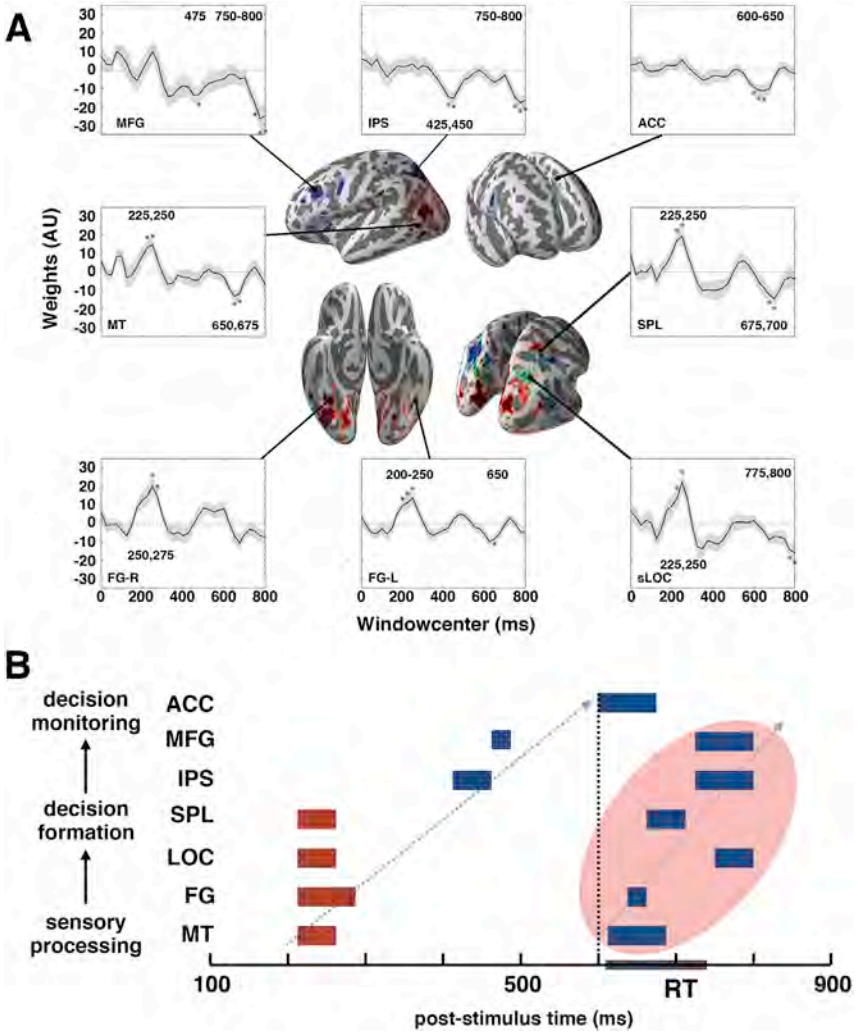

Fig. 6. Spatiotemporal event-related activations show coordinated reactivations. A, Union across time windows of significant voxels for high (red) and low (blue) stimulus evidence activations. Voxels with activations for both high and low conditions (at different time windows) are displayed in green (also see Supplementary Material Fig. S2). Also shown are the encoding model weights for specific voxels, including fusiform gyrus (FG-R):36/-51/-18, (FG-L):-42/-42/-18, superior lateral occipital cortex (sLOC):28/-64/38, superior parietal lobule (SPL):32/-52/54, anterior cingulate cortex (ACC):-8/24/28, intraparietal sulcus (IPS):-30/-60/39, middle frontal gyrus (MFG):-45/27/30, middle temporal gyrus (MT):-57/-60/0. Asterisks indicate significant windows. **B**, Sequence of significant weights showing a "replay" of the network after the onset of ACC activation (shaded ellipse). "Replay" is faster than the initial stimulus driven sequence and strongest for low evidence trials.

dof = 206, $p = 0.029$) (see Fig. 7 and Supplementary Material Fig. S5a). Additionally, reactivation amplitude correlates with behavioral accuracy (Supplementary Material Fig. S5b) (high stimulus evidence, slope = 0.0115 ± 0.0047, $t = 2.41$, dof = 206, $p = 0.016$; low stimulus evidence, slope = 0.0104 ± 0.0047, $t = 2.19$, dof = 206, $p = 0.028$). Additional control analyses from early time windows (175–250 ms) shows no significant correlations between activation and confidence proxy (Supplementary Material Fig. S5).

