ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 28 (2005) 342 – 353

Cortical origins of response time variability during rapid discrimination of visual objects

Adam D. Gerson,[a] Lucas C. Parra,[b] and Paul Sajda[a,*]

[a]*Department of Biomedical Engineering, Columbia University, New York, NY 10027, USA*
[b]*Department of Biomedical Engineering, City College of New York, New York, NY 10031, USA*

Received 8 November 2004; revised 23 May 2005; accepted 1 June 2005
Available online 19 September 2005

In this paper, we use single-trial analysis of electroencephalography (EEG) to ascertain the cortical origins of response time variability in a rapid serial visual presentation (RSVP) task. We extract spatial components that maximally discriminate between target and distractor stimulus conditions over specific time windows between stimulus onset and the time of a motor response. We then compute the peak latency of this differential activity on a trial-by-trial basis, and correlate this with response time. We find, for our nine participants, that the majority of the latency is introduced by component activity which begins far-frontally 200 ms prior to the response and proceeds to become parietally distributed near the time of response. This activity is consistent with the hypothesis that cortical networks involved in generating the late positive complexes may be the origins of the observed response time variability in rapid discrimination of visual objects.

Keywords: Electroencephalogram (EEG); Response time (RT); Response variability; Rapid serial visual presentation (RSVP); Linear discrimination; Single-trial

Introduction

Important for understanding the relationship between neural activity and behavior is identifying the cortical origins of behavioral variability. In many visual discrimination and recognition tasks, significant variability in response time is observed across trials. This variability may be due to a variety of factors ranging from the difficulty in discriminating an object on any given trial, trial-by-trial variability of the subject's engagement in the task, or intrinsic variability of neural processing. Identifying neural activity that is correlated with response time variability may shed light on the underlying cortical networks responsible for perceptual decision making processes and the processing latencies that these networks may introduce for a given task.

Neural origins of response time variability have been extensively studied in animals, including primates, using single and multi-unit recordings. For example, Roitma and Shadlen (2002) have shown that for a two alternative forced choice (2-AFC) motion discrimination task, the latency observed in a monkey's response (saccades to the target) is directly related to the integration of evidence (percent coherence) by neurons in lateral inter-parietal cortex (LIP). LIP activity can thus be used to directly predict the time of the saccade, with variability in response time correlated with the time at which the LIP activity reaches a given threshold. Though important for understanding the relationship between neural activity and behavioral variability, this work has been confined to the study of response time variability in animals, given the invasiveness of the recordings, as well as highly spatially localized (single neurons) activity predictive of response time. We therefore wanted to consider whether the neural origins of response time variability could be identified in human subjects and the specific spatial and temporal distribution of that activity.

To study response time variability in humans requires a non-invasive approach. Modalities such as functional magnetic resonance imaging (fMRI) have been used extensively to study cortical activity and its relationship to behavior (Toga and Mazziotta, 2000). Though by current non-invasive functional imaging standards fMRI possesses excellent spatial resolution, it has limited temporal resolution (Kim et al., 1997). Since response times across trials can vary on a millisecond time scale, fMRI does not possess sufficient temporal resolution to disentangle the cortical sources of response time variability.

Traditionally, psychophysiologists have exploited electrophysiological responses to stimuli to infer the nature of cognitive processes with millisecond precision. Long-studied event-related potentials (ERPs), such as the visually evoked potential, N200, P300, readiness potential, and error related negativity, arise from neural tissue in regions believed to be associated with visual stimulus encoding, classification, recognition, motor planning, and response evaluation. These potentials, derived by stimulus or response-locked

* Corresponding author.
E-mail addresses: adg71@columbia.edu (A.D. Gerson), parra@ccny.cuny.edu (L.C. Parra), ps629@columbia.edu (P. Sajda).
Available online on ScienceDirect (www.sciencedirect.com).

1053-8119/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2005.06.026

trial-averaging, are used to develop and validate models describing the timing and interactions between mental processes.

Interaction between stimulus characteristics and ERP features that covaries with response time provides evidence for hypotheses concerning the nature of information transfer between mental processes. For instance, one study examined the effects of luminance on response time and early components of the visually evoked potential during a simple reaction task. Correlating response times at different luminances with peak latencies of N80, P100, and N130 reveals that the slope of linear regression increases with successive ERP peaks (Kammer et al., 1999). This finding indicates dispersion along early visual processing stages and suggests that successive processes introduce greater delays which result in increased response times.

Surprisingly, few studies have yet to explore the correlation between single-trial variability of response times and late potentials (>150 ms) associated with visual recognition. Those that do exist have focused primarily on characterizing the functional significance of P300. For instance, during a forced choice visual discrimination task, P300 latency and response time were significantly influenced by stimulus discriminability, while only response time was significantly affected by stimulus–response compatibility indicating that P300 is sensitive to stimulus evaluation processes rather than processes related to response selection and execution (McCarthy and Donchin, 1981).

In this paper, we consider target detection using a rapid serial visual presentation (RSVP) paradigm and use single-trial spatial integration of high-density electroencephalography to identify the time course and cortical origins leading to response time variability. High spatial density EEG permits the extraction of spatial components that optimally discriminate between task conditions (Parra et al., 2002). We use linear discrimination to determine spatial weighting coefficients that optimally discriminate between EEG resulting from different RSVP task conditions over specific temporal windows between stimulus and response. Spatial integration enhances signal quality without loss of temporal precision common to trial averaging in ERP studies. The resulting discriminating components describe activity specific to target recognition and subsequent response. We estimate peak amplitude and latency of discriminating components by fitting a parametric function for the component to individual trials. The parameters of these fits are then used to estimate the fraction of response time variability introduced by processing stages involved with target recognition.

Methods

Participants

Nine participants (3 females and 6 males, mean age 29 years, range 21–37 years) from the Columbia University community volunteered and received no payment. All had normal or corrected to normal vision and reported no history of neurological problems. Informed consent was obtained from all participants in accordance with the guidelines and approval of the Columbia University Institutional Review Board.

Data acquisition

EEG data were acquired in an electrostatically shielded room (ETS-Lindgren, Glendale Heights, IL) using a Sensorium EPA-6 Electrophysiological Amplifier (Charlotte, VT) from 60 Ag/AgCl scalp electrodes mounted in a standard electrode cap (Electro-Cap, Eaton, OH) at locations based on the International 10–20 system and from three periocular electrodes placed below the left eye and at the left and right outer canthi. All channels were referenced to the left mastoid with input impedance <15 kΩ and chin ground. Data were sampled at 1000 Hz with an analog pass band of 0.01–300 Hz using 12 dB/octave high pass and 8th order Elliptic low pass filters. Subsequently, a software-based 0.5-Hz high pass filter was used to remove DC drifts and 60-Hz and 120-Hz (harmonic) notch filters were applied to minimize line noise artifacts. These filters were designed to be linear-phase to prevent delay distortions. Motor response and stimulus events recorded on separate channels were delayed to match latencies introduced by the digital filtering of the EEG.

Behavioral paradigm

During the RSVP task, participants were presented with a continuous sequence of natural scenes. Participants completed four blocks of 50 sequences each with a rest period lasting no more than 5 min between blocks. Each sequence consisted of 50 images and had a 50% chance of containing one target image with one or more people in a natural scene. These target images could only appear within the middle 30 images of each 50-image sequence. The remaining natural scenes without a person are referred to as distractor images. Each image was presented for 100 ms. A fixation cross was displayed for 2 s between sequences. Participants were instructed to press the left button of a generic 3-button mouse with their right index finger while the fixation cross was present, and release the button as soon as they recognized a target image. An illustration of this paradigm is given in Fig. 1.

Stimuli

Images were selected randomly with replacement from a queue of 251 non-target and 33 target gray scale images. Images were obtained with a Kodak DCS420 digital camera with a 28-mm camera lens (Rochester, NY) (van Hateren and van der Schaaf, 1998). Images were deblurred and the size of the images was reduced from 1536 × 1024 pixels to 640 × 426 pixels to ensure precise timing of stimulus presentation. The images were non-linearly transformed via gamma correction to match their mean luminance. The fixation cross display had the same mean luminance as the images. A Dell Precision 530 Workstation (Round Rock, TX) with nVidia Quadro4 900XGL graphics card (Santa Clara, CA) and E-Prime software (Psychological Software Tools, Pittsburgh, PA) controlled stimulus display. An LCD projector (InFocus LP130, Wilsonville, OR) projected stimuli through an RF shielded window onto a front projection screen. Stimuli subtended 33° ± 3° × 25° ± 3° of visual field. Target images were visually inspected to ensure that target objects did not comprise more than 25% of the area (fraction of pixel) in the scene.

