Behavioral/Sytems/Cognitive

Neural Representation of Task Difficulty and Decision Making during Perceptual Categorization: A Timing Diagram

Marios G. Philiastides,[1] **Roger Ratcliff,**[2] **and Paul Sajda**[1]

[1]Laboratory for Intelligent Imaging and Neural Computing, Department of Biomedical Engineering, Columbia University, New York, New York 10027, and [2]Department of Psychology, Ohio State University, Columbus, Ohio 43210

When does the brain know that a decision is difficult to make? How does decision difficulty affect the allocation of neural resources and timing of constituent cortical processing? Here, we use single-trial analysis of electroencephalography (EEG) to identify neural correlates of decision difficulty and relate these to neural correlates of decision accuracy. Using a cued paradigm, we show that we can identify a component in the EEG that reflects the inherent task difficulty and not simply a correlation with the stimulus. We find that this decision difficulty component arises ≈220 ms after stimulus presentation, between two EEG components that are predictive of decision accuracy [an "early" (170 ms) and a "late" (≈300 ms) component]. We use these results to develop a timing diagram for perceptual decision making and relate the component activities to parameters of a diffusion model for decision making.

Key words: task difficulty; decision making; electroencephalography (EEG); single-trial analysis; diffusion model; timing diagram

Introduction

Perceptual decision making is a complex process engaging sensory processing, attention, evidence accumulation, and motor-response networks (Schall and Hanes, 1993; Kim and Shadlen, 1999; Shadlen and Newsome, 2001; Egner and Hirsch, 2005; Huk and Shadlen, 2005). Experiments on both primates and humans have focused on identifying the neural correlates of decision making and its underlying component processes. For instance, neural correlates of the time-dependent accumulation of stimulus evidence have been localized to the lateral intraparietal area with additional decision-making processing also identified in the frontal eye fields and the dorsal lateral prefrontal cortex. Fronto-parietal networks have also been implicated, particularly with respect to attention and allocation of relevant resources (Posner and Peterson, 1990; Bisley and Goldberg, 2003; Romo et al., 2003; Egner and Hirsch, 2005). Although much of this work has been done in primates using single and multiunit recordings, recent work in humans using functional magnetic resonance imaging (fMRI) (Heekeren et al., 2004) has attempted to address some of the same questions. However, because of the time resolution of fMRI, little can be said about the relative timing of the cortical processes underlying decision making in humans.

Our previous work using single-trial analysis of the electroencephalography (EEG) focused on the issue of the timing of neural components predictive of psychophysical performance during decision making in humans. Specifically, we have shown that for a face-versus-car perceptual categorization task (see Fig. 1A), two stimulus-locked EEG components resulted in neurometric functions indistinguishable from their corresponding psychometric curves (Philiastides and Sajda, 2006), with the timing of these components being 170 ms ("early") and ≥300 ms ("late") relative to stimulus onset. Noteworthy was that although the early component showed little variability in its timing as a function of stimulus evidence (i.e., phase coherence of the image), the late component systematically shifted forward in time as the phase coherence was reduced and the task became more difficult. In addition, for most subjects, the late component was a better match to the corresponding psychometric curve and yielded larger (and more significant) choice probabilities, relative to the early component, suggesting this component was more directly linked to the decision made by the subject.

One interpretation of the shift in time of the late component is that during perceptual decision making, the brain must engage additional resources (e.g., attention networks) and/or prolong processing when a decision is hard (relative to when it is easy) and that this results in the delay of the timing of the late component. Here, we investigate the cortical activity related to decision difficulty and relate it to the two components predictive of psychophysical performance. Once again, we use EEG to uncover components covarying with decision difficulty, analyzing their timing and strength relative to the task. We use these results to construct a timing diagram for perceptual decision making and show that the late component can be interpreted as representing the mean drift rate in a diffusion-based decision model.

Materials and Methods

Subjects. Thirteen subjects (six females; age range, 20–37 years) participated in the study. 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.

Stimuli. We used a set of 12 face (face database; Max Planck Institute for Biological Cybernetics, Tuebingen, Germany) and 12 car grayscale images (image size, 512 × 512 pixels; 8 bits/pixel). All images were equated for spatial frequency, luminance, and contrast. They 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 (Dakin, 2002) technique to generate a set of images characterized by their percentage of phase coherence. For the first experiment, we processed each image to have six different phase coherence values (20, 25, 30, 35, 40, and 45%) (Fig. 1*B*). In addition, for the second experiment, we colorized our images with subtle red and green tones. We performed this adjustment by manipulating the hue (H), saturation (S), and value (V) color space of the original images (red: H = 0.04, S = 0.17 V, unchanged; green: H = 0.34, S = 0.23 V, unchanged). A Dell (Round Rock, TX) Precision 530 Workstation with an nVidia (Santa Clara, CA) Quadro4 900XGL graphics card and E-Prime software (Psychological Software Tools, Pittsburgh, PA) controlled the stimulus display. A liquid crystal display projector (LP130; InFocus, Wilsonville, OR) was used to project the images through radio frequency-shielded window onto a front projection screen. Each image was subtended to 33 × 22° of the visual angle.

Behavioral paradigms. Subjects performed two versions of a simple categorization task. In the first experiment, they had to discriminate between grayscale images of faces and cars. Within a block of trials, face and car images over a range of phase coherences were presented in random order. The range of phase coherence

Figure 1. Schematic representation of the behavioral paradigms used in this study. *A*, In the first experiment, subjects were instructed to fixate on the center of the screen and were subsequently presented, in random order, with a series of different face and car images at one of six phase coherence levels. *B*, A sample face image at the six different phase coherence levels (20, 25, 30, 35, 40, and 45%) used in this experiment. *C*, In the second experiment, colorized images (red and green tones), at 30 and 45% phase coherence, were presented in random order. At each trial, subjects were initially shown one of two possible visual cues that lasted for 400 ms, followed by a 600 ms delay before the presentation of the image. Based on the cue, subjects were required to either discriminate between face and car images or simply categorize the image as red or green. In both experiments, each image was presented for 30 ms, followed by an ISI lasting between 1500 and 2000 ms, during which time subjects were required to make a decision and respond by pressing a button.

levels was chosen to span psychophysical thresholds. All subjects performed nearly perfectly at the highest phase coherence but performed near chance for the lowest one. In the second experiment, colorized face and car trials of 30 and 45% phase coherence were presented in random order. In this version of the experiment, subjects were presented with a visual cue for 400 ms that was followed by a 600 ms delay before the next image presentation. Based on the cue, subjects had to either discriminate face versus car or the color of the image (i.e., red vs green). Subjects reported their decisions by pressing one of two mouse buttons, left for faces (and red) and right for cars (and green), using their right index and middle fingers, respectively. All images were presented for 30 ms, followed by an interstimulus interval (ISI) that was randomized in the range of 1500–2000 ms. Subjects were instructed to respond as soon as they formed a decision and before the next image was presented. In both experiments, a total of 50 trials per behavioral condition were presented (i.e., an overall of 600 trials for each experiment). Schematic representations of the two behavioral paradigms are given in Figure 1, *A* and *C*, respectively. Trials in which subjects failed to respond within the ISI were marked as no-choice trials and were discarded from additional analysis.

Data acquisition. EEG data were acquired in an electrostatically shielded room (ETS-Lindgren, Glendale Heights, IL) using a Sensorium (Charlotte, VT) EPA-6 Electrophysiological Amplifier from 60 Ag/AgCl scalp electrodes 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 of <15 kΩ, and the chin electrode was used as ground. Data were sampled at 1000 Hz with an analog pass band of 0.01–300 Hz using 12 dB/octave high-pass and eighth-order elliptic low-pass filters. Subsequently, a software-based 0.5 Hz high-pass filter was used to remove DC drifts, and 60 and 120 Hz (harmonic) notch filters were applied to minimize line-noise artifacts. These filters were

designed to be linear phase to minimize delay distortions. Motor response and stimulus events recorded on separate channels were delayed to match latencies introduced by digitally filtering the EEG.