We next analyzed which cortical regions most contributed to this reactivation. Specifically, we used a recursive feature elimination to identify those regions contributing the most to the correlation between the level of reactivation and the confidence proxy. We found that the IPS/SPL and dorsal lateral prefrontal cortex (DLPFC) contributed the most (Fig. 7C). Interestingly, these two areas have been previously implicated in metacognitive judgments of confidence (Fleming et al., 2012; Steinhauser and Yeung, 2010; Yeung and Summerfield, 2012).

4. Discussion

In this paper we have shown that linking simultaneously acquired EEG and fMRI using a novel encoding model enables us to probe the high-resolution spatiotemporal dynamics in the human brain. We demonstrate the methodology using neural data acquired during a rapid perceptual decision-making task. Our method, which resolves whole-brain activity with EEG-like temporal resolution, uncovers what appears to be a reactivation of neural activity in perceptual decision-making, dynamics that would otherwise be masked by the temporal averaging and slow dynamics of traditional fMRI. More broadly, our results demonstrate a general non-invasive data-driven methodology for measuring, at high spatiotemporal resolution, latent neural processes underlying human behavior. Below we discuss the methodology and our specific findings for perceptual decision-making with respect to the current literature.

The trial-to-trial variability we leverage in our methodology can be viewed as a combination of exogenous and endogenous neural variability expressed in the EEG and fMRI BOLD. The exogenous variability is related to the stimulus presentation while the endogenous variability is subject and trial specific. Endogenous trial-to-trial variability during perceptual decisions has been shown to be associated with attention (Walz et al., 2014), reward (Fouragnan et al., 2015), confidence (Gherman and Philiastides, 2015) and other internal mental states that vary at different times during the perceptual decision. The classifier applied to the EEG data is thus sensitive to both the exogenous variability (stimulus differences) and the endogenous variability at the multiple time steps throughout the decision process.