Eye movement artifact reduction

Immediately prior to the RSVP task, participants completed an eye motion calibration experiment during which they were instructed to blink repeatedly upon the appearance of a white on black fixation cross and then make several horizontal and vertical eye movements according to the position of a fixation cross

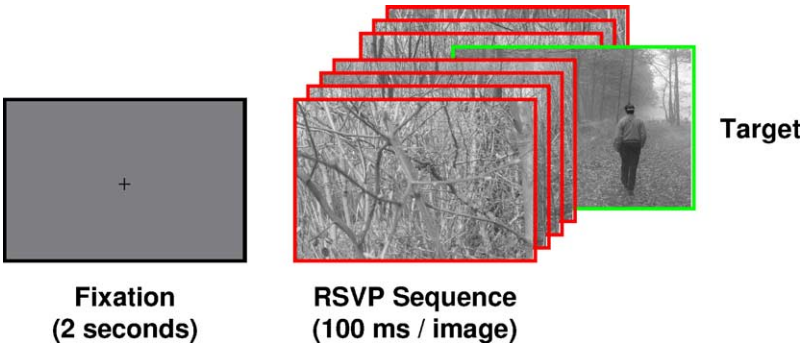

Fig. 1. Example Rapid Serial Visual Presentation (RSVP) trial. A fixation cross lasting 2 s is followed by a sequence of 50 images. Each sequence has a 50% probability of containing one target image. This target can only appear within the middle 30 images to ensure that a 1-s image buffer precedes and follows the target.

subtending 1° × 1° of the visual field. Horizontal eye movements subtended 33° ± 3° and vertical eye movements subtended 25° ± 3°. The timing of these visual cues was recorded simultaneously with EEG. This enabled determination of linear components associated with eye blinks and eye movements that were subsequently projected out of EEG recorded during the RSVP task.

Components for horizontal and vertical eye movements were derived from the difference in the means across samples during which participants held left/right and up/down eye positions.

$$\mathbf{v_h} = \langle \mathbf{x}(t) \rangle_{\text{left}} - \langle \mathbf{x}(t) \rangle_{\text{right}}$$

$$\mathbf{v_v} = \langle \mathbf{x}(t) \rangle_{\text{up}} - \langle \mathbf{x}(t) \rangle_{\text{down}} \quad (1)$$

$\mathbf{x}(t)$ is a column vector representing N scalp potentials sampled at time t, and $\langle \mathbf{x}(t) \rangle_{\text{condition}}$ represents a sample average over times corresponding to the respective conditions. The eye blink component, $\mathbf{V_b}$, was derived as the orientation in N dimensional space that captured most of the power in the signal, i.e., the first principal component, during a series of eye blinks. With this information eye movement artifacts can be subtracted from the data with:

$$\mathbf{x}(t) \leftarrow \left[\mathbf{I} - \mathbf{V}\mathbf{V}^{\#}\right]\mathbf{x}(t) \quad (2)$$

where, $\mathbf{V} = [\mathbf{v_b}, \mathbf{v_h}, \mathbf{v_v}]$, is an $N \times 3$ matrix containing the vectors corresponding to the components of the eye-blink and left/right and up/down saccades, and $\mathbf{V}^{\#}$ is the pseudo-inverse of $\mathbf{V}$. In the new EEG data $\mathbf{x}(t)$, the activity due to eye movement artifacts has been subtracted (Parra et al., 2005).

Spatial integration for identifying discriminating components

Conventional event-related potential (ERP) analysis relies on averaging across trials to boost the amplitude of EEG components associated with event-related neural activity while minimizing the contribution from uncorrelated neural activity and artifacts due to muscle activity and environmental noise. Averaging across trials inevitably conceals variability in amplitude and latency of EEG components which may provide insight into interactions between cortical sources responsible for behavioral responses. Clearly, single-trial analysis of EEG is difficult due to the presence of interference and artifacts. One common approach is to sum the activity across electrodes rather than across trials. Typically, a weighted sum of the electrode activity is used to generate a new composite signal, sometimes referred to as a component of the EEG signal. For instance, independent component analysis (ICA) (Hyvarinen et al., 2001) finds a set of weight vectors such that the corresponding component signals are statistically independent. Applying this method to EEG typically finds several task- and artifact-related components (Makeig et al., 1996,1999,2004). The identity of each component must be determined by inspection. This problem is alleviated when discriminating between EEG evoked by multiple conditions. The method of common spatial patterns (Ramoser et al., 2000) finds a set of weight vectors such that the power of the resulting components are maximally discriminative. Typically, 10% of these components are associated with cortical activity describing differences between conditions. The temporal relationship between such components may not be clear.

Logistic regression can be used to find an optimal basis for discriminating between two conditions over a specific temporal window (Parra et al., 2002). Discrimination permits us to study differences between EEG resulting from task conditions. Specifically, we define a training window, shown in Fig. 2, starting at a relative onset time τ, with a duration of δ and orientation θ (relative to response time) and use logistic regression to estimate a spatial weighting vector $\mathbf{w}_{\tau,\delta,\theta}$ which maximally discriminates between sensor array signals for two conditions, $c \in (0,1)$, where $c = 1$ represents a target trial and $c = 0$ a non-target trial. The spatial weighting vector generates a component[1]

$$y_c(t) = \mathbf{w}_{\tau,\delta,\theta}^T \mathbf{x}_c(t), \quad (3)$$

that best separates the EEG signals by making $y_0(t) < y_1(t)$ for as many samples in the training window as possible. Here the result is a discriminating component specific to target recognition activity while minimizing activity correlated with both task conditions such as early visual processing.

The assumption in logistic regression is that the data when projected onto coordinate, $y = \mathbf{w}^T\mathbf{x} + b$, is distributed according to a logistic function, i.e., the likelihood that sample x belongs to the class of positive examples, $c = 1$, follows

$$p(c = 1|\mathbf{x}) = f(y) = \frac{1}{1 + e^{-y}} = \frac{1}{1 + e^{-(\mathbf{w}^T\mathbf{x}+b)}}. \quad (4)$$

This likelihood is parameterized by the projection vector $\mathbf{w}$ and a bias b. The likelihood for negative examples, $p(c = 0 \mid \mathbf{x}) = 1 - f(x)$, is also a logistic. The hyper-plane orthogonal to $\mathbf{w}$ and

[1] We use the term "component" instead of "source" to make it clear that this is a projection of all the activity correlated with the underlying source.

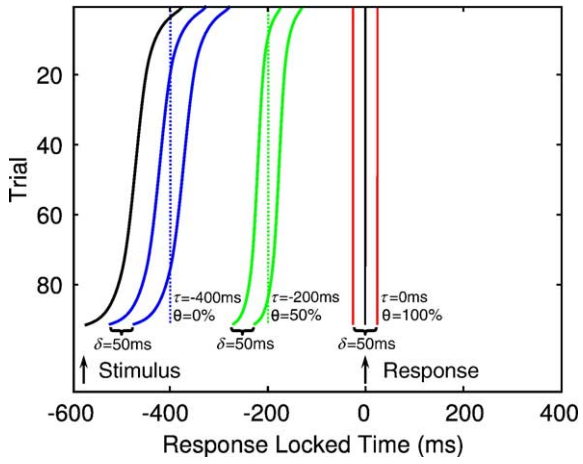

Fig. 2. Parameters of the training window defined for logistic regression (Eq. (3)). We define a training window starting at a relative onset time τ, with a duration of δ and orientation θ (relative to response time). Shown are three examples of training windows with $\tau = -400$ ms, $\theta = 0\%$ (blue sigmoids), $\tau = -200$ ms, $\theta = 50\%$ (green sigmoids) and $\tau = 0$ ms, $\theta = 100\%$ (red vertical lines). The duration of all training windows is $\delta = 50$ ms.

shifted by b from the origin divides the two classes. $\mathbf{w}$ is estimated by minimizing the negative log-likelihood of the data with respect to the model parameters:

$$\mathbf{w}_{\mathrm{LR}} = \underset{\mathbf{w}}{argmin} - L(\mathbf{w}, b) = \underset{\mathbf{w}}{argmin} \left(- \sum_t \log\, p(c_t | y_t) \right) \tag{5}$$

There are no closed form solutions to this optimization problem. However, the maximum can be computed using a fast algorithm based on Iteratively Re-weighted Least Squares (IRLS) (McCullagh and Nelder, 1989). It is a type of Newton–Raphson gradient descent algorithm where the Hessian is given by the Fisher information matrix:

$$\mathbf{w}^{(k+1)} = \mathbf{w}^{(k)} - E\left[\frac{\partial^2 L(\mathbf{w})}{\partial \mathbf{w} \partial \mathbf{w}^T}\right]^{-1} \frac{\partial L(\mathbf{w})}{\partial \mathbf{w}}. \tag{6}$$

A more detailed description of this method and its relationship to other linear classification techniques for EEG can be found in Parra et al. (2005)[2].