Movement artifact removal. Before the main experiment, subjects completed an eye movement calibration experiment during which they were instructed to blink repeatedly on the appearance of a white-on-black fixation cross and then to make several horizontal and vertical saccades according to the position of the fixation cross subtending 1 × 1° of the visual field. Horizontal saccades subtended 33°, and vertical saccades subtended 22°. The timing of these visual cues was recorded with EEG. This enabled us to determine linear components associated with eye blinks and saccades (using principal component analysis) that were subsequently projected out of the EEG recorded during the main experiment (Parra et al., 2003). Trials with strong eye movements or other movement artifacts were manually removed by inspection.

Data analysis. We used a single-trial analysis of the EEG to discriminate between any two given experimental conditions (i.e., face vs car or easy vs hard). Logistic regression was used to find an optimal basis for discriminating between the two conditions over a specific temporal window (Parra et al., 2002, 2005). Specifically, we defined a training window starting at a poststimulus onset time τ, with a duration of δ, and used logistic regression to estimate a spatial weighting vector $\mathbf{w}_{\tau,\delta}$, which maximally discriminates between sensor array signals X for the two conditions as follows:

$$y = \mathbf{w}_{\tau,\delta}^T X, \tag{1}$$

in which X is an $N \times T$ matrix (N sensors and T time samples). The result is a "discriminating component" y, which is specific to activity correlated with one condition while minimizing activity correlated with both task

conditions such as early visual processing. 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. For our experiments, the duration of the training window (δ) was 60 ms and the window onset time (τ) was varied across time. We used the reweighted least-squares algorithm to learn the optimal discriminating spatial weighting vector $\mathbf{w}_{\tau,\delta}$ (Jordan and Jacobs, 1994).

The discrimination vector $\mathbf{w}_{\tau,\delta}$ can be seen as the orientation (or direction) in the space of the EEG sensors that maximally discriminates between the two experimental conditions. Thus, the time dimension defines the time of a window (relative to the either the stimulus or response) used to compute this discrimination vector. Given a fixed window width (60 ms in this case), sweeping the training window from the onset of the visual stimulation to the earliest response time represents the evolution of the discrimination vector across time. Within a window, at a fixed time, all samples are treated as independent and identically distributed to train the discriminator. Once the discriminator is trained, it is applied across all time so as to visualize the projection of the trials onto that specific orientation in EEG sensor space. A discriminating component is defined as one such discrimination vector, with its activity visualized by projecting the data across all time onto that orientation. We call this visualization a discriminant component map. For instance, for recurring components, one would expect activity trained during one window time to also be present at another time.

To visualize the profile of these components (stimulus or response locked) across all trials, we constructed discriminant component maps. We aligned all trials of an experimental condition of interest to the onset of visual stimulation and sorted them by their corresponding reaction times (RTs). Therefore, each row of one such discriminant component map represents a single trial across time [i.e., $y_i(t)$]. The discriminant component maps used in this study (see Fig. 7) represent face trials with the mean of the car trials subtracted (i.e., $y_{\text{faces}} - \bar{y}_{\text{car}}$).

To provide a functional neuroanatomical interpretation of the resultant discriminating activity, and given the linearity of our model, we computed the electrical coupling coefficients for the linear model as follows:

$$a = \frac{Xy}{y^T y}. \qquad (2)$$

Equation 2 describes the electrical coupling a of the discriminating component y that explains most of the activity X. To compute these coefficients, y is computed for only times during the specific window used to calculate the weights for that component. Strong coupling indicates low attenuation of the component and can be visualized as the intensity of the "sensor projections" a. a can also be seen as a forward model of the discriminating component activity (Parra et al., 2002, 2005).

We quantified the performance of the linear discriminator by the area under the receiver operator characteristic (ROC) curve, referred to as A_z, with a leave-one-out approach (Duda et al., 2001). We used the ROC A_z metric to characterize the discrimination performance while sliding our training window from stimulus onset to response time (varying τ). Finally, to assess the significance of the resultant discriminating component, we used a bootstrapping technique to compute an A_z value leading to a significance level of $p = 0.01$. Specifically, we computed a significance level for A_z by performing the leave-one-out test after randomizing the truth labels of our face and car trials. We repeated this randomization process 100 times to produce an A_z randomization distribution and compute the A_z leading to a significance level of $p = 0.01$.

Traditional event-related potential (ERP) analysis was also performed by aligning the data to paradigm-specific events and averaging across trials as well as across subjects where appropriate. When ERP activity was used for additional analysis (e.g., ERP amplitude correlation with other experimental parameters), we averaged activity across short-length temporal windows (typically 40 ms in width) to make our estimates more robust. To visualize the spacial extent of the ERP activity across time, we computed average ERP scalp maps by interpolating the ERP activity across all electrode locations. We used a biharmonic spline interpolation (Sandwell, 1987) that is designed for irregularly spaced data points. All scalp maps were plotted using EEGLAB (Delorme and Makeig, 2004).
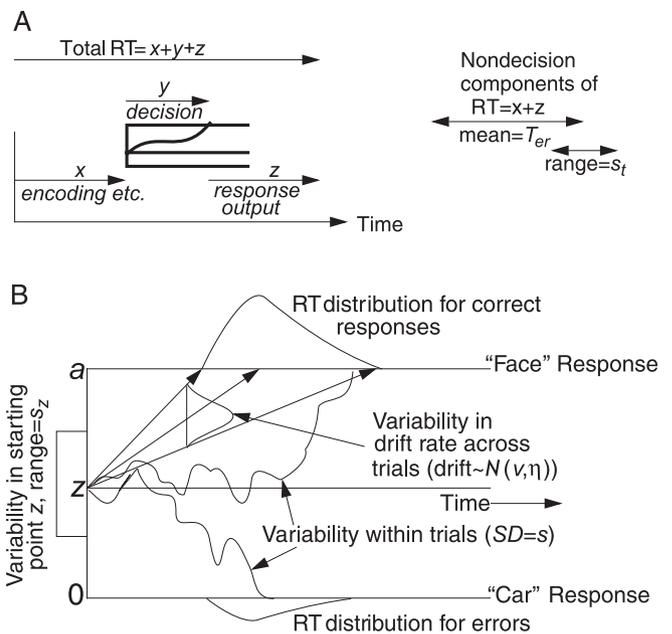
Figure 2. An illustration of the diffusion model. Parameters of the model are as follows: a, boundary separation; z, starting point; T_{er}, mean value of the nondecision component of RT; η, SD in drift across trials; s_z, range of the distribution of starting point (z) across trials; v, drift rate; s_t, range of the distribution of nondecision times across trials; s, SD in variability in drift within trials. **A**, Encoding processes (x) and response output processes (y) combine to give the nondecision component (z) with mean T_{er}. **B**, The diffusion process with two sample paths, RT distributions for correct and error responses, and all of the relevant parameters outlined above.

Discriminant component peak detection. To quantify the spread/duration of a discriminant component, we performed single-trial peak detection by fitting a parametric function to the spatially integrated discriminating component $y(t)$. For simplicity, we used a Gaussian profile [as in the study by Gerson et al. (2005)] that is parameterized by its height β, width σ, delay μ, and baseline offset α as follows:

$$\hat{y}(t) = \alpha + \frac{\beta}{\sigma \sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}}. \qquad (3)$$

We computed the optimal parameters for each trial by using a nonlinear least-squares Gauss–Newton optimization (Gill et al., 1981). The center and width of the discriminating training window was used to initialize the optimization.

Diffusion model simulations. The diffusion model assumes that fast two-choice decisions are made by a noisy process that accumulates information over time from a starting point toward one of two response criteria or boundaries, as in Figure 2B. The starting point is labeled z, and the boundaries are labeled "a" and "0." When one of the boundaries is reached, a response is initiated. The rate of accumulation of information is called drift rate v, and it is determined by the quality of the information available from the stimulus. The better the information quality, the larger the drift rate toward the appropriate decision boundary and the faster and more accurate the response. Within-trial variability in the accumulation of information results in processes with the same mean drift rate terminating at different times (producing RT distributions) and sometimes at different boundaries (producing errors). Speed–accuracy tradeoffs are modulated by the positions of the boundaries as follows: moving boundaries closer to the starting point speeds responses and decreases accuracy. Response time distributions in two-choice tasks are positively skewed, which occurs naturally in the model by simple geometry: the increase in RT is larger if a lower value of drift rate is decreased by some amount than if a larger value of drift rate is decreased by the same amount. Besides the decision process, there are nondecision components of processing such as encoding and response execution (Fig.