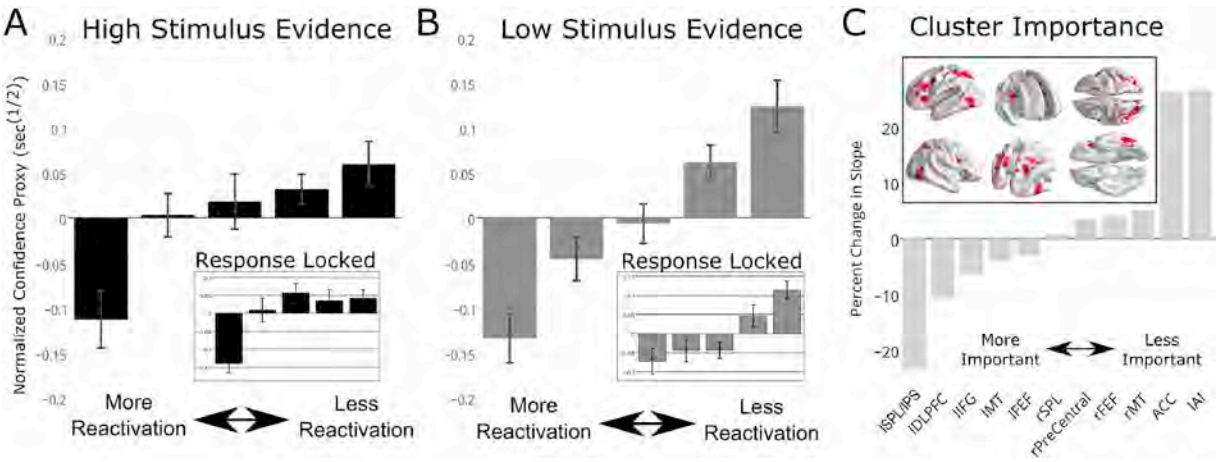

Fig. 7. Trial-to-trial reactivation correlates with decision confidence. Trial-to-trial reactivation amplitude (Y_{ji}^{R} see Methods) of "replay" correlates with confidence proxy for both high (**A**) and low (**B**) stimulus evidence conditions. Error bars represent standard errors across subjects. Response locked analysis results are shown in the insets for (**A**) and (**B**) as well as Supplementary Material Figs. S6–7. **C**, Stimulus-locked replay activation clusters and feature importance. (Inset) Regions of interest used in computing the reactivation values for computing confidence proxy correlations. These regions were taken from significant group activations from 600 to 800 ms post stimulus. Regions were then clustered (> 48 voxels) and a secondary analysis for feature importance was performed. Specifically, we removed each cluster before computing trial-to-trial reactivations and compared the slope of reactivation x confidence proxy when all clusters were present. Panel **C** shows the ranking of feature importance for each cluster (more negative % change = region is more important in contributing to reactivation). A negative change in slope indicates that removing that cluster causes the slope of the correlation between reactivation and confidence to decrease.

The approach we present requires that EEG and BOLD data be collected simultaneously and not in separate sessions in order to exploit the correlations in exogenous and endogenous trial-to-trial variability between EEG and BOLD to temporally "tag" voxels in the space of the fMRI. To show the importance of collecting the data simultaneously, we ran a control analysis that randomly permuted the trials within their stimulus evidence class, thus effectively simulating an EEG and BOLD dataset collected separately. By destroying the link between the EEG and BOLD trials, the encoding model failed to find any consistent activation (Supplementary Material Fig. S8), indicating the necessity of simultaneous acquisition.

Alternative techniques for fusing simultaneous EEG-fMRI typically do not exploit EEG across the trial and instead only analyze specific ERP components or time windows of interest (Huster et al., 2012; Warbrick et al., 2013; De Martino et al., 2010; Jann et al., 2009; Jaspers-Fayer et al., 2012; Mayhew et al., 2013; Novitskiy et al., 2011; Omata et al., 2013; Warbrick et al., 2009; Goldman et al., 2009; Walz et al., 2013, 2014; Fouragnan et al., 2015). Results from these techniques identify regions that modulate with the specific components, but yield limited information about the timing of other task-relevant regions seen in traditional fMRI contrasts. The methodology developed here extends the previously reported work of Goldman et al. (2009) and Walz et al. (2014) by combining their EEG data reduction techniques with techniques developed for encoding stimulus features onto BOLD data (Cukur et al., 2013; Hansen et al., 2007; Kay et al., 2008; Naselaris et al., 2011; Nishimoto et al., 2011; Stansbury et al., 2013), ultimately providing a framework for labeling voxels in task-relevant fMRI contrasts with their timing information (Supplementary Material Fig. S2C/E/F).

Clearly, other EEG components that are task-related can be isolated and could potentially be used to "tag" BOLD data. The sliding window linear classification used here acts to reduce the EEG data along a dimension that categorizes stimulus evidence; however, this could be replaced by any other data reduction technique, such as temporally windowed ICA or PCA. Variability along these component directions could then be used in the encoding model to link with the simultaneously collected BOLD data. The choice of data reduction technique (i.e. feature space) would be highly dependent on the nature of the inferences.

In terms what our methodology is able to potentially reveal about the spatiotemporal brain dynamics of perceptual decision making, we observe what appears to be reactivation of the pre-response network, spatiotemporal dynamics that would be masked using traditional fMRI analysis. Interestingly, the reactivation terminated in a network that included the MFG, SPL, and IPS, similar areas previously reported to be reactivated in metacognitive judgments of confidence in perceptual decisions (Fleming et al., 2012; Steinhauser and Yeung, 2010; Yeung and Summerfield, 2012). In addition, these areas contributed the most to the correlation to our confidence proxy. Gherman and Philiastides (2015) observed this network using a multivariate single-trial EEG approach, coupled with a distributed source reconstruction technique. Fleming et al. (2012) and Heereman et al. (2015) used BOLD fMRI to show that areas in this network negatively correlate with subjective certainty ratings. Unique to our findings, we saw this reactivation on a single-trial basis after engagement of the ACC, which has been shown to be involved in decision monitoring (Gherman and Philiastides, 2015; Botvinick et al., 2001), and also observed the dynamic sequence leading up to this network reactivation. Our results showed that reactivation/replay occurred on a trial-to-trial basis after a decision, was stronger for difficult decisions, and correlated with a proxy for decision confidence (Fig. 7).