The choice of the discrimination time window is crucial as this selects the time frame for which the evoked response difference is extracted. The time window could be chosen at a fixed time after the stimulus onset (stimulus locked), at a fixed time prior to the response (response locked), or at some intermediate form of locking, which we refer to as the orientation of the window (θ) and quantify with, $0\% \leq \theta \leq 100\%$. Specifically the time window for the jth trial is chosen as $t_j \in [T_j - \delta/2,\ T_j + \delta/2]$ with $T_j = \tau + (r_j - \langle r_j \rangle)^* \theta$, where r_j represents the subject response time and $\langle r_j \rangle$ is the mean response time. The duration of the training window was $\delta = 50$ ms and the window onset τ was varied in 10-ms increments to extract the progression of the activity over time. In Fig. 2, the training windows are indicated by the white curves.

[2] Plugins for EEGLAB can be downloaded from http://liinc.bme.columbia.edu/downloads.

Only correctly identified trials were included in the analysis. The 25th image in non-target sequences were used to extract non-target EEG data for analysis. Given our linear model, determination of sensor projections from the discriminating activity is straightforward, namely

$$\mathbf{a} = \frac{\langle \mathbf{x}(t), y(t) \rangle}{\langle y(t), y(t) \rangle}, \tag{7}$$

where $\langle \cdot, \cdot \rangle$ denotes an inner product. Eq. (7) describes the projection $\mathbf{a}$ of the discriminating component $y(t)$ that explains most of the activity $\mathbf{x}(t)$. A strong projection indicates low attenuation. Therefore, the intensity of sensor projections $\mathbf{a}$ indicates proximity of the component to the sensors.

Performance of the linear classifiers was characterized by the area under receiver operator characteristic (ROC) curves referred to as A_z (Green and Swets, 1966). Given the limited number of trials, rather than dividing data into training and validation sets, a leave-one-out procedure was used (Duda et al., 2001). A classifier was trained excluding one trial, then this excluded trial was classified. This procedure was repeated for all trials to determine leave-one-out performance.

Sliding the training window (by varying τ) from stimulus to response results in variability of discrimination which is reflected in A_z. A significance level for A_z was determined using a bootstrapping method, whereby leave-one-out performance was computed after randomizing the trial labels. This randomization procedure was repeated 100 times to produce an A_z randomization distribution. The randomization distribution was assumed to have a Gaussian profile, and the 99th percentile of this distribution corresponds to a significance level of $P = 0.01$.

We search for the most discriminating component for each τ by finding the orientation (θ) of the training window that leads to the largest A_z. Fig. 3 shows discriminating component activities and scalp projections for one subject aligned to both the time of the stimulus (Fig. 3(a)) and time of the response (Fig. 3(b)). As one moves from the top row to the bottom, one sees the temporal progression of the training window (white curves) and the resultant evolution of the discriminating component activities and corresponding scalp projections.

Peak detection and response time correlation

The variability introduced by visual processing stages associated with recognition has yet to be rigorously investigated with EEG. Critical to any such study is single-trial analysis of EEG features. While enhancing signal to noise, averaging across trials introduces biases. For example, comparison of onset latencies with peak latencies of ERP components may result in misinterpretations as to the degree of overlap and interaction between processing stages. Such ERP features obscure inter-trial variability of EEG morphology which when analyzed on a single-trial basis may support alternative hypotheses. Initial attempts to avoid such biases involved sub-ensemble averaging over trials from response time percentiles. Some investigators have derived single-trial peak and amplitude estimates through visual inspection (Ritter et al., 1979) and template matching methods (Kutas et al., 1977). Visual inspection of individual trials from several electrodes is impractical and prone to errors while adaptive template matching methods such as Woody filters (Woody, 1967) are susceptible to errors induced by irrelevant external events.