2*A*). These processes are combined in the model, and their contribution to RT has mean T_{er} (Ratcliff and Tuerlinckx, 2002).

In the diffusion model, components of processing are assumed to vary from trial to trial. Variability in drift rate across trials (normally distributed with SD η) gives rise to error responses that are relatively slow compared with correct responses, and variability in starting point across trials (uniformly distributed with range s_z) gives rise to relatively fast errors. Whether errors are faster or slower than correct responses for an experimental condition depends on the relative amounts of drift rate and starting point variability, drift rate values, and boundary positions (Ratcliff and Rouder, 1998; Ratcliff et al., 1999). Across-trial variability in T_{er} is uniformly distributed with range s_t (Ratcliff and Tuerlinckx, 2002).

The diffusion model serves to map performance characteristics onto underlying processes. From the probability of a correct response and the RT distributions for correct and error responses for each of the experimental conditions, the model extracts estimates of the quality of the stimulus information that enters the decision process for each condition (drift rate), the amount of information that must be accumulated before a decision can be made (boundary positions), the time taken by nondecision components of RT (T_{er}), and the amount of variability across trials in each of the processing components.

Here we fit the model to the data from the six subjects in the first experiment in which a range of experimental conditions (i.e., different phase coherence levels) were available. We used a χ^2 method (Ratcliff and Tuerlinckx, 2002) to perform the fits. We used the simulation results to relate different parameters of the diffusion model to our experimental observations.

Results

To identify the components related to task difficulty and perceptual decision making, we measured the psychophysical performance of 13 subjects on two separate versions of a simple categorization task (Fig. 1) while simultaneously recording neuronal activity using a high-density EEG electrode array. We changed the stimulus evidence by manipulating the phase coherence (Dakin, 2002) of our images (Fig. 1*B*). During the first experiment (six subjects), face and car images were presented in random order over a range of phase coherences, and subjects were asked to report their decision regarding the type of image by pressing a button (Fig. 1*A*). All subjects performed nearly perfectly at the highest phase coherence but performed near chance at the lowest coherence.

Early and late face-selective components

We compared the EEG activity obtained for the two types of images at each phase coherence level on a single-trial basis, using a linear discriminator that integrates EEG over space rather than across trials (see Materials and Methods) and searched for components that maximally discriminated between the two experimental conditions. We used ROC analysis to quantify the discriminator's performance. Furthermore, taking advantage of the linearity of our model, we computed sensor projections of the discriminating components activity as a means for interpreting the neuroanatomical significance of the resultant discriminating components.

In the interval between the onset of the visual stimulation and the earliest reaction time, we identified two maximally discriminating face-selective components. The early component was consistent with the well known N170 (Jeffreys, 1996; Halgren et al., 2000; Liu et al., 2000, 2002; VanRullen and Thorpe, 2001; Rossion et al., 2003), and its temporal onset appeared to be consistent across all subjects. The late component appeared at least 130 ms after the first, and its temporal onset varied across subjects and across coherences in the range between 300 and 450 ms from the stimulus onset (Philiastides and Sajda, 2006). To characterize the

neuronal performance at these two times, in a manner compatible to the description of the psychophysical sensitivity as captured by the psychometric functions (Green and Swets, 1966), we constructed neurometric functions by plotting the area under the ROC curves (A_z values) against the corresponding phase coherence levels (Britten et al., 1992). We showed that for all subjects in our dataset, these neurometric functions were statistically indistinguishable from their corresponding psychometric functions. In addition, the late face-selective component resulted in a better match to the psychophysical data than the early one (Philiastides and Sajda, 2006).

We also investigated the relationship between the temporal onset of the early and the late face-selective components and task difficulty. We found that for the early component (N170), there was no significant shift in time, whereas the optimal onset of the late component had a systematic forward shift in time as the task became progressively more difficult (Philiastides and Sajda, 2006). This observation raised the interesting possibility that there is yet another component that is more closely associated with task difficulty.

Task-difficulty component

We computed average ERPs at each of six different phase coherence levels. Figure 3, *A* and *B*, illustrates the results for two sensors of interest (FCz and PO8, respectively). To construct these ERPs, we averaged across all subjects and across both face and car trials and found a component ≈220 ms after stimulus, the amplitude of which is inversely proportional to the strength of the stimulus as follows: the lower the coherence (harder task), the greater the amplitude of the D220 component ("D" for "difficulty"). In addition, we constructed average ERP scalp maps to visualize the spatial distribution of this component. To compute these maps, we considered ERP activity centered at the peak of the D220 component. Figure 3*C* demonstrates these results. We found that this component is distributed over a range of electrode locations as indicated by the increased activations at several centrofrontal (high negative activations) and occipitoparietal (high positive activations) sites in the lower coherences. It is also clear from Figure 3*C* that the magnitude of the effect deteriorates as the percentage of phase coherence increases (i.e., as the task becomes easier).

To determine whether this neural signature at 220 ms after stimulus is indeed associated with task difficulty, as captured by our subjects' behavioral performance, we defined a metric we termed "difficulty index" (*DI*) as follows:

$$DI_j = (P_{max} - P_j)/(P_{max} - P_{min}), \qquad (4)$$

where P_{min} and P_{max} are the subject's proportion correct at the lowest and highest phase coherence levels, respectively, and P_j is the subject's behavioral performance at a given phase coherence level. Note that our subjects performed nearly perfectly at the highest coherence and near chance at the lowest one. We subsequently used our single-trial analysis techniques (Parra et al., 2002, 2005) to discriminate between the different difficulty levels. Unlike our previous use of the single-trial classifier (i.e., to discriminate between face and car trials), we now pool data across both face and car trials and discriminate between the highest and each progressively lower coherence set of trials.

During each comparison, we applied our classifier at several time intervals and each time computed an A_z value as a metric for the classifier's performance. We finally correlated these A_z values with the corresponding *DI*s. We found a significant correlation
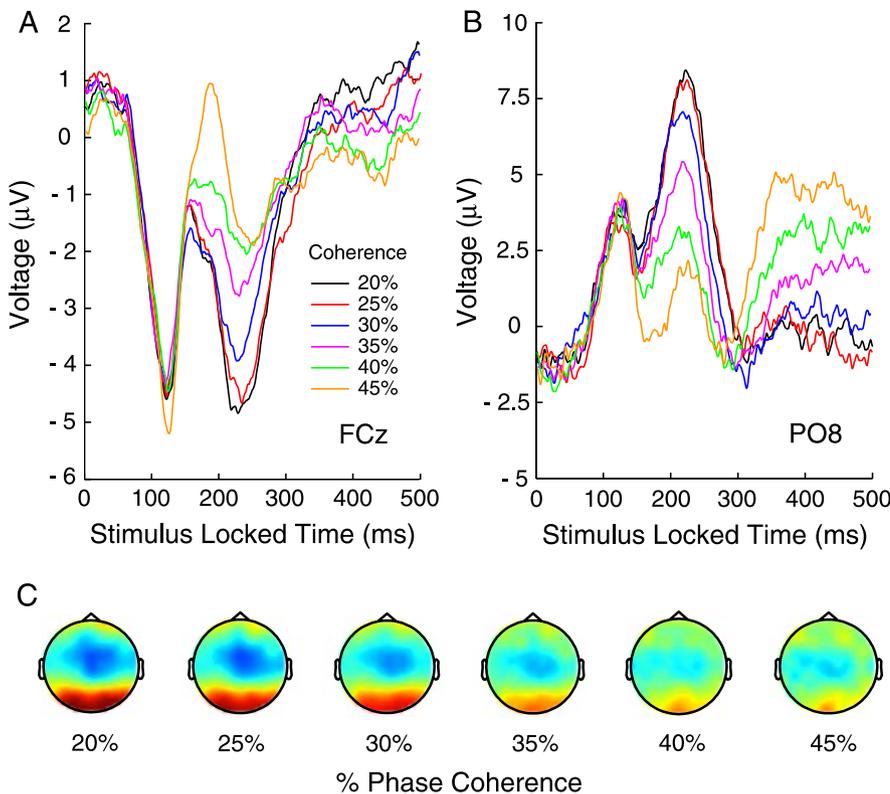
Figure 3. Average ERP analysis of the original face-versus-car categorization task. ***A***, Average ERPs at each phase coherence level, across all subjects and across both face and car trials, for an electrode at a centrofrontal site (FCz). A high negativity, 220 ms after the onset of visual stimulation, can be seen for hard trials (i.e., low phase coherence trials). The amplitude of this negativity is reduced as the task becomes progressively easier. ***B***, Average ERPs for an occipitoparietal electrode location (PO8). Although the effect at 220 ms remains unchanged, its orientation has changed sign (positive activation). ***C***, Average ERP scalp maps at the peak of the D220 component at all phase coherence levels. Red indicates high positive activation, and blue indicates high negative activation.