To ensure that our findings were not simply an artifact of a stimulus-locked analysis, we performed the same analysis response-locked. We found that the reactivation, both its correlation with the confidence proxy and the cortical areas implicated, to be consistent with the stimulus locked analysis (see Supplementary Material Fig. S6). In addition we found that the reactivation clearly begins pre-response and extends through and past the timing of the behavioral response (Supplementary Material Fig. S7). Thus our complete analysis suggests that while in the process of making and executing a perceptual decision, the brain may replay activity to develop a metacognitive representation of its decision.

We note several caveats in terms of the interpretation of our novel results as they relate to perceptual decision-making. One potential criticism is that the variability we discriminate and encode in the BOLD is related to the stimulus-evidence and not a decision variable such the choice (e.g. trial-to-trial variability of a subjects decision of a face, car or house). Though our previous work has shown that there are EEG components that are sensitive to stimulus category or choice, these components tend to be isolated at specific time windows (Philiastides and Sajda, 2006; Philiastides et al., 2006). The level of stimulus evidence, on the other hand, spans most of the trial and thus can be used to provide a more complete picture of brain dynamics during the task. Ultimately the reactivation we observe in these dynamics is intriguing and thus we conducted additional analyses that points to a novel "replay" hypothesis,

namely that the reactivation is a replay of activity associated with the human brain generating, on a trial-to-trial basis, a meta-cognitive judgment of confidence immediately after the decision.

Another potential source of bias in our results comes from the autocorrelation of the beta estimates in our deconvolution procedure. Since there is temporal autocorrelation in the beta's, our leave-one-out cross validation is not the optimal analysis design (Mumford et al., 2014). However we decided to run the leave-one-out cross validation to remain consistent with previous analyses.

Of course additional experiments must be done to more thoroughly test our "replay" hypothesis. For example, future work will investigate other measures of confidence that are more direct than the proxy we estimated from the DDM. We will also consider different types of rapid perceptual decision-making tasks to investigate whether the results we see generalize across stimulus types and/or the modality of stimulus presentation. In general, we have shown that simultaneously acquired EEG/fMRI data enables a novel non-invasive approach to visualize high resolution spatial and temporal processing in the human brain with the potential for providing a more comprehensive understanding of the neural basis of complex behaviors.

Acknowledgements

We would like to thank Jianing Shi for assistance in collecting the EEG/fMRI data. This work was funded by National Institutes of Health Grant R01-MH085092, DARPA under Contract NBCHC090029 and the Army Research Laboratory and under Cooperative Agreement Number W911NF-10-2-0022.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.neuroimage.2017.06.059.

References

Alexander, D.M., Trengove, C., van Leeuwen, C., 2015. Donders is dead: cortical traveling waves and the limits of mental chronometry in cognitive neuroscience. Cogn. Process. 16 (4), 365–375. http://dx.doi.org/10.1007/s10339-015-0662-4.

Banko, E.M., Gal, V., Kortvelyes, J., Kovacs, G., Vidnyanszky, Z., 2011. Dissociating the effect of noise on sensory processing and overall decision difficulty. J. Neurosci. 31 (7), 2663–2674. http://dx.doi.org/10.1523/JNEUROSCI.2725-10.2011.

Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. J. Stat. Softw. 67 (1), 48. http://dx.doi.org/10.18637/jss.v067.i01.

Baumeister, S., Hohmann, S., Wolf, I., Plichta, M., Rechtsteiner, S., Zangl, M., Ruf, M., Holz, N., Boecker, R., Meyer-Lindenberg, A., Holtmann, M., Laucht, M., Banaschewski, T., Brandeis, D., 2014. Sequential inhibitory control processes assessed through simultaneous EEG-fMRI. NeuroImage 94, 349–359. http://dx.doi.org/10.1016/j.neuroimage.2014.01.023.

Botvinick, M.M., Braver, T.S., Barch, D.M., Carter, C.S., Cohen, J.D., 2001. Conflict monitoring and cognitive control. Psychol. Rev. 108 (3), 624–652.

Conroy, B.R., Walz, J.M., Sajda, P., 2013. Fast bootstrapping and permutation testing for assessing reproducibility and interpretability of multivariate FMRI decoding models. PloS one 8 (11), e79271. http://dx.doi.org/10.1371/journal.pone.0079271.