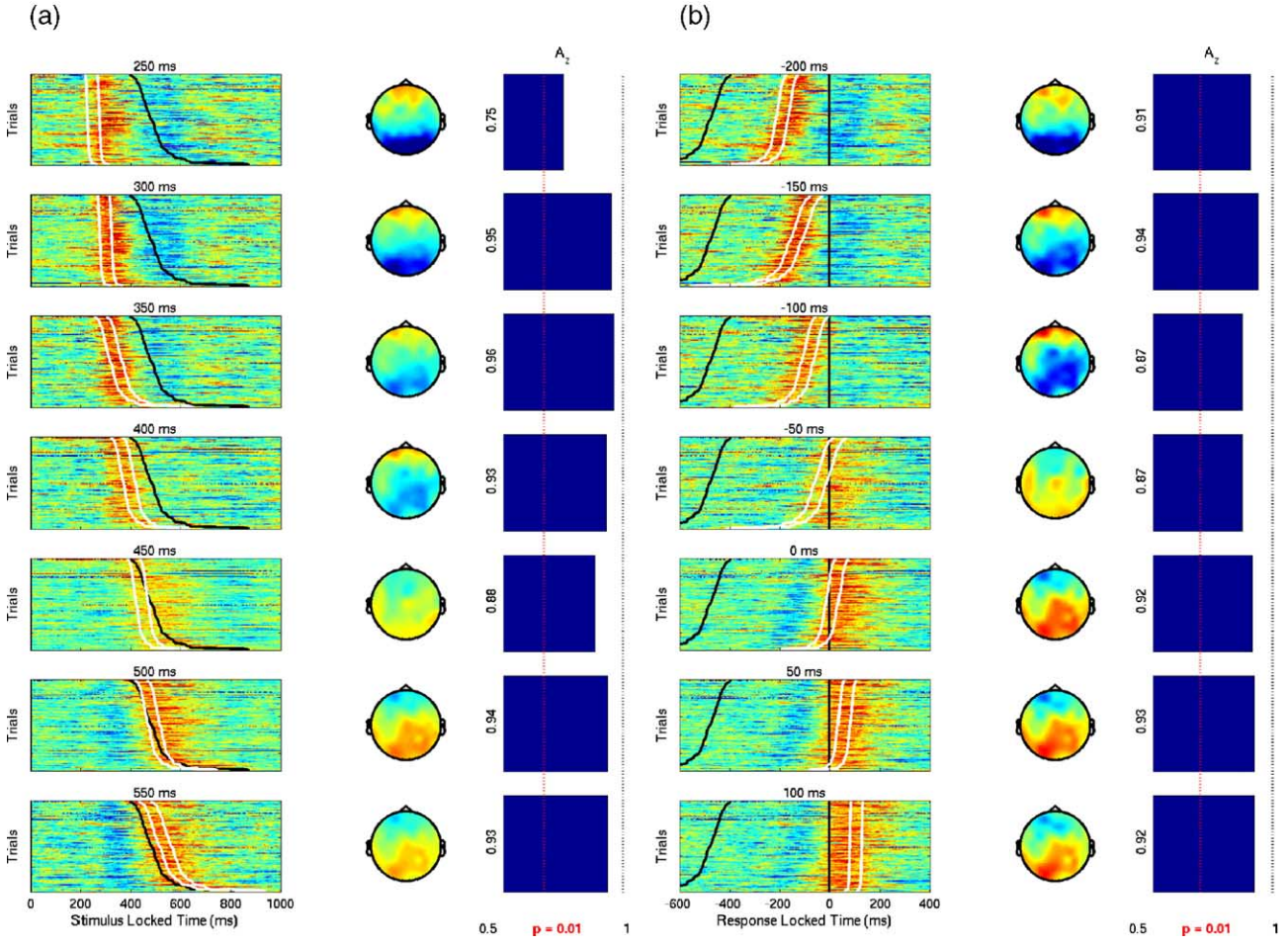

Fig. 3. Discriminating component activity shows the difference between target and non-target trials. Stimulus-locked (a) and response-locked (b) component activity $\mathbf{y}$ (Eq. (3)) for subject 2. The activity is color coded (positive activity is red, negative activity is blue). Trials are sorted by response time and aligned vertically so that stimulus or button release time is at 0 ms for stimulus- or response-locked panels, respectively (vertical black line). The black "S"-shaped curve indicates the response or stimulus for stimulus- or response-locked plots, respectively. Each row shows the activity extracted for a window of $\delta = 50$ ms duration. The beginning and end of this window is displayed with two white curves. The profile (i.e., orientation θ) of this window is derived by scaling response times across trials. The degree of scaling was determined by searching for the scaled training window yielding maximum discrimination performance. Moving the window by varying τ, enables discriminating activity for various latencies to be extracted. A representation of the topology of the extracted activity (Eq. (7)) is shown to the right (dorsal view). The color code indicates (red) positive correlation of the sensor readings with the extracted activity and (blue) negative correlation. These scalp plots can be thought of as a coupling of the discriminating activity with the sensors, reflecting the proximity and orientation of the discriminating activity. The A_z value shown as a bar graph indicates the significance of the differential activity. The red dotted line corresponds to $P = 0.01$.

A maximum likelihood reformulation of adaptive template matching filters demonstrated more accurate estimates (Pham et al., 1987). This approach has been extended to incorporate estimates of single-trial amplitudes (Jaśkowski and Verleger, 1999). Most recently, a Bayesian framework for modeling single-trial multi-component event-related potentials generalizes these filtering approaches to multiple components. This differentially variable component analysis (dVCA) model incorporates estimates of optimal latency shift and scaling factors for each trial and has been used to compare single-trial component latencies with response time for intracortically recorded local field potentials from monkeys performing a Go/No-Go visual discrimination task (Truccolo et al., 2003). However, the method does not estimate peak times, rather lag for maximum correlation between trials, and thus cannot be used to determine the timing of the component. We estimate inter-trial variability by extracting features from discriminating components. The stochastic nature of EEG precludes the robust extraction of component onset from individual trials; however, there is evidence of strong correlation between ERP peak and onset times (Scheffers et al., 1991). The peaks of spatially integrated discriminating components were found by fitting a parametric function to the extracted component $y(t)$. For simplicity, we use a Gaussian profile that is parameterized by its height β, width σ, delay μ, and baseline offset α:

$$\hat{y}(t) = \alpha + \frac{\beta}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}}. \tag{8}$$

The optimal parameters for each trial were found using a nonlinear least-squares Gauss–Newton optimization. The center and width of the discrimination training window was used to initialize optimization. Fig. 4 illustrates the fit to response-locked data.

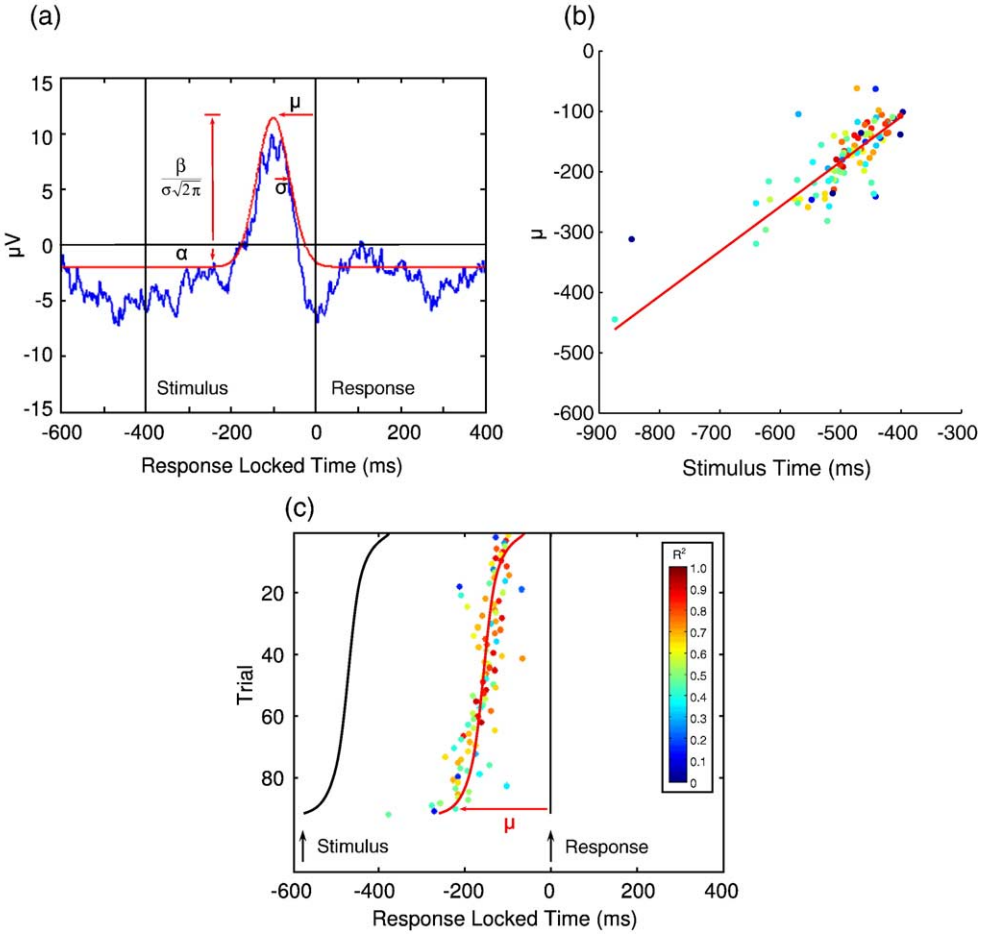

Fig. 4. Illustration of single-trial fitting procedure and estimation of percent response locking. (a) Fit $\hat{y}(t)$ (Eq. (8)) for component activity $y_c(t)$ (Eq. (3)) for one trial. Parameters of single-trial fit: μ corresponds to discriminating component peak latency relative to the response and σ indicates component duration. (b) Robust regression of component latencies μ and response times r as described by Eq. (9) for subject 2, response locked $\tau = -200$ ms. Predicted $\hat{\mu}$ is shown by the red line. In response-locked analysis, stimulus time $(-r)$ is used in the regression rather than response time r. For response-locked analysis, the degree of response locking is then defined as $1 - \phi$ rather than ϕ as in stimulus-locked analysis. Subtracting the slope of the red line ($\hat{\mu}$) from one corresponds to the degree of response locking shown in Figs. 5–8. (c) Displaying the fit from this regression across trials helps to visualize the degree of response locking. Component peak latencies μ across trials are shown by colored dots. The color of the peaks corresponds to the R^2 value indicating the goodness–of–fit for $\hat{y}(t)$. Note that points with a large R^2 tend to lie along the fitted curve. Stimulus onset time (black sigmoid) is fit to peak latencies to quantify degree of component response locking across trials (red sigmoid).

To determine response locking, we compute the linear regression coefficients that predict the latency of the component activity as measured by μ from the response times given by r as described by Eq. (9). The slope from the response time peak latency regression (ϕ) is defined to be the degree of response locking (percentage) for each component. This metric quantifies the extent to which the component is correlated with the response across trials. It ranges from 0% for pure stimulus lock to 100% for pure response lock. A slope $\phi = 100\%$ indicates that slow responses show a corresponding late activity, and fast responses show a corresponding early activity. A slope of $\phi = 0\%$ indicates that the timing of the activity does not change with response time and is therefore stimulus locked. Robust linear regression using iteratively reweighted least-squares (Holland and Welsch, 1977) was used to determine this projection in order to minimize the effects of outliers[3] (see online Supplemental material for animations illustrating these methods). An example of this regression is shown in Fig. 4 for subject 2.

$$\hat{\mu}_j = \phi r_j + b \tag{9}$$

where $\hat{\mu}_j$ and r_j are the predicted peak latencies and response times for the jth trial. $100 \times \phi$ is then the % response lock shown in Figs. 5–8. Robust regression tries to find the optimum ϕ such that $\hat{\mu}_j$ approaches the observed μ_j which has been estimated with the curve fit described by Eq. (8).