($p < 0.01$) at ≈220 ms after the onset of the stimulus, as can be seen in Figure 4*A*. The correlation between the A_z and *DI*s that yielded the highest correlation coefficient is shown in Figure 4*B*. Note that although there was also a pronounced separation in the average ERPs at an earlier time representing the N170 component (Fig. 3*A*,*B*), the correlation analysis did not yield a significant observation at that time interval (Fig. 4*B*, inset).

Difficulty component predicts onset of late component

As mentioned previously, the temporal onset of our late face-selective component systematically shifts forward in time as the task becomes more difficult. We wanted to investigate the potential relationship of this temporal shift and the strength of the D220 component. Therefore, we performed an ERP amplitude correlation of this component with the onset of the late one. To emphasize the magnitude of the effect, we used data from several occipitoparietal and centrofrontal electrode sites for this analysis. We identified these sites based on the ERP activation maps in Figure 3*C*. Specifically, we averaged activity across these sensors around a 40 ms window centered at the peak of the D220 component on a subject-by-subject basis and for each different phase coherence level. We subsequently correlated this activity with the onset of each subject's late component at these different coherences. Figure 5 summarizes the results of this analysis and demonstrates that the D220 component can, in fact, predict the temporal onset of the late face-selective component.

Top-down influence on difficulty component

To investigate the possibility that the D220 component represents a true top-down influence on decision making rather than a mere bottom-up processing of the stimulus, we designed a modified version of our original face-versus-car categorization task. We used the same set of images as in the original paradigm, with the only exception that we now colorized the images with subtle red and green tones. At any given trial, subjects were initially presented with one of two possible visual cues. The cues indicated whether subjects would perform the original face-versus-car categorization task or simply discriminate the color of the image (i.e., red vs green) that was presented shortly thereafter (Fig. 1*C*). There were seven new participants for this version of the experiment. The ultimate goal of this new design was to manipulate the task/decision difficulty while keeping the stimulus evidence unchanged (for instance, for the same low coherence stimulus, we could either make the task "hard" by cueing the subjects to perform the face-vs-car categorization task or "easy" by cueing them to perform the color discrimination task).

We used only two phase coherence levels for this study; a high one (45%), for which subjects performed nearly perfectly at the face-versus-car discrimination, and a low one (30%) that caused a significant reduction in behavioral performance. For each cued condition, we sorted the trials based on the strength of the stimulus evidence (i.e., low-vs-high-coherence trials) and constructed average ERPs (Fig. 6). In the face-versus-car cueing condition, we found, just like in the original experiment, a significant separation ≈220 ms after the onset of visual stimulation between hard (low-coherence) and easy (high-coherence) decisions (mean performances of 65 and 95% correct, respectively). These results can be seen in Figure 6*A*. In the red-versus-green cueing condition, for which subjects performed nearly perfectly for both the low- and high-coherence trials (mean performances of 96 and 96% correct, respectively), the effect of this component was eliminated (Fig. 6*B*).

To visualize the spatial extent of this effect, we devised a bootstrap test that allowed us to assess the significance of the difference between low- and high-coherence trials for the two tasks across all electrode locations. For each task, we permuted the trial labels so that the original relationship between low- and high-coherence trials was abolished. We then computed the difference in ERP amplitudes among the newly labeled trials around the D220 component. We repeated this test 5000 times to construct a difference ERP distribution. We finally checked whether the true ERP amplitude difference was outside of the 99% confidence interval of this distribution, in which case we concluded that the difference was statistically significant. For the face-versus-car discrimination task, significant effects can be seen in virtually all electrode locations (Fig. 6*C*). Interestingly, the effect was not nearly as pronounced when subjects were cued to discriminate

A

B
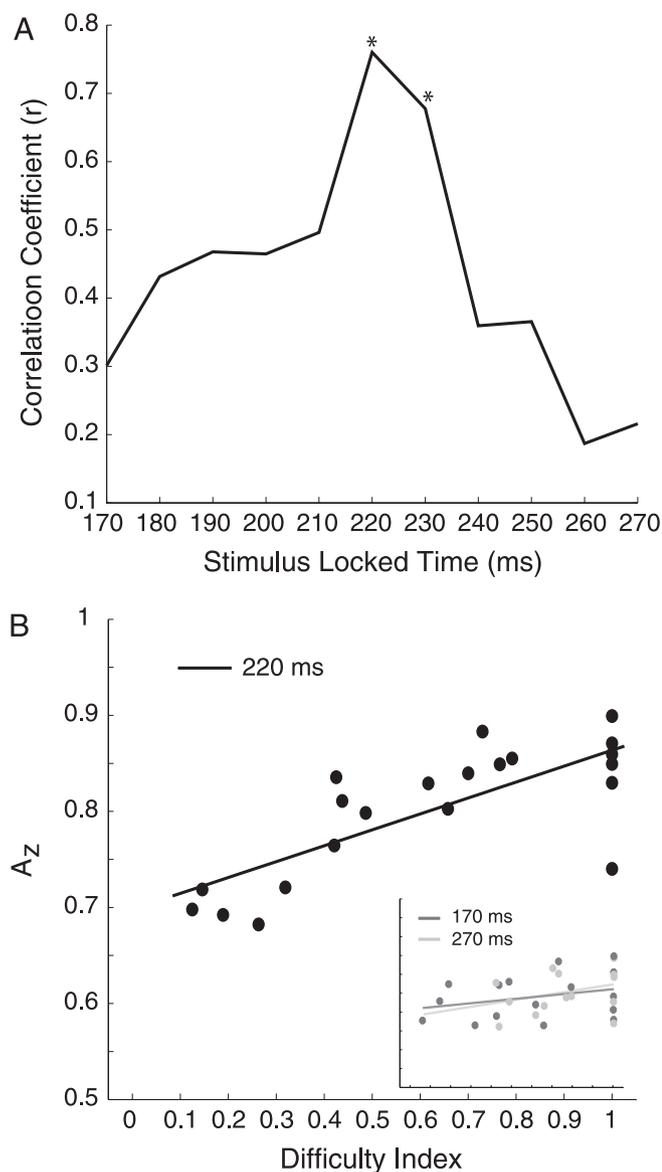
Figure 4. Correlation analysis between the single-trial performance of our classifier (A_z values) for easy-vs-hard discrimination and DI. **A**, The correlation coefficients are plotted as a function of time. A significant correlation between A_z and DI can be seen in the time interval ≈220 ms after stimulus (asterisks). **B**, The actual correlation results for the highest point of the correlation analysis (at 220 ms). Inset, Correlation analysis for two other time points (170 and 270 ms). Both the number of significant A_z values and the correlation itself is significantly reduced. Specifically, the number of significant A_zs is reduced by 30 and 47%, respectively, for the two time windows, respectively.

color instead (Fig. 6D). In addition, we used this permutation test to check which electrode locations demonstrated a significant difference between the difference ERP amplitudes across the two tasks. Not surprisingly, a distributed network that includes several centrofrontal and occipitoparietal regions was revealed (Fig. 6E).