Cukur, T., Nishimoto, S., Huth, A.G., Gallant, J.L., 2013. Attention during natural vision warps semantic representation across the human brain. Nat. Neurosci. 16 (6), 763–770. http://dx.doi.org/10.1038/nn.3381.

Dahne, S., Biessmann, F., Samek, W., Haufe, S., Goltz, D., Gundlach, C., Villringer, A., Fazli, S., Muller, K.-R., 2015. Multivariate machine learning methods for fusing multimodal functional neuroimaging data. Proc. IEEE 103 (9), 1507–1530. http://dx.doi.org/10.1109/JPROC.2015.2425807.

Dakin, S.C., Hess, R.F., Ledgeway, T., Achtman, R.L., 2002. What causes non-monotonic tuning of fMRI response to noisy images? Curr. Biol. 12 (14), 476–477. http://dx.doi.org/10.1016/S0960-9822(02)00960-0.

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. Magn. Reson. Imaging 28 (8), 1104–1112. http://dx.doi.org/10.1016/j.mri.2009.12.026.

Deuker, L., Olligs, J., Fell, J., Kranz, T.A., Mormann, F., Montag, C., Reuter, M., Elger, C.E., Axmacher, N., 2013. Memory consolidation by replay of stimulus-specific neural activity. J. Neurosci. 33 (49), 19373–19383. http://dx.doi.org/10.1523/JNEUROSCI.0414-13.2013.

Fleming, S., Huijgen, J., Dolan, R., 2012. Prefrontal contributions to metacognition in perceptual decision making. J. Neurosci. 32 (18), 6117–6125. http://dx.doi.org/10.1523/JNEUROSCI.6489-11.2012.

Fouragnan, E., Retzler, C., Mullinger, K., Philiastides, M.G., 2015. Two spatiotemporally distinct value systems shape reward-based learning in the human brain. Nat. Commun. 6, 8107. http://dx.doi.org/10.1038/ncomms9107.

Gherman, S., Philiastides, M.G., 2015. Neural representations of confidence emerge from the process of decision formation during perceptual choices. NeuroImage 106, 134–143. http://dx.doi.org/10.1016/j.neuroimage.2014.11.036.

Goldman, R.I., Wei, C.-Y., 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 (1), 136–147. http://dx.doi.org/10.1016/j.neuroimage.2009.03.062.

Greve, D.N., Fischl, B., 2009. Accurate and robust brain image alignment using boundary-based registration. NeuroImage 48, 63–72. http://dx.doi.org/10.1016/j.neuroimage.2009.06.060.

Hansen, K.A., Kay, K.N., Gallant, J.L., 2007. Topographic organization in and near human visual area V4. J. Neurosci. 27, 11896–11911. http://dx.doi.org/10.1523/JNEUROSCI.2991-07.2007.

Heekeren, H.R., Marrett, S., Bandettini, P.A., Ungerleider, L.G., 2004. A general mechanism for perceptual decision-making in the human brain. Nature 431 (7010), 859–862. http://dx.doi.org/10.1038/nature02966.

Heereman, J., Walter, H., Heekeren, H.R., 2015. A task-independent neural representation of subjective certainty in visual perception. Front. Hum. Neurosci. 9, 551. http://dx.doi.org/10.3389/fnhum.2015.00551.

Horikawa, T., Tamaki, M., Miyawaki, Y., Kamitani, Y., 2013. Neural decoding of visual imagery during sleep. Sci. (New York, N.Y.) 340 (6132), 639–642. http://dx.doi.org/10.1126/science.1234330.

Huster, R.J., Debener, S., Eichele, T., Herrmann, C.S., 2012. Methods for simultaneous EEG-fMRI: an introductory review. J. Neurosci. 32 (18), 6053–6060. http://dx.doi.org/10.1523/JNEUROSCI.0447-12.2012.

Huth, A.G., de Heer, W.A., Griffiths, T.L., Theunissen, F.E., Gallant, J.L., 2016. Natural speech reveals the semantic maps that tile human cerebral cortex. Nature 532 (7600), 453–458. http://dx.doi.org/10.1038/nature17637.