In defining this response locking metric, regression is more appropriate than correlation. The goal of regression is to minimize the error of the prediction as the latency μ carries most of the measurement error while the response time can be measured rather reliably. Robust regression in particular accounts for outliers in the variable that is to be predicted, in this case μ. Again, a sensible choice as it is μ that may at times be off altogether (when there was no clear peak to be fitted in the single

[3] For a demonstration of robust regression, please refer to the robustdemo function provided with the Matlab Statistics Toolbox.

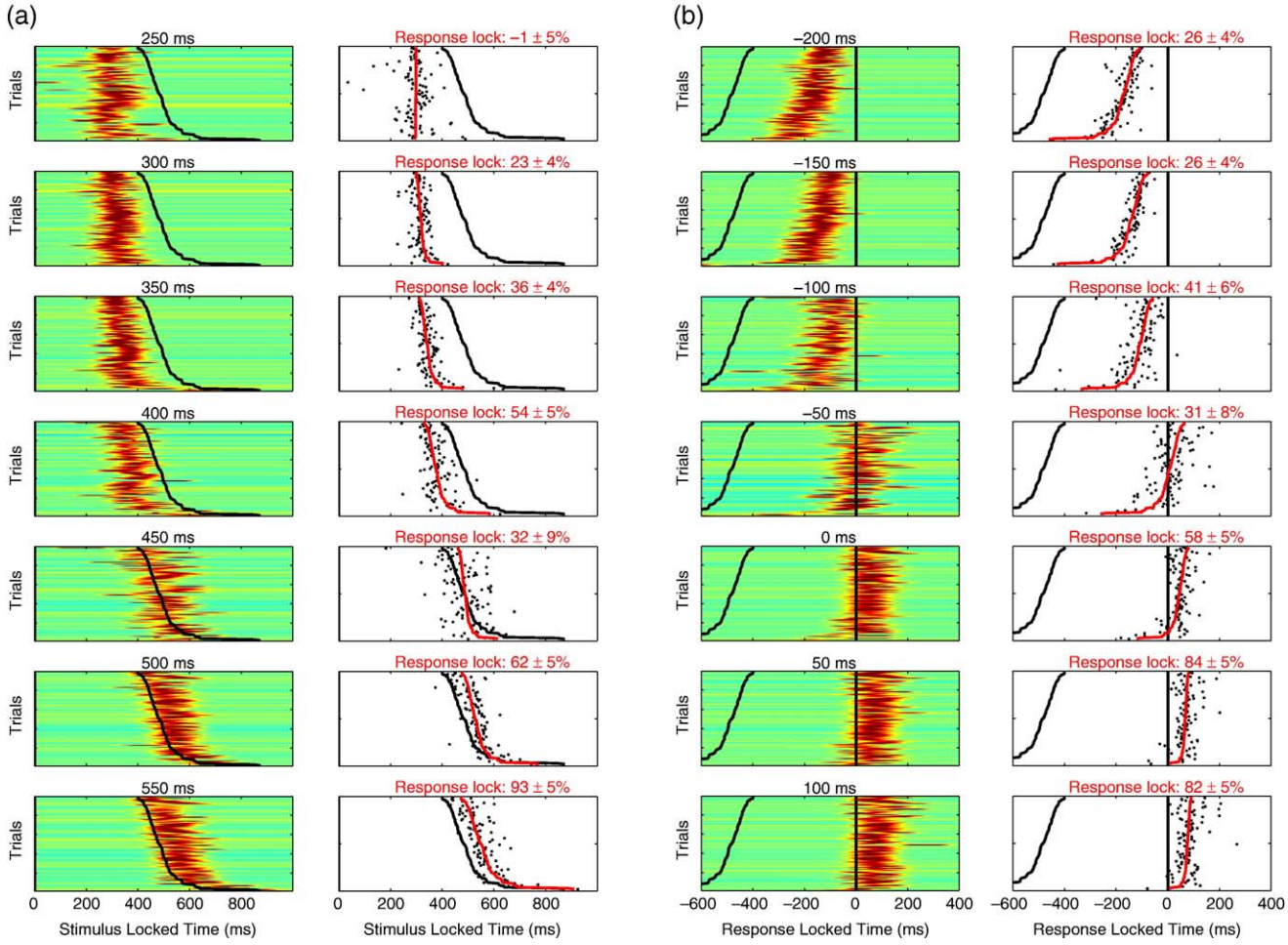

Fig. 5. Detailed temporal analysis of stimulus-locked (a) and response-locked (b) discriminating activity for subject 2. Each row in the left column shows the fit of discriminating activity from Fig. 3 to a Gaussian profile described by Eq. (8). On the top of each of these panels is the onset time of the window used for discrimination. Right columns of each panel display the peak latency (μ) (black dots) of each trial. The projection of response times onto these peak latencies is shown with a red curve (stimulus times were projected onto peak latencies for the response-locked condition). The parameters for this projection indicate the degree of response locking for each component. Purely stimulus- and response-locked conditions are indicated by 0% and 100% response lock, respectively. On top of these panels are reported the percent response lock and corresponding error in the fit of the peak latencies across trials as well as the mean onset time of the component. The standard deviation of peak latencies is 62 ms for stimulus-locked and 66 ms for response-locked results shown.

trial response). One would like to know the trial-to-trial variability within this locking. This variability is captured by the residual error of the regression (unexplained variability of the dependent variable). To put this variability into context, one will naturally divide by the variability of the dependent variable, in this case the variability of r. Incidentally, this ratio is proportional to the standard error of the slope estimate. This standard error is shown along with the slope estimates in Figs. 6 and 7. These values provide an indication of the relative trial-to-trial variability—up to $\sqrt{N}$, where N is the number of trials.

Results

Task performance

In the RSVP task, participants correctly responded to 91.8% of sequences containing targets and 98.3% of distractor only sequences. Response times for correctly identified targets ranged from 342 to 874 ms, mean response time of 500 ± 77 ms. While there are reports of significant correlation between response time and accuracy (Thorpe et al., 1996), our results did not have sufficient significance to support correlation between mean response time and fraction of correct responses to target or distractor only sequences.

Discriminating components

Fig. 3 shows projections, $y(t)$, and scalp maps, **a**, of components significantly discriminating between target and distractor trials for one subject (subject 2). Results are similar across subjects. Visual inspection of the single-trial components resulting from discrimination of stimulus-locked versus response-locked trials provides some intuition as to whether components are correlated with stimulus or response. Based on the bootstrapping results, significant discriminating components ($P < 0.01$) were found beginning 200 ms prior to the response for all subjects. In most cases, components before 200 ms were not significant. This indicates that the neuronal

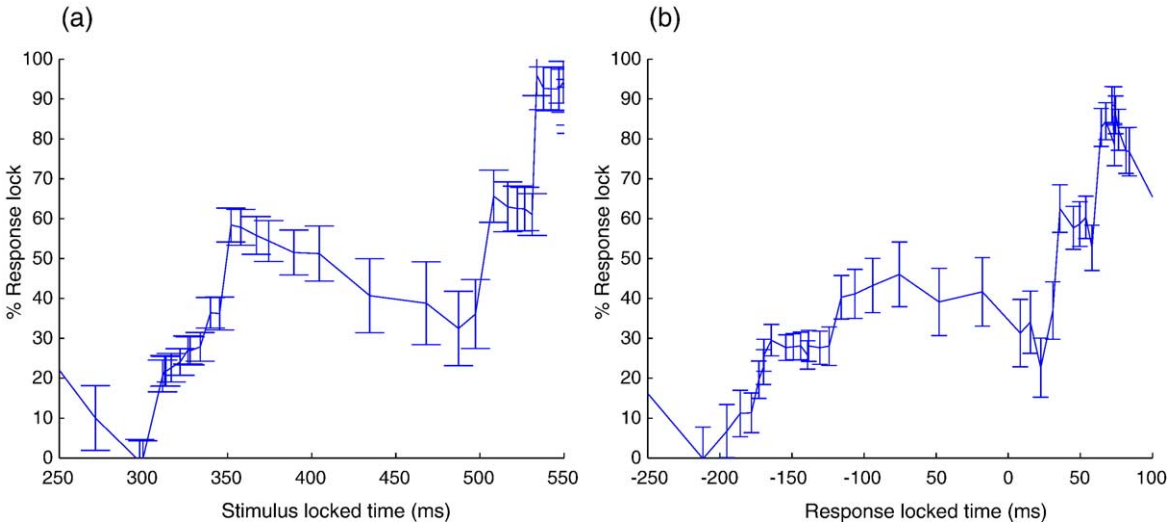

Fig. 6. Degree of response locking over time derived from stimulus-locked (a) and response-locked (b) discriminating components for subject 2. Response lock percentages are from projections shown in Fig. 5. Error bars reflect standard error of the regression parameter associated with response locking %.

processing, measured via EEG, up to this point does not differ between targets and distractors.