Decision-related late component

Using the color version of our behavioral paradigm, we came across another interesting finding that provides more evidence that our late component, which discriminates highly between face and car trials, is related to a postsensory/decision event rather than an early visual perception/detection event. When we used our single-trial discrimination analysis to classify between
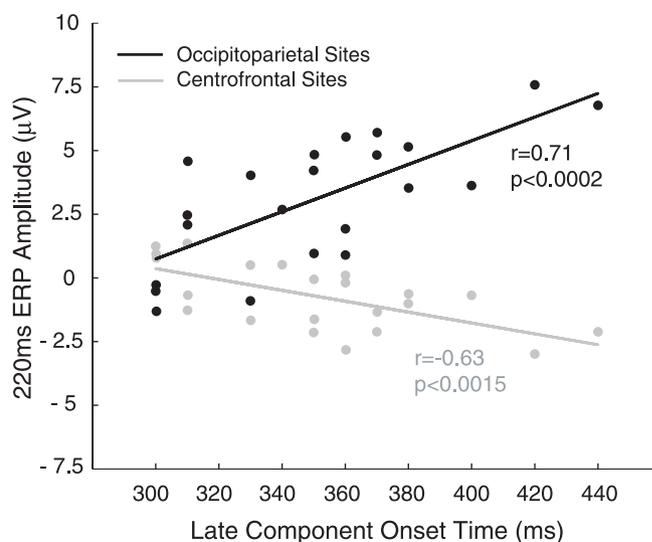
Figure 5. ERP amplitude correlation of the D220 component and the onset of the late face-selective component. The ERP amplitudes of the D220 component (averaged across a 40 ms window centered at the peak of the component) from six sensors at occipitoparietal sites (black) and six sensors at centrofrontal locations (gray) were correlated with the onset of every subject's late component at each different phase coherence level. In both cases, the result was a statistically significant correlation ($r = 0.71$, $p < 0.0002$ and $r = -0.63$, $p < 0.0015$, respectively), indicating that the strength of the D220 component can predict the onset of the late face-selective component. Note that the negative correlation results from the negativity seen ≈220 ms at centrofrontal locations (see Fig. 3A).

face and car images for trials in which subjects were cued to merely discriminate the color of an image, we found that the early face-selective component remained present, whereas the late component was significantly reduced. Figure 7 shows single-trial data for one subject at the 45% phase coherence level. The top row shows that the early and the late components are both present when the subject performed the face-versus-car discrimination task, in line with our original findings. The bottom row demonstrates the extinction of the late component when the discriminator is classifying between face and car trials while the subject is categorizing images based on their color.

We performed a group analysis to quantify the significance in the reduction of the late component. Our subjects' performance was comparable at both the high coherence (45% phase coherence) face-versus-car and color discrimination tasks with 95 and 96% correct, respectively (two-tailed paired t test, $p > 0.25$). Their mean RTs on the two tasks were 667 and 669 ms, respectively, for which they were statistically indistinguishable (two-tailed, paired t test, $p > 0.89$). Our group analysis revealed that when we discriminated between face and car trials, the early component (N170) was present in both tasks. The mean A_z values in each case were 0.771 and 0.773, and their corresponding distributions did not differ significantly from one another (two-tailed, paired t test, $p > 0.94$). For the late component, however, there was a significant reduction in the classifier's performance during the color discrimination task (mean A_z, 0.675) compared with the face-versus-car discrimination task (mean A_z, 0.905) (left-tail paired t test, $p < 0.0006$). These findings are summarized in Figure 8.

Evidence changes the temporal spread of late component

To quantify the spread/duration of the late component, we performed single-trial peak detection by fitting a parametric function to the spatially integrated discriminating component $y_{late}(t)$ (see Materials and Methods for details).
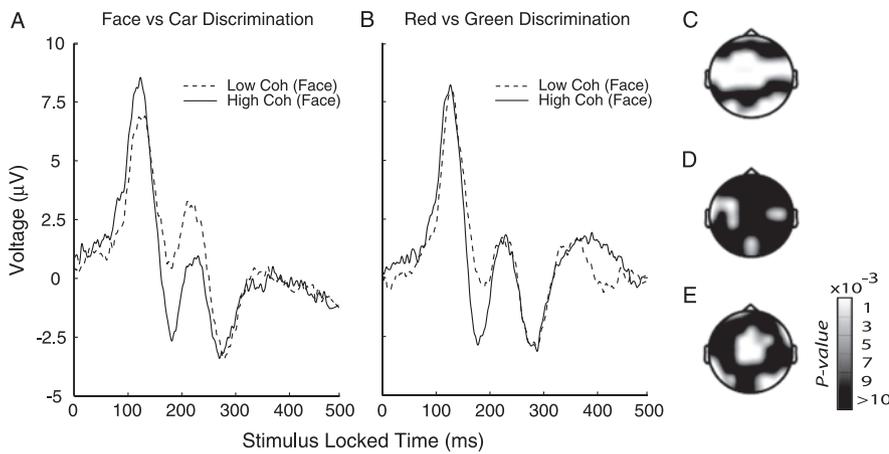
Figure 6. *A*, Average ERPs at electrode PO8 for 30% (dashed line) and 45% (solid line) coherence face trials from a face-versus-car discrimination task. Consistent with our previous findings, a D220 component can be identified. *B*, Average ERPs, at electrode PO8, for low (30%) and high (45%) coherence face trials for the red-versus-green discrimination task. Although the stimulus evidence remained unchanged in this case, the task difficulty was reduced for the low coherence trials (and essentially equalized with that of the high coherence trials). This manipulation allowed the D220 component to be eliminated. Note that although the effect on the D220 component remains unchanged when we included car trials in this analysis, we present results from face trials only to emphasize that the N170 component is present in both discrimination tasks. *C*, Interpolated electrode locations with a significant difference in ERP amplitudes between low and high coherence trials (at 220 ms after stimulus) during a face-versus-car discrimination task as assessed by a bootstrap test. *D*, Interpolated electrode locations with significant difference ERP amplitudes between low and high coherence trials during a red-versus-green discrimination task as assessed by a bootstrap test. *E*, Interpolated electrode sites with a significant difference between the difference in ERP amplitudes of *C* and *D*. Low Coh, Low coherence; High Coh, high coherence.
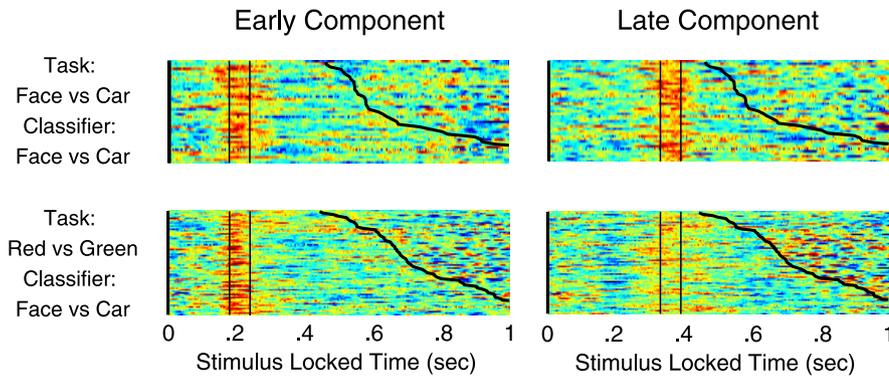
Figure 7. Discriminant component maps for one subject during the face-versus-car (top row) and red-versus-green (bottom row) discrimination tasks. In both tasks, the classifier was trained to discriminate between face and car trials. An early component was identified for both tasks, whereas a significant late component was only present for the face-versus-car categorization task. Red represents positive activity, and blue represents negative activity. Results shown are those of face trials with the mean of the car trials subtracted. All trials were aligned to the onset of visual stimulation, as indicated by the vertical black line at time 0, and sorted by response time. The black sigmoidal curves represent the subjects' response times for each condition.