Jann, K., Dierks, T., Boesch, C., Kottlow, M., Strik, W., Koenig, T., 2009. BOLD correlates of EEG alpha phase-locking and the fMRI default mode network. NeuroImage 45 (3), 903–916. http://dx.doi.org/10.1016/j.neuroimage.2009.01.001.

Jaspers-Fayer, F., Ertl, M., Leicht, G., Leupelt, A., Mulert, C., 2012. Single-trial EEG-fMRI coupling of the emotional auditory early posterior negativity. NeuroImage 62 (3), 1807–1814. http://dx.doi.org/10.1016/j.neuroimage.2012.05.018.

Jorge, J., van der Zwaag, W., Figueiredo, P., 2014. EEG-fMRI integration for the study of human brain function. NeuroImage 102, 24–34. http://dx.doi.org/10.1016/j.neuroimage.2013.05.114.

Kay, K.N., Naselaris, T., Prenger, R.J., Gallant, J.L., 2008. Identifying natural images from human brain activity. Nature 452 (7185), 352–355. http://dx.doi.org/10.1038/nature06713.

Kiani, R., Corthell, L., Shadlen, M.N., 2014. Choice certainty is informed by both evidence and decision time. Neuron 84 (6), 1329–1342. http://dx.doi.org/10.1016/j.neuron.2014.12.015.

Logothetis, N.K., 2008. What we can do and what we cannot do with fMRI. Nature 453 (7197), 869–878. http://dx.doi.org/10.1038/nature06976.

Martino, F.D., Valente, G., Staeren, N., Ashburner, J., Goebel, R., Formisano, E., 2008. Combining multivariate voxel selection and support vector machines for mapping and classification of fMRI spatial patterns. NeuroImage 43 (1), 44–58. http://dx.doi.org/10.1016/j.neuroimage.2008.06.037.

Mayhew, S.D., Ostwald, D., Porcaro, C., Bagshaw, A.P., 2013. Spontaneous EEG alpha oscillation interacts with positive and negative BOLD responses in the visual-auditory cortices and default-mode network. NeuroImage 76, 362–372. http://dx.doi.org/10.1016/j.neuroimage.2013.02.070.

Mumford, J.A., Turner, B.O., Ashby, F.G., Poldrack, R.A., 2012. Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. NeuroImage 59, 2636–2643. http://dx.doi.org/10.1016/j.neuroimage.2011.08.076.

Mumford, J.A., Davis, T., Poldrack, R.A., 2014. The impact of study design on pattern estimation for single-trial multivariate pattern analysis. NeuroImage 103, 130–138. http://dx.doi.org/10.1016/j.neuroimage.2014.09.026.

Muraskin, J., Sherwin, J., Sajda, P., 2015. Knowing when not to swing: EEG evidence that enhanced perception-action coupling underlies baseball batter expertise. NeuroImage 123, 1–10. http://dx.doi.org/10.1016/j.neuroimage.2015.08.028.

Muraskin, J., Sherwin, J., Lieberman, G., Garcia, J.O., Verstynen, T., Vettel, J.M., Sajda, P., 2017. Fusing multiple neuroimaging modalities to assess group differences in perception action coupling. Proc. IEEE (105), 83–100. http://dx.doi.org/10.1109/jproc.2016.2574702.

Naselaris, T., Kay, K.N., Nishimoto, S., Gallant, J.L., 2011. Encoding and decoding in fMRI. NeuroImage 56 (2), 400–410. http://dx.doi.org/10.1016/j.neuroimage.2010.07.073.

Nguyen, V.T., Cunnington, R., 2014. The superior temporal sulcus and the N170 during face processing: single trial analysis of concurrent EEG-fMRI. NeuroImage 86, 492–502. http://dx.doi.org/10.1016/j.neuroimage.2013.10.047.

Nichols, T., Hayasaka, S., 2003. Controlling the familywise error rate in functional neuroimaging: a comparative review. Stat. Methods Med. Res. 12, 419–446. http://dx.doi.org/10.1191/0962280203sm341ra.

Nishimoto, S., Vu, A.T., Naselaris, T., Benjamini, Y., Yu, B., Gallant, J.L., 2011. Reconstructing visual experiences from brain activity evoked by natural movies. Curr. Biol. 21, 1641–1646. http://dx.doi.org/10.1016/j.cub.2011.08.031.