The component 200 ms prior to the response corresponds to the activity observed approximately 250 ms after stimulus presentation. The corresponding topological scalp map indicates that this component is positively correlated with bilateral far-frontal regions and negatively correlated with parietal regions. Subsequent discriminating components derived by shifting the training window toward the response time indicate that these components remain positively correlated with far-frontal regions for approximately 100 ms. This frontal activity is followed by positive correlation with posterior parietal regions. Between 200 and 100 ms prior to response, negative correlation of discriminating activity shifts from occipital to parietal regions.

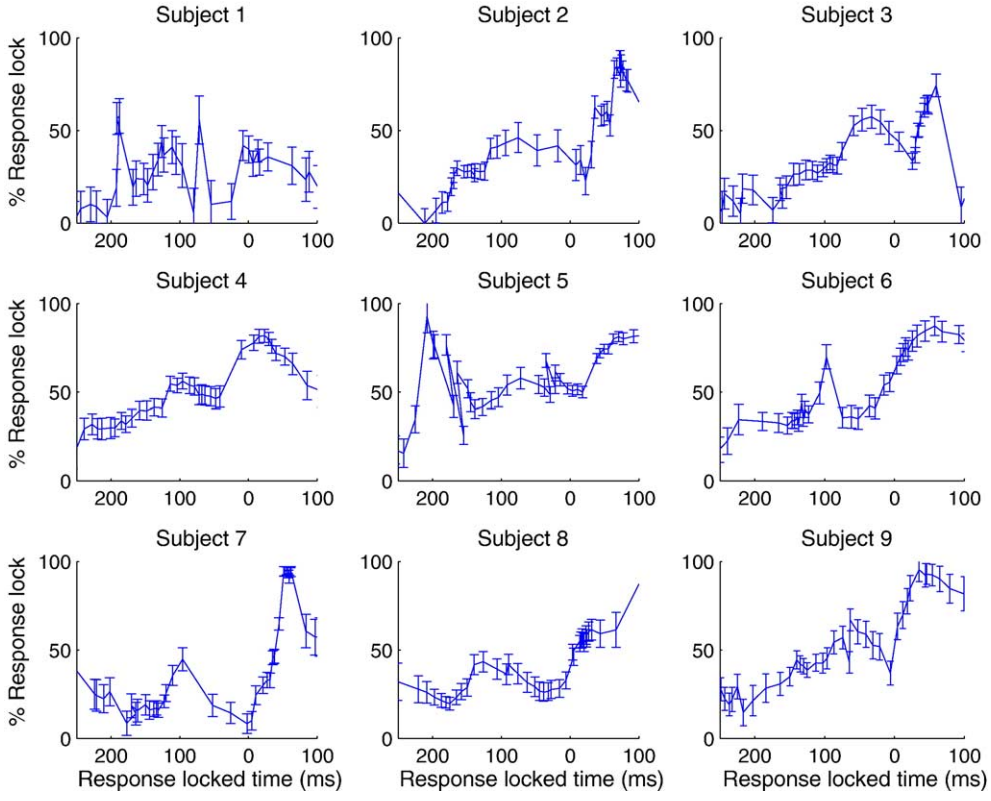

Fig. 7. Degree of response locking for our nine subjects, with data response locked. Seven of the nine subjects show a systematic increase in response locking as the component onset time nears the response time. Subject 1, and to a lesser sense subject 7, does not show this trend.

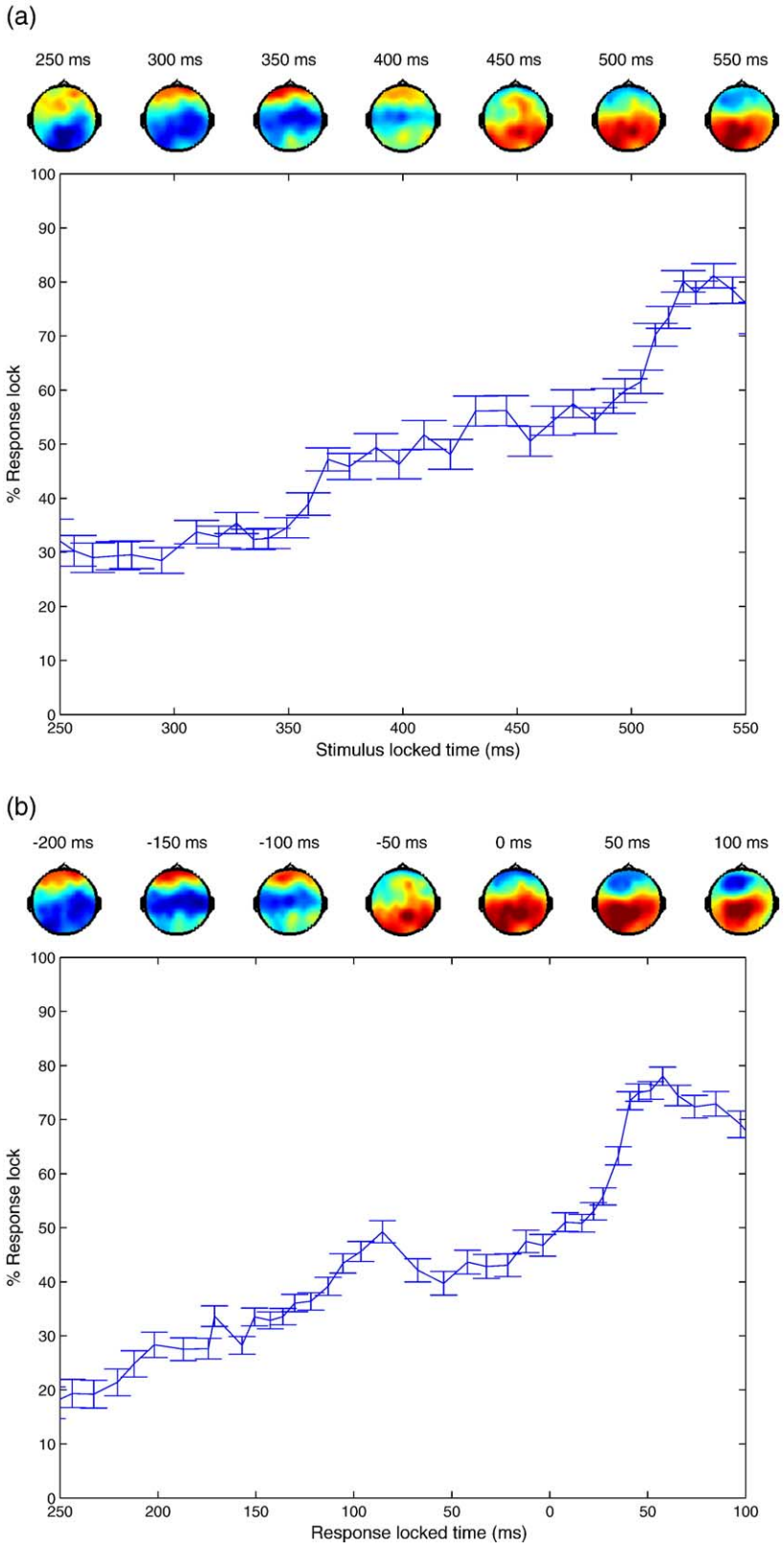

Fig. 8. Group results over all 9 subjects for stimulus-locked (a) and response-locked (b) discriminating components. Top row shows scalp distribution of discriminating activity averaged over all subjects. Bottom row shows the degree of response locking over time. Error bars reflect standard error of the regression parameter associated with response locking %. For all subjects, the first discriminating activity is frontal and correlated more with the stimulus than response. By the time it arrives in parietal areas, a delay has been introduced.