Figure 9*A* illustrates a single-trial fit (black) of the late discriminating component (gray). We applied this procedure on all trials across the three highest phase coherence levels (i.e., 35, 40, and 45%) to investigate the extent to which the duration of the late component varies with the task difficulty. For every single-trial, we used the SD of the Gaussian fit as a metric to quantify the spread of the component. We then computed the mean SD across all trials and all subjects for each one of the three phase coherence conditions. For every fit, we computed an r^2 value to characterize the goodness-of-fit for $\hat{y}_{late}(t)$. We considered parameters only from fits with $r^2 > 0.5$. As can be seen in Figure 9*B*, we found that there is a systematic increase in the duration of the late component as the percentage of phase coherence is reduced. We observed a significant increase in the mean SD between the 45 and

40% phase coherence conditions ($\sigma_{45\%}$, 40 ms; $\sigma_{40\%}$, 53 ms; paired *t* test, $p < 10^{-12}$) as well as a significant increase between the 40 and 35% conditions ($\sigma_{40\%}$, 53 ms; $\sigma_{35\%}$, 62 ms; paired *t* test, $p < 4 \times 10^{-6}$).

Finally, we point out that there is also a corresponding increase in mean RT variance across all subjects as the percentage of phase coherence is reduced (i.e., as the task becomes more difficult). Figure 9*B* also illustrates this point.

Association between the late component and the diffusion model

Using only behavioral data (i.e., behavioral performance and RT distributions) from all subjects of our original face-versus-car categorization task, we estimated parameters for the diffusion model of decision making (Ratcliff and Rouder, 1998, 2000; Ratcliff and Tuerlinckx, 2002). Specifically, we estimated diffusion rates for five different phase coherence levels (25–45% in 5% increments). We also computed a χ^2 goodness-of-fit value (mean $\chi^2 = 40.9$, with df = 45) which showed that the fits of the model did not significantly deviate from the data. The fits are shown in quantile probability plots in Figure 10*A*. In this plot, the quantiles of the RT distributions are plotted on the *y*-axis and the response proportion is plotted on the *x*-axis, with correct responses to the right and the corresponding error responses to the left.

To provide evidence that our late component could, in fact, represent the mean drift rate in a diffusion model, we correlated the strength of this component (i.e., A_z value), for each subject and at each coherence level, with the estimated accumulation rates from the diffusion model simulations. Figure 10*B* shows that there is a significant correlation ($r = 0.8575$; $p < 1.0 \times 10^{-8}$) between these two measures. This finding is especially interesting considering that the first measure (A_z) is derived using only neural responses and the second one (mean drift rate) purely behavioral responses.

Finally, we looked at the total nondecision time (T_{er}) to estimate the relative position of the late component with respect to the actual decision-making process. Note that the time for processes such as stimulus encoding, response output, memory access, and so on, are all combined in T_{er}. For our subjects, the mean T_{er} was estimated to be $\approx$460 ms. Assuming a 100 ms motor response output, which follows the decision-making process in our paradigm, leaves a total processing time of $\approx$360 ms for any early stimulus encoding/sensory event. Note that the peak of the activity of our late component occurs, on average, $\approx$350 ms after the onset of the stimulus at the highest phase coherence (i.e., easiest condition), and it increases with task difficulty [Philiastides and Sajda (2006), their Fig. 5*A*]. This finding places essentially all of our detected late components directly after early visual
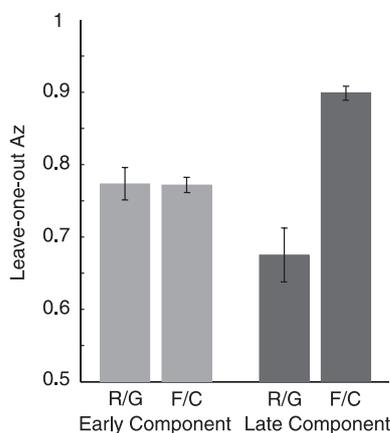
Figure 8. A_z values computed for the early (light gray bars) and the late (dark gray bars) components for the face-versus-car (F/C) and red-versus-green (R/G) discrimination tasks. In both tasks, the classifier was trained to discriminate between face and car trials. Consistent with our previous findings, we identified, for the face-versus-car discrimination task, an early and a late component with significant A_z values. Moreover, the A_z values for the late component were higher than that of the early component. For the color discrimination task, however, we observed a significant reduction in the strength of the late component. The early component, on the other hand, remained equally strong as in the face-versus-car discrimination task. Error bars indicate ±1 SEM.

perception and at the onset of the actual decision-making process. We interpret this result as additional evidence that our late component is unlikely to represent purely early perceptual/sensory events.

Discussion

Figure 11 summarizes our findings in the form of timing diagrams. Discriminating between face and car trials, we initially identified an early face-selective component (N170), the strength of which is proportional to the strength (percentage of coherence) of the stimulus. This component is present, in equal strength, for both a face-versus-car categorization task and a color discrimination task (Fig. 11A,C). Its presence in the color discrimination task, although our subjects were not explicitly discriminating the type of image, seems to indicate that this component represents an early perceptual event and is not directly linked to the actual decision. The late component, on the other hand, is clearly more closely linked to the decision-making process, as evidenced by it being a better predictor of behavioral accuracy (at any given coherence level, its strength, A_z value, is higher than that of the early one) (Fig. 11A,C), the delay in its onset time when a decision is difficult (Fig. 11B), and its suppression when the decision is unrelated to a face-versus-car discrimination (Fig. 11C,D).

In addition, we found a difficulty component that is situated somewhere between the early and the late components. Specifically, we found a component ∼220 ms after the onset of visual stimulation, the strength of which is inversely proportional to the stimulus evidence (Fig. 11A,C). Moreover, the amplitude of this component is highly correlated to the difficulty of the task, as captured by the behavioral performance of our subjects, and it also predicts the onset of the late face-selective component. As such, this component appears more closely linked to the late rather than the early component. We also showed that this difficulty component (D220) is likely to represent a true top-down influence on decision making, rather than a mere bottom-up processing of the stimulus, by virtue of the fact that it disappears during an easy color (red-vs-green) discrimination (Fig. 11C,D).
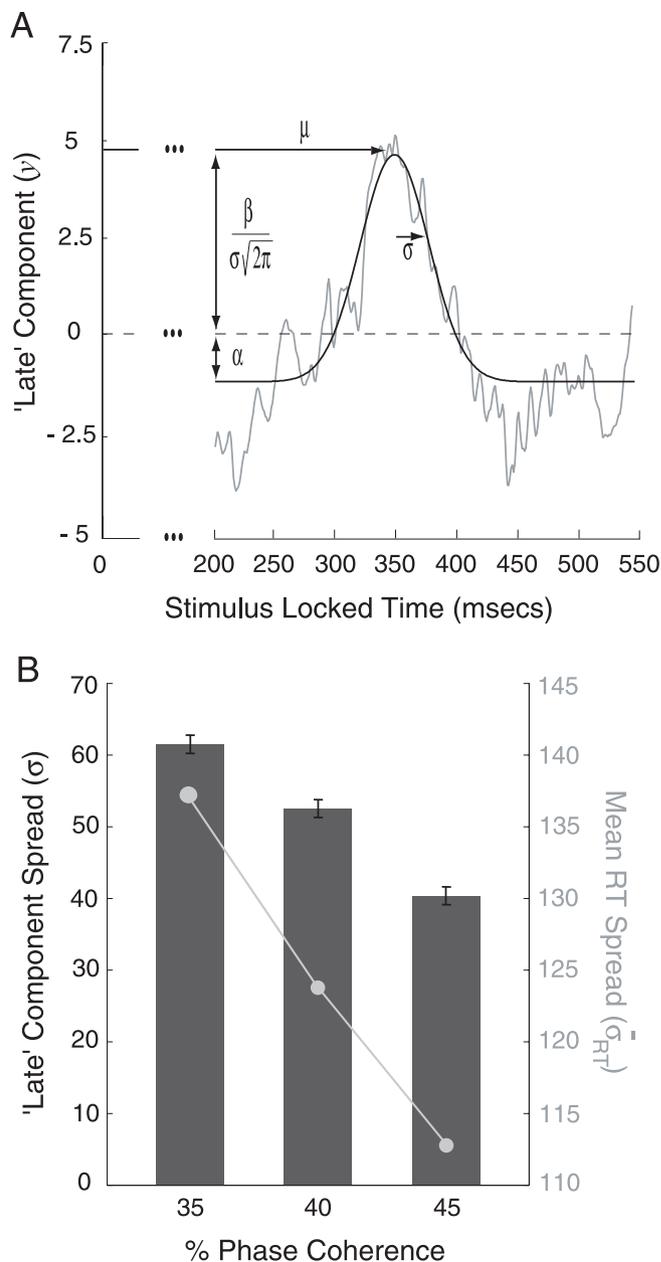
Figure 9. Illustration of our single-trial peak fitting procedure and estimation of the late component duration as a function of stimulus evidence. **A**, Fit $\hat{y}(t)$ (black) for component activity $y(t)$ (gray) from one trial. Parameters of a single-trial fit include the following: discriminating component peak latency relative to the onset of visual stimulation (μ) and component spread (σ). **B**, Results indicating increased mean component duration (i.e., mean σ across trials and subjects) as a function of the percentage of phase coherence. Note also that there is a corresponding increase in mean RT variance as the task becomes more difficult. Error bars indicate ±1 SEM.