Novitskiy, N., Ramautar, J.R., Vanderperren, K., De Vos, M., Mennes, M., Mijovic, B., Vanrumste, B., Stiers, P., Van den Bergh, B., Lagae, L., Sunaert, S., Van Huffel, S., Wagemans, J., 2011. The BOLD correlates of the visual P1 and N1 in single-trial

analysis of simultaneous EEG-fMRI recordings during a spatial detection task. NeuroImage 54 (2), 824–835. http://dx.doi.org/10.1016/j.neuroimage.2010.09.041.

Omata, K., Hanakawa, T., Morimoto, M., Honda, M., 2013. Spontaneous slow fluctuation of EEG alpha rhythm reflects activity in deep-brain structures: a simultaneous EEG-fMRI study. PloS one 8 (6), e66869. http://dx.doi.org/10.1371/journal.pone.0066869.

Parra, L.C., Spence, C.D., Gerson, A.D., Sajda, P., 2005. Recipes for the linear analysis of EEG. NeuroImage 28, 326–341. http://dx.doi.org/10.1016/j.neuroimage.2005.05.032.

Philiastides, M.G., Sajda, P., 2006. Temporal characterization of the neural correlates of perceptual decision making in the human brain. Cereb. Cortex 16 (4), 509–518. http://dx.doi.org/10.1093/cercor/bhi130.

Philiastides, M.G., Sajda, P., 2007. EEG-informed fMRI reveals spatiotemporal characteristics of perceptual decision making. J. Neurosci. 27 (48), 13082–13091. http://dx.doi.org/10.1523/JNEUROSCI.3540-07.2007.

Philiastides, M.G., Ratcliff, R., Sajda, P., 2006. Neural representation of task difficulty and decision making during perceptual categorization: a timing diagram. J. Neurosci. 26 (35), 8965–8975. http://dx.doi.org/10.1523/JNEUROSCI.1655-06.2006.

Philiastides, M.G., Heekeren, H.R., Sajda, P., 2014. Human scalp potentials reflect a mixture of decision-related signals during perceptual choices. J. Neurosci. 34 (50), 16877–16889. http://dx.doi.org/10.1523/JNEUROSCI.3012-14.2014.

Plichta, M.M., Wolf, I., Hohmann, S., Baumeister, S., Boecker, R., Schwarz, A.J., Zangl, M., Mier, D., Diener, C., Meyer, P., Holz, N., Ruf, M., Gerchen, M.F., Bernal-Casas, D., Kolev, V., Yordanova, J., Flor, H., Laucht, M., Banaschewski, T., Kirsch, P., Meyer-Lindenberg, A., Brandeis, D., 2013. Simultaneous EEG and fMRI reveals a causally connected subcortical-cortical network during reward anticipation. J. Neurosci. 33 (36), 14526–14533. http://dx.doi.org/10.1523/JNEUROSCI.0631-13.2013.

Ratcliff, R., McKoon, G., 2008. The diffusion decision model: theory and data for two-choice decision tasks. Neural Comput. 20 (4), 873–922. http://dx.doi.org/10.1162/neco.2008.12-06-420.

Sajda, P., Goldman, R., Philiastides, M., Gerson, A., Brown, T., 2007. A system for single-trial analysis of simultaneously acquired EEG and fMRI. In: 2007 3rd International IEEE/EMBS Conference on Neural Engineering, Kohala Coast, HI, pp. 287–290. http://dx.doi.org/10.1109/CNE.2007.369667.

Sajda, P., Philiastides, M.G., Parra, L.C., 2009. Single-trial Analysis of neuroimaging data: inferring neural networks underlying perceptual decision-making in the human brain. IEEE Rev. in Biomed. Eng. 2, 97–109. http://dx.doi.org/10.1109/RBME.2009.2034535.

Sajda, P., Goldman, R.I., Dyrholm, M., Brown, T.R., 2010. Signal processing and machine learning for single-trial analysis of simultaneously acquired EEG and fMRI. In: Statistical Signal Processing for Neuroscience and Neurotechnology. Academic Press, pp. 311–334. http://dx.doi.org/10.1016/B978-0-12-375027-3.00009-0.