If our estimator is unbiased, then we would expect similar results for stimulus and response-locked conditions. However, tests of our method, using simulated EEG data, demonstrated a small bias largely due to the method for finding the optimal orientation of the training window (θ) and the fact that the training window is incremented in 10-ms steps relative to the stimulus or response so the data selected for analysis are not identical. We show results for both to indicate the degree of bias. Clearly, the major trends are conserved.

Correlation between component latency and response time

Response times were projected onto peak latencies of discriminating components as shown for one subject in Fig. 5. These regression slopes quantify the degree of response locking of each processing stage. Purely stimulus and response-locked conditions are indicated by 0% and 100%, respectively. Note that as the training window onset advances in time (i.e., as τ increases), the discriminating component advances and is more response locked. However, the window onset time and component time slightly differ. In our analysis, we always consider response locking as a function of the component timing.

Fig. 6 summarizes these response locking results for subject 2, as a function of component onset time. Together with the scalp projections in Fig. 3, these results indicate that component activity systematically shifts from stimulus locked to response locked as activity progresses from frontal to parietal regions. Note that though the overall trend is a transition from stimulus to response locking, the curves are not monotonic. This is due to noise in the EEG signal and the resulting errors in estimating single-trial peak latencies. The trends seen for subject 2 are consistent across seven of our nine subjects, as seen by the response-locked results in Fig. 7. The trend is less evident in subjects 1 and 7.

Fig. 8 shows the group results for the discriminating component activity across nine participants. Scalp projections of discriminating components were normalized prior to averaging. Group averaged results are consistent with scalp topology from Fig. 3 showing a shift of activity from frontal to parietal regions over the course of 200 ms. In order to estimate the progression of response locking across all subjects, it is necessary to account for response time variability between subjects. Averaging results is not appropriate since components are not temporally aligned across subjects. Consequently, histograms of response times were equalized to one subject (subject 2), and component peak times were scaled accordingly. Scaled response times and component peak times were concatenated across subjects. These registered group response times were then projected onto the scaled component peak times to estimate the degree of response locking across subjects. The group response lock increases from 28% at − 200 ms to 78% 50 ms after the response.

Discussion

In this study, we used linear discrimination of single-trial high-density EEG to identify components correlated with visual object recognition and then analyze these components in terms of their onset relative to response time. Rather than integrate across trials, logistic regression was used to find an optimal set of weighting coefficients to integrate across sensors. Scalp projections derived from these coefficients describe the proximity of discriminating components to the sensors. Discrimination performance of the classifier defined in terms of ROC analysis was used to gauge the significance of components using a bootstrap method. Area under ROC curves indicates that significantly discriminating activity is observed within 250 ms following stimulus presentation. Discriminating activity was initially observed over frontal regions and progressed to parietal regions around the response time.

The RSVP task requires high vigilance and emulates natural saccadic scene acquisition. The structure of our task, with varied scale, pose, and position of target objects (people) requires subjects to perform object recognition rather than simply recognizing low-level features. The RSVP paradigm has been used to study the response of neurons in higher visual areas such as superior temporal sulcus through invasive recordings (Keysers et al., 2001). A rapid visual presentation task was subsequently used in an event-related potential study and revealed discriminating activity in frontal electrodes about 150 ms following stimulus onset (Thorpe et al., 1996). A more recent EEG RSVP study revealed this early component is likely associated with discrimination of low-level features while a later response-locked component arising about 300 ms following stimulus presentation is most likely associated with object recognition (Johnson and Olshausen, 2003). This study analyzed the significance of differences between task conditions for each electrode to derive estimates of processing time.

The methods we present permit the study of processing stages without manipulating stimulus characteristics or stimulus–response compatibility as described by the additive factors method. Traditionally, investigators of “mental chronometry” have relied on experimental manipulations of response time to characterize mental processing stages (Rugg and Coles, 1995, Meyer et al., 1988). For instance, Donders’ subtraction method varies the complexity of experimental tasks (Donders, 1868/1969). Comparison of response time distributions resulting from the canonical simple, forced-choice and go/no-go response time tasks yields estimates of the duration of mental processes associated with stimulus perception, discrimination, response selection, and motor response. The validity of the subtraction approach rests on the assumptions that the durations of mental processes combine additively and are not influenced by experimental changes. Sternberg’s additive factors method (AFM) circumvents these assumptions by modulating stimulus characteristics such as luminance, coherence, and frequency content (Sternberg, 1969). The effects of stimulus factor manipulation are assumed to be limited to temporal variations in a stimulus encoding stage. Independent of stimulus encoding processes, a response stage is influenced by changes in stimulus–response compatibility induced, for instance, by adding stimulus classes. The AFM is only valid assuming that processing stages are arranged serially without temporal overlap and information transmission between stages is discrete. This is not necessarily the case as indicated by McClelland’s cascade model which proposes that information flows continuously through several parallel processes (McClelland, 1979). The varied-priming and speed–accuracy decomposition methods described by Meyer et al. (1988) were developed to address the parallel–serial and discrete–continuous questions. These methods rely on manipulating the timing and number of stimuli presented during an experiment which may invalidate assumptions governing the model and methods used to analyze results (van der Molen et al., 1991).

While our methods do not rely on experimental manipulations as in AFM to characterize processing stages, the validity of our

Table 1
Cortical areas engaged in visual object recognition during the RSVP task

Cortical region	Occipital	Frontal	Parietal	Motor and somatosensory
Function	Feature extraction and object encoding	Attention and decision making	Memory and decision making	Response preparation and execution
Earliest processing time (ms)	80	150	300	350
Typical processing time (ms)	150	300	450	500
Minimum delay introduced (%)	5	25	40	100
Typical delay introduced (%)	20	70	85	100

Processing times indicate the latency of earliest neuronal response to target image presentation and a typical average latency. Percentages approximate the cumulative amount of delay introduced by successive processing stages. Values were determined from discriminating component activity and the associated response lock percentages from nine subjects. "Function" describes the commonly ascribed functionality of each area.

methods rests on several experimental factors. It is critical that differences within each stimulus class are minimized in order for features observed in EEG trials evoked by a particular stimulus class to describe the same phenomena. In addition, the cognitive state (i.e., attentional resources) of participants must be consistent during presentation of stimuli from each class. We made an effort to mitigate the impact of these requirements by screening stimuli and using a high vigilance task.

Specific to our findings, current models for frontal–parietal lobe interactions induced by target stimuli during an oddball paradigm indicate that anterior cingulate, inferior-temporal lobe, and hippocampus are primarily responsible for P3a and P3b potentials observed on the scalp (Polich, 2003). According to this model, frontal lobe activation associated with P3a reflects allocation of attentional resources. Stimulus induced updating of working memory results in activation of anterior cingulate and communication of representational changes to infero-temporal cortex. Parietal P3b scalp potentials reflect memory operations induced in hippocampus by frontal attention related activity. Such memory updates are subsequently transmitted by hippocampus to parietal cortex. Our current results differ in that the far-frontal activity we report here is more frontal than the conventional anterior P300 (P3a). It also lacks a central component that typically results from anterior cingulate (Makeig et al., 2004). While the scalp projections of our discriminating components do not match P3a and P3b, the timing and frontal–parietal progression of discriminating activity suggest that cortical activity underlying these components may be related. Preliminary fMRI results (not shown) indicate components in frontal and parietal Brodmann's areas 6, 7, 8, and 40. These areas correspond to pre-motor cortex involved with motor planning, parietal cortex associated with visuo-motor integration, frontal eye fields involved with attention, decision making and controlling eye movement, and lateral inferior parietal cortex associated with decision making. These results were not recorded simultaneously with EEG; therefore, it is difficult to determine correspondence with presented results. We are currently conducting simultaneous EEG/fMRI studies to gather additional evidence for spatially localizing these components.