We speculate that the difficulty component could be implicated in the recruitment of the relevant attentional and other neuronal resources required to make a difficult decision. For instance, an EEG component with similar latency (N2pc) was shown to capture the moment-by-moment rapid shift of attention in a demanding visual search task (Woodman and Luck, 1999). The N2pc component is also implicated in covert orientation of visual attention during object recognition (Luck and Hillyard, 1994) and was shown to resemble the attention-related modulations of activity seen in monkeys (Luck et al., 1997).

All three components are stimulus locked and thus, within the
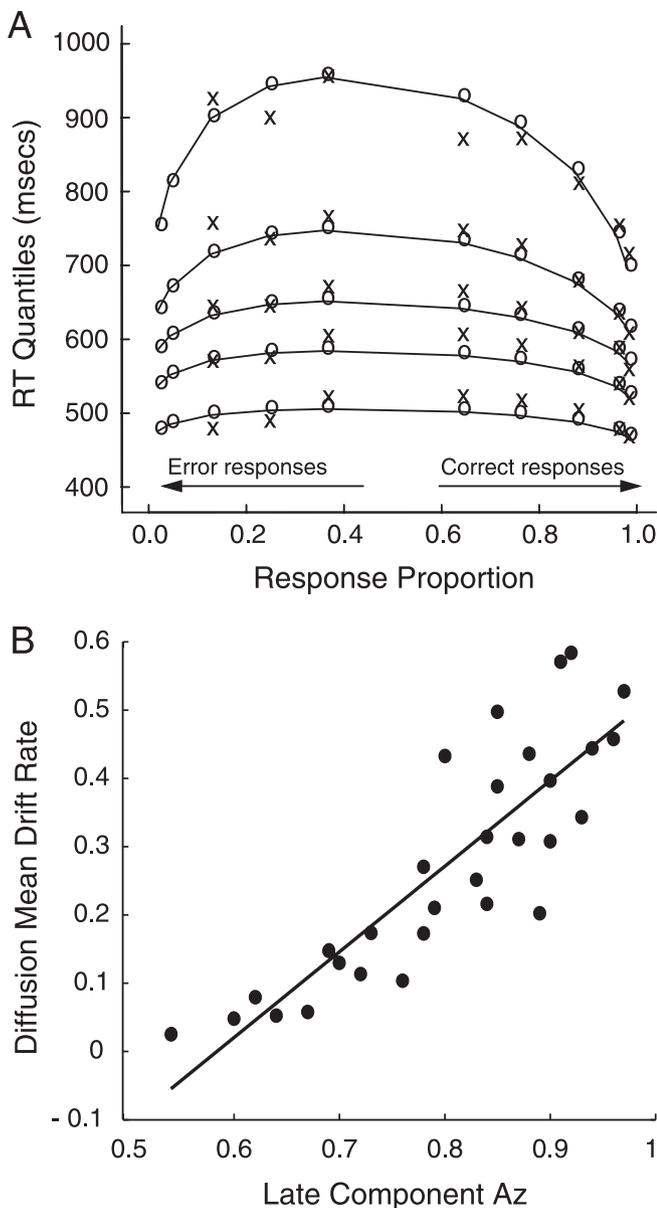
Figure 10. Summary of the results from the diffusion model simulations. **A**, Theoretical (×) and empirical (○) quantile-probability plots for our data. The lines represent the theoretical fits of the diffusion model. The quantile RTs in order from the bottom to the top are 0.1, 0.3, 0.5, 0.7, and 0.9, and in each vertical line, the quantiles must have this order. In some conditions, a moderate proportion of subjects do not have enough responses to allow computation of quantiles, and so no quantiles are presented for these extreme error conditions. **B**, Correlation between the strength of our late component (A_z metric) and the mean drift rate as derived from a diffusion model simulation using behavioral responses (i.e., behavioral performance and RTs) from six subjects at five different phase coherence levels. The results demonstrate significant correlation ($r = 0.8575$; $p < 1.0 \times 10^{-8}$) between the two measures.

context of a neural integrator (Shadlen and Newsome, 2001; Mazurek et al., 2003; Huk and Shadlen, 2005) or diffusion model (Ratcliff, 1978; Luce, 1986; Smith and Ratcliff, 2004; Palmer et al., 2005), all components, including the late one, are unlikely to represent the process of evidence accumulation to threshold which can account for both accuracy and RT variability (i.e., because the late component is stimulus locked and does not persist until the response, it does not predict the trial-by-trial RT distribution within a coherence level).

Additional analysis, however, allows us to relate the late component with other parameters of an integrator or diffusion model. For example, the late component is highly predictive of behavioral accuracy (i.e., a good match between neurometric and psychometric functions) and shifts in time as a function of difficulty. This time shift correlates with the shift of the mean RT at a given coherence level. In addition, the duration of the late component gradually increases as coherence is reduced (Fig. 11 B) and is strongly correlated with RT variance. Finally, there is a significant correlation between the drift rates computed from the diffusion model simulations and the strength (A_z values) of the late component. These findings are all consistent with the late component being analogous to the mean drift rate in a diffusion model (Ratcliff, 1978; Palmer et al., 2005), much the same way as noisy, but sustained, activity in the middle temporal area would drive the integration process in neural integration models (Mazurek et al., 2003) (although, in that case, the mean drift rate would typically be thought of as a signed quantity). In other words, the late component represents the postsensory evidence that is fed into the diffusion process that ultimately determines the decision. Last, the fact that we find no evidence of a component that discriminates a red-versus-green decision suggests that the late component activity is not indicative of a general purpose decision-making process (Heekeren et al., 2004) and is more likely part of a face-selective network that drives such a process.

Our choice of stimuli, in particular faces, was made because faces are known to activate strong neural generators measurable via EEG/magnetoencephalography (Botzel et al., 1995; Bentin et al., 1996; Jeffreys, 1996; Liu et al., 2000, 2002; Rossion et al., 2003). One question is whether our timing diagram generalizes for other stimulus classes. Unpublished work by our group investigates this question using motion coherence stimuli to find neural correlates of decision making in EEG. However, we found that the neural generators are not strong enough to enable construction of neurometric functions on more than one or two relative motion coherences that are well above perceptual threshold.

The issue of localization is one that we have not addressed. A number of single-unit and fMRI studies have attempted to localize neural activity related to perceptual decision making (Kim and Shadlen, 1999; Platt and Peterson, 1999; Shadlen and Newsome, 2001; Heekeren et al., 2004; Ridderinkhof et al., 2004), with several studies specifically addressing the issue of decision difficulty and uncertainty (Critchley et al., 2001; Paulus et al., 2002; Volz et al., 2003, 2004; Huettel et al., 2005; Grinband et al., 2006). These studies have found a variety of areas modulated by decision difficulty, including areas in the frontal and parietal cortices, the thalamus, as well as the striatum. Our work complements these previous findings by focusing on the timing of the underlying neural processes. For example, our results show a change in the timing of a late component that is predictive of behavioral accuracy, suggesting that for difficult decisions, not only are different brain areas engaged or modulated, but also additional latency is introduced. This latency may represent the need for, or allocation of, additional processing time required to evaluate the evidence. Consistent with this last hypothesis is the presence of a difficulty component in our data that appears after the initial evaluation of the evidence (early component) and before the late face-selective component that is mostly linked to a postsensory/decision event.