Sherwin, J., Sajda, P., 2013. Musical experts recruit action-related neural structures in harmonic anomaly detection: evidence for embodied cognition in expertise. Brain Cogn. 83 (2), 190–202. http://dx.doi.org/10.1016/j.bandc.2013.07.002.

Sherwin, J., Muraskin, J., Sajda, P., 2012. You can't think and hit at the same time: neural correlates of baseball pitch classification. Front. Neurosci. 6, 177. http://dx.doi.org/10.3389/fnins.2012.00177.

Sherwin, J.S., Muraskin, J., Sajda, P., 2015. Pre-stimulus functional networks modulate task performance in time-pressured evidence gathering and decision-making. NeuroImage 111, 513–525. http://dx.doi.org/10.1016/j.neuroimage.2015.01.023.

Silbert, L.J., Honey, C.J., Simony, E., Poeppel, D., Hasson, U., 2014. Coupled neural systems underlie the production and comprehension of naturalistic narrative speech. Proc. Natl. Acad. Sci. 111, E4687–E4696. http://dx.doi.org/10.1073/pnas.1323812111.

Smith, S.M., Nichols, T.E., 2009. Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. NeuroImage 44 (1), 83–98. http://dx.doi.org/10.1016/j.neuroimage.2008.03.061.

Stansbury, D., Naselaris, T., Gallant, J., 2013. Natural scene statistics account for the representation of scene categories in human visual cortex. Neuron 79, 1025–1034. http://dx.doi.org/10.1016/j.neuron.2013.06.034.

Steinhauser, M., Yeung, N., 2010. Decision processes in human performance monitoring. J. Neurosci. 30 (46), 15643–15653. http://dx.doi.org/10.1523/JNEUROSCI.1899-10.2010.

Walz, J.M., Goldman, R.I., Carapezza, M., Muraskin, J., Brown, T.R., Sajda, P., 2013. Simultaneous EEG-fMRI reveals temporal evolution of coupling between supramodal cortical attention networks and the brainstem. J. Neurosci. 33 (49), 19212–19222. http://dx.doi.org/10.1523/JNEUROSCI.2649-13.2013.

Walz, J.M., Goldman, R.I., Carapezza, M., Muraskin, J., Brown, T.R., Sajda, P., 2014. Simultaneous EEGfMRI reveals a temporal cascade of task-related and default-mode activations during a simple target detection task. NeuroImage 102, 229–239. http://dx.doi.org/10.1016/j.neuroimage.2013.08.014.

Warbrick, T., Mobascher, A., Brinkmeyer, J., Musso, F., Richter, N., Stoecker, T., Fink, G.R., Shah, N.J., Winterer, G., 2009. Single-trial P3 amplitude and latency informed event-related fMRI models yield different BOLD response patterns to a target detection task. NeuroImage 47 (4), 1532–1544. http://dx.doi.org/10.1016/j.neuroimage.2009.05.082.

Warbrick, T., Arrubla, J., Boers, F., Neuner, I., Shah, N.J., 2014. Attention to detail: why considering task demands is essential for single-trial analysis of BOLD correlates of the visual P1 and N1. J. Cogn. Neurosci. 26 (3), 529–542. http://dx.doi.org/10.1162/jocn.

Wiecki, T.V., Sofer, I., Frank, M.J., 2013. HDDM: hierarchical bayesian estimation of the drift-diffusion model in Python. Front. Neuroinfo. 7, 14. http://dx.doi.org/10.3389/fninf.2013.00014.

Yeung, N., Summerfield, C., 2012. Metacognition in human decision-making: confidence and error monitoring. Philos. Trans. R. Soc. Lond B Biol. Sci. 367 (1594), 1310–1321. http://dx.doi.org/10.1098/rstb.2011.0416.

Yuan, H., Liu, T., Szarkowski, R., Rios, C., Ashe, J., He, B., 2010. Negative covariation between task-related responses in alpha/beta-band activity and BOLD in human sensorimotor cortex: an EEG and fMRI study of motor imagery and movements. NeuroImage 49 (3), 2596–2606. http://dx.doi.org/10.1016/j.neuroimage.2009.10.028.

Zhang, Y., Brady, M., Smith, S., 2001. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. IEEE Trans. Med. Imaging 20, 45–57. http://dx.doi.org/10.1109/42.906424.