Interestingly, work by Makeig and colleagues (Makeig et al., 1999,2004) using ICA to identify subcomponents of the P3 complex, identifies two components that have nearly identical scalp projections to our far-frontal and parietal components. Makeig labels these P3f (far-frontal) and P3b (parietal). Timing of these components is shown to indicate that P3f precedes the P3b component. Our results indicate that these components found using ICA are in fact also the most discriminating components for the task at their corresponding onset times. In addition, our results indicate that the P3 complex in fact is responsible for introducing the major fraction (at least 60%) of response time variability seen in our visual oddball task. This seems to be further evidence that the P3 complex is composed of distinct components which dynamically interact to form the much studied P3 (P3a and P3b) seen in trial-averaged ERPs.

Since the features of discriminating components are believed to reflect visual processing, attention and decision stages, modeling the peak latency, amplitude and duration of each trial enables the study of covariability of each stage with response time. Consistent with Kammer et al. (1999), Fig. 8 indicates that significant processing delays may be introduced by early processing stages. Within 200 ms prior to response (≈250 ms following stimulus), activity is already, on average, between 25–35% response locked. Due to our method, it is not possible to determine whether this response locking is a result of components at this onset time or earlier onset times, since discriminating components were not significant for earlier onset (peak) times. Thus, we conclude it is possible that some of this early response locking may be due to early visual processes (0–250 ms poststimulus). For our nine subjects, correlation analysis reveals that discriminating component activity progressively becomes more response locked with subsequent processing stages. Along with scalp projections derived from discriminant analysis, the covariability of peak latency with response time describes which cortical regions introduce processing delays. If we assume early visual processing to be mostly occipital, then we can roughly estimate the delay introduced by different processing stages for this task (see Table 1). We conclude that the time course of discriminating component activity provides insight into the nature of information flow through the brain during visual discrimination.

Acknowledgments

This work was funded by the Defense Advanced Research Projects Agency (DARPA) under the Augmented Cognition program the DoD Multidisciplinary University Research Initiative (MURI) administered by the Office of Naval Research (N00014-01-1-0625) and by the NIH (EB004730).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2005.06.026.

References

Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. J. Neurosci. 134, 9–21.

Donders, F.C., 1868/1969. On the speed of mental processes. In: Koster, W.-G. (Ed.), Attention and Performance, vol. II. North-Holland Publishing Company, Amsterdam, pp. 412–413 (translation by W. G. Koster.)

Duda, R., Hart, P., Stork, D., 2001. Pattern Classification. John Wiley and Sons Inc, New York.

Green, D.M., Swets, J.A., 1966. Signal Detection Theory and Psychophysics. John Wiley and Sons Inc, New York.

Holland, P.W., Welsch, R.E., 1977. Robust regression using iteratively reweighted least-squares. Commun. Stat., Theory Methods A6, 813–827.

Hyvärinen, A., Karhunen, J., Oja, E., 2001. Independent Component Analysis. John Wiley and Sons Inc, New York.

Jaśkowski, P., Verleger, R., 1999. Amplitudes and latencies of single-trial ERP's estimated by a maximum-likelihood method. IEEE Trans. Biomed. Eng. 46 (8), 987–993.

Johnson, J.S., Olshausen, B.A., 2003. Timecourse of neural signatures of object recognition. J. Vis. 3 (7), 499–512.

Kammer, T., Lehr, L., Kirschfeld, K., 1999. Cortical visual processing is temporally dispersed by luminance in human subjects. Neurosci. Lett. 263, 133–136.

Keysers, C., Xiao, K.-K., Földiák, P., Perrett, D.I., 2001. The speed of sight. J. Cogn. Neurosci. 13 (1), 90–101.

Kim, S.-G., Richter, W., Ugurbil, K., 1997. Limitations of temporal resolution in functional MRI. Magn. Reson. Med. 37, 631–636.

Kutas, M., McCarthy, G., Donchin, E., 1977. Augmenting mental chronometry: the P300 as a measure of stimulus evaluation time. Science 197 (4305), 792–795.

Makeig, S., Bell, A.J., Jung, T.-P., Sejnowski, T., 1996. Independent component analysis of electroencephalographic data. In: Touretzky, D., Mozer, M., Hasselmo, M. (Eds.), Advances in Neural Information Processing Systems (NIPS 96), vol.8. MIT Press, pp. 145–151.

Makeig, S., Westerfield, M., Jung, T.-P., Covington, J., Townsend, J., Sejnowski, T., Courchesne, E., 1999. Independent components of the late positive response complex in a visual spatial attention task. J. Neurosci. 19, 2665–2680.

Makeig, S., Delorme, A., Westerfield, M., Townsend, J., Courchense, E., Sejnowski, T., 2004. Electroencephalographic brain dynamics following visual targets requiring manual responses. PLoS Biology 2 (6), 747–762 (June).

McCarthy, G., Donchin, E., 1981. A metric for thought: a comparison of P300 latency and reaction time. Science 211, 77–80.

McClelland, J.L., 1979. On the time relations of mental processes: an examination of systems of processes in cascade. Psychol. Rev. 86, 287–330.

McCullagh, P., Nelder, J.A., 1989. Generalized Linear Models, 2nd ed. Chapman and Hall, London.

Meyer, D.E., Osman, A.M., Irwin, D.E., Yantis, S., 1988. Modern mental chronometry. Biol. Psychol. 26, 3–67.

Parra, L., Alvino, C., Tang, A., Pearlmutter, B., Yeung, N., Osman, A., Sajda, P., 2002. Linear spatial integration for single-trial detection in encephalography. NeuroImage 17 (1), 223–230.

Parra, L.C., Spence, C.D., Gerson, A., Sajda, P., 2005. Recipes for the linear analysis of EEG. NeuroImage 28, 326–341 (this issue).

Pham, D.T., Mocks, J., Kohler, W., Gasser, T., 1987. Variable latencies of noisy signals: estimation and testing in brain potential data. Biometrika 74 (3), 525–533.

Polich, J. (Ed.), Detection of Change: Event-Related Potential and fMRI Findings. Kluwer Academic Publishers, Boston.

Ramoser, H., Müller-Gerking, J., Pfurtscheller, G., 2000. Optimal spatial filtering of single trial EEG during imaging hand movement. IEEE Trans. Rehabil. Eng. 8 (4), 441–446 (December).

Ritter, W., Simon, R., Vaughan, H., Friedman, D., 1979. A brain event related to the making of sensory discrimination. Science 203 (4387), 1358–1361.

Roitma, J.D., Shadlen, M.N., 2002. Response of neurons in posterior parietal cortex (area LIP) during a combined reaction-time direction-discrimination task. J. Neurosci. 22, 9475–9489.

Rugg, M.D., Coles, M.G.H. (Eds.), Electrophysiology of Mind. Oxford Univ. Press, New York.

Scheffers, M.K., Johnson, R., Ruchkin, D.S., 1991. P300 in patients with unilateral temporal lobectomies: the effects of reduced stimulus quality. Psychophysiology 28, 274–284.

Sternberg, S., 1969. The discovery of pressing stages: extensions of Donders' method. In: Koster, W.G. (Ed.), Attention and Performance, Vol. II. North-Holland Publishing Company, Amsterdam, pp. 276–315.

Thorpe, S., Fize, D., Marlot, C., 1996. Speed of processing in the human visual system. Nature 381, 520–522.

Toga, A.W., Mazziotta, J.C. (Eds.), Brain Mapping: The Systems. Academic Press, San Diego.

Truccolo, W.A., Knuth, K.H., Shah, A.S., Bressler, S.L., Schroeder, C.E., Ding, M., 2003. Estimation of single-trial multi-component ERPs: differentially variable component analysis. Biol. Cybern. 89, 426–438.

van der Molen, M.W., Bashore, T.R., Halliday, R., Callaway, E., 1991. Chronopsychophysiology: mental chronometry augmented by psychophysiological time markers. In: Jennings, J.R., Coles, M.G.H. (Eds.), Handbook of Cognitive Psychophysiology: Central and Autonomic Nervous System Approaches. Wiley, Chichester, pp. 9–178.

van Hateren, J., van der Schaaf, A., 1998. Independent component filters of natural images compared with simple cells in primary visual cortex. Proc. R. Soc. London, B 265, 359–366.

Woody, C.D., 1967. Characterization of an adaptive filter for the analysis of variable latency neuroelectric signals. Med. Biol. Eng. 5, 539–553.