Intriguing in our results is that the scalp plots for the early and late components are similar in topology, except for sign (Philiastides and Sajda, 2006). One might consider using inverse methods, such as dipole fitting, to localize the component activities. However, given the ill-posed nature of the inverse problem in EEG, we are instead focusing on simultaneous EEG and fMRI to localize the component activities in our timing diagram.

In summary, we show that with single-trial analysis of EEG, we can construct a timing diagram that begins to separate perceptual processing from decision-making processing in human subjects. Our results suggest that cortical networks selective to task difficulty could play an integral role in perceptual decision making by dynamically allocating resources, such as additional processing time, for difficult decisions.

Figure 11. Timing diagrams summarizing the temporal evolution of component activity during the perceptual categorization tasks introduced in this study. These diagrams are given the same dimensions as the discrimination component maps so that they can be directly related to the experimental data (see Fig. 7). Each column represents the type of experiment (left, face-vs-car categorization; right, color discrimination) at two hypothetical phase coherence levels, a high one (top row) and a low one (bottom row). We represent each temporally specific component as a color-coded vertical bar to resemble the stimulus-locked profile of these components as seen in an actual discriminant component map. The color represents the strength of the component. Hypothetical RT curves are shown by the black sigmoidal curves. See Discussion for a complete summary of the fundamental relationships between these components.

References

Bentin S, Allison T, Puce A, Perez A, McCarthy G (1996) Electrophysiological studies of face perception in humans. J Cogn Neurosci 8:551–565.

Bisley J, Goldberg M (2003) Neuronal activity in the lateral intraparietal area and spatial attention. Science 299:81–86.

Botzel K, Schulze S, Stodieck R (1995) Scalp topography and analysis of intracranial sources of face-evoked potentials. Exp Brain Res 104:135–143.

Britten K, Shadlen M, Newsome W, Movshon J (1992) The analysis of visual motion: a comparison of neuronal and psychophysical performance. J Neurosci 12:4745–4765.

Critchley H, Mathias C, Dolan R (2001) Neural activity in the human brain relating to uncertainty and arousal during anticipation. Neuron 29:537–545.

Dakin S (2002) What causes non-monotonic tuning of fMRI response to noisy images? Curr Biol 12:476–477.

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

Duda R, Hart P, Stork D (2001) Pattern classification. New York: Wiley.

Egner T, Hirsch J (2005) Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. Nat Neurosci 8:1784–1790.

Gerson AD, Parra LC, Sajda P (2005) Cortical origins of response time variability during rapid discrimination of visual objects. NeuroImage 28:342–353.

Gill P, Murray W, Wright M (1981) Practical optimization. New York: Academic.

Green D, Swets J (1966) Signal detection theory and psychophysics. New York: Wiley.

Grinband J, Hirsch J, Ferrera VP (2006) A neural representation of categorization uncertainty in the human brain. Neuron 49:757–763.

Halgren E, Raij T, Marinkovic K, Jousmaki V, Hari R (2000) Cognitive response profile of the human fusiform face area as determined by MEG. Cereb Cortex 10:69–81.

Heekeren H, Marrett S, Bandettini P, Ungerleider L (2004) A general mechanism for perceptual decision-making in human brain. Nature 431:859–862.

Huettel S, Song A, McCarthy G (2005) Decisions under uncertainty: probabilistic context influences activation of prefrontal and parietal cortex. J Neurosci 25:3304–3311.

Huk A, Shadlen M (2005) Neural activity in the macaque parietal cortex reflects temporal integration of visual motion signals during perceptual decision making. J Neurosci 25:10420–10436.

Jeffreys D (1996) Evoked potential studies of face and object processing. Vis Cogn 3:1–38.

Jordan M, Jacobs R (1994) Hierarchical mixtures of experts and the EM algorithm. Neural Comput 6:181–214.

Kim J, Shadlen M (1999) Neural correlates of decision making in the dorsolateral prefrontal cortex of the macaque. Nat Neurosci 2:176–185.

Liu J, Higuchi M, Marantz A, Kanwisher N (2000) The selectivity of the occipitotemporal M170 for faces. NeuroReport 11:337–341.

Liu J, Harris A, Kanwisher N (2002) Stages of processing in face perception: an MEG study. Nat Neurosci 5:910–916.

Luce RD (1986) Response times: their role in inferring elementary mental organization. In: Oxford psychology series, Vol VIII. New York: Oxford UP.

Luck S, Hillyard S (1994) Electrophysiological correlates of feature analysis during visual search. Psychophysiology 31:291–308.

Luck S, Girelli M, McDermott M, Ford M (1997) Bridging the gap between monkey neurophysiology and human perception: an ambiguity resolution theory of visual selective attention. Cognit Psychol 33:64–87.

Mazurek M, Roitman J, Ditterich J, Shadlen M (2003) A role for neural integrators in perceptual decision making. Cereb Cortex 13:1257–1269.

Palmer J, Huk AC, Shadlen MN (2005) The effect of stimulus strength on the speed and accuracy of a perceptual decision. J Vision 5:376–404.

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

Parra L, Spence C, Gerson A, Sajda P (2003) Response error correction: a demonstration of improved human-machine performance using real-time EEG monitoring. IEEE Trans Neural Syst Rehabil Eng 11:173–177.

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

Paulus M, Hozack N, Frank L, Brown G (2002) Error rate and outcome predictability affect neural activation in prefrontal cortex and anterior cingulate during decision-making. NeuroImage 15:836–846.

Philiastides M, Sajda P (2006) Temporal characterization of the neural correlates of perceptual decision making in the human brain. Cereb Cortex 16:509–518.

Platt M, Peterson S (1999) Neural correlates of decision variables in parietal cortex. Nature 400:233–238.

Posner M, Peterson S (1990) The attention system of the human brain. Annu Rev Neurosci 13:25–42.

Ratcliff R (1978) A theory of memory retrieval. Psychol Rev 85:59–108.

Ratcliff R, Rouder J (1998) Modeling response times for two choice decisions. Psychol Sci 9:347–356.

Ratcliff R, Rouder J (2000) A diffusion model account of masking in two-choice letter identification. J Exp Psychol Hum Percept Perform 26:127–140.

Ratcliff R, Tuerlinckx F (2002) Estimating parameters of the diffusion model: approaches to dealing with contaminant reaction times and parameters variability. Psychon Bull Rev 9:438–481.

Ratcliff R, Zandt TV, McKoon G (1999) Connectionist and diffusion models of reaction time. Psychol Rev 106:261–300.

Ridderinkhof K, Ullsperger M, Crone E, Nieuwenhuis S (2004) The role of the medial frontal cortex in cognitive control. Science 306:543–447.

Romo R, Hernandez A, Zainos A, Salinas E (2003) Correlated neuronal discharges that increase coding efficiency during perceptual discrimination. Neuron 38:649–657.

Rossion B, Joyce C, Cottrell G, Tarr M (2003) Early lateralization and orientation tuning for face, word, and object processing in the visual cortex. NeuroImage 20:1609–1624.

Sandwell D (1987) Biharmonic spline interpolation of GEOS-3 and SEA-SAT altimeter data. Geophys Res Lett 14:139–142.

Schall J, Hanes D (1993) Neural basis of saccade target selection in frontal eye field during visual search. Nature 366:467–469.

Shadlen M, Newsome W (2001) Neural basis of perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. J Neurophysiol 86:1916–1936.

Smith P, Ratcliff R (2004) Psychology and neurobiology of simple decisions. Trends Neurosci 27:161–168.

VanRullen R, Thorpe S (2001) The time course of visual processing: from early perception to decision-making. J Cogn Neurosci 13:454–461.

Volz K, Schubotz R, von Cramon D (2003) Predicting events of varying probability: uncertainty investigated by fMRI. NeuroImage 19:271–280.

Volz K, Schubotz R, von Cramon D (2004) Why am I unsure? Internal and external attributions of uncertainty dissociated by fMRI. NeuroImage 21:848–857.

Woodman G, Luck S (1999) Electrophysiological measurement of rapid shifts of attention during visual search. Nature 400:867–869.