Single-Trial Analysis of Neuroimaging Data: Inferring Neural Networks Underlying Perceptual Decision-Making in the Human Brain

Paul Sajda, *Senior Member, IEEE*, Marios G. Philiastides, and Lucas C. Parra, *Senior Member, IEEE*

Methodological Review

Abstract—**Advances in neural signal and image acquisition as well as in multivariate signal processing and machine learning are enabling a richer and more rigorous understanding of the neural basis of human decision-making. Decision-making is essentially characterized behaviorally by the variability of the decision across individual trials—e.g., error and response time distributions. To infer the neural processes that govern decision-making requires identifying neural correlates of such trial-to-trial behavioral variability. In this paper, we review efforts that utilize signal processing and machine learning to enable single-trial analysis of neural signals acquired while subjects perform simple decision-making tasks. Our focus is on neuroimaging data collected noninvasively via electroencephalograpy (EEG) and functional magnetic resonance imaging (fMRI). We review the specific framework for extracting decision-relevant neural components from the neuroimaging data, the goal being to analyze the trial-to-trial variability of the neural signal along these component directions and to relate them to elements of the decision-making process. We review results for perceptual decision-making and discrimination tasks, including paradigms in which EEG variability is used to inform an fMRI analysis. We discuss how single-trial analysis reveals aspects of the underlying decision-making networks that are unobservable using traditional trial-averaging methods.**

Index Terms—**Decision-making, electroencephalography, functional magnetic resonance imaging, machine learning, single-trial analysis.**

I. INTRODUCTION

OUR ability to make rapid decisions based on transient stimuli is a unique aspect of our brains' capacity to process information. Recent advances in signal processing and neuroimaging have begun to provide tools for answering questions related to how we make rapid decisions by enabling the identification of cortical networks underlying rapid decision-making in the human brain. Typically, such studies utilize

functional magnetic resonance imaging (fMRI) and have identified areas in frontal, parietal, and thalamic cortex in which metabolic activity correlates with decision-related variables. However, decision-making is a dynamic process, and the localized activations found with fMRI must be part of cortical networks defined by the relative timing of these activations and their causality. Electroencephalography (EEG) offers the ability to resolve neural processing at the millisecond level, however, with a relatively poor spatial resolution. Combinations of these modalities have been recently developed enabling more precise identification of cortical networks underlying rapid decision-making, including those processes involved in integration of contextual information and accumulation of multiple sources of evidence.

At the core of the scientific and engineering challenge is how to analyze neuroimaging data on a trial-by-trial basis, since perceptual decision-making is in fact characterized by variations in behavior across trials, even given nominally identical stimuli. In this paper, we review experimental results in animals and humans with respect to perceptual decision-making, focusing on the role neuroimaging has to play in shedding light on the underlying information processing. We also review how signal processing and machine learning play a role, specifically in our ability to identify meaningful neural signatures, on a trial-by-trial basis, to characterize the underlying network processes. We also review our own specific results, which focus on using EEG, fMRI, and both simultaneously for inferring the networks underlying perceptual decision-making.

II. DECISION-MAKING AND NEUROIMAGING

Perceptual decision-making is the process by which incoming sensory information is combined and used to influence our behavior [1]. A number of ground-breaking electrophysiology studies have already laid the foundation upon which future research on perceptual decision-making will be conducted [2], [3].

A. Perceptual Decision-Making in Animals

Experiments in primates have already established a relationship between behavioral choices and neuronal activity in sensory regions. In a motion discrimination task during which monkeys had to discriminate the direction of motion from random

dot kinetograms consisting of varying amounts of coherent motion, Newsome *et al.* showed that the activity of direction-selective neurons in middle temporal area (MT) can provide an adequate account of behavioral performance [4]–[6] and that trial-to-trial variability in these neuronal signals could predict the monkeys' actual choices [7]. In addition, electrical microstimulation of MT neurons biased the monkeys choices towards the neurons' preferred direction [8]–[10]. Interestingly, this pattern of neural responses appears to extend even to highly complex visual stimuli such as faces [11], [12].

Similarly, in the somatosensory domain, Romo *et al.* used a task in which monkeys had to discriminate the vibration frequency of two sequentially presented tactile stimuli and report which one was the highest. They showed that, on average, the responses in primary somatosensory cortex (S1) correlated with the behavioral performance of the monkeys and the trial-to-trial fluctuations in these responses predicted the animals' choices [13]. Furthermore, microstimulation of S1 neurons, in the absence of physical stimulation, was sufficient to reproduce the behavioral patterns seen under normal conditions [14], [15]. These findings support the idea that neuronal signals in areas such as MT and S1 provide the sensory evidence upon which monkeys base their decisions regarding motion and vibrotactile discrimination, respectively.

How, then, can activity in these areas be used to influence decision-making? Theoretical models of decision-making have modeled binary perceptual choices as a race-to-barrier diffusion process [16]–[20]. The diffusion model for two-choice decisions in particular assumes that in the decision process, evidence is integrated over time to one of two decision thresholds corresponding to the two choices. The rate of accumulation is called drift rate, and it is determined by the quality of the sensory information. The better the information quality, the larger the drift rate toward the appropriate decision boundary and the faster and more accurate the response.

In the monkey brain, a region in the intraparietal sulcus, commonly known as the lateral intraparietal (LIP) area, is shown to exhibit some of the characteristics proposed by the aforementioned models. Specifically, electrophysiological recordings in LIP during the motion discrimination task [21]–[24] have shown that, for choices made towards the response field of the neuron under consideration, neural activity increased in a ramp-like fashion, consistent with a temporal integration process. The rate at which responses increased was proportional to the amount of coherent motion in the stimulus. This buildup of activity is believed to represent the accumulated difference in firing rates of two opposing pools of direction selective MT neurons. Just prior to the response, neural responses in LIP achieved a common firing rate, independent of motion strength, consistent with the idea of a common decision threshold. More recently, these results were extended from a two- to a four-choice direction discrimination task [25]. The same pattern of activity was reported in the frontal eye fields (FEFs) and the dorsolateral prefrontal cortex (DLPFC) [26] and is seen as evidence that all of these areas are involved in the conversion of an analog motion representation into a binary decision variable.

Likewise, during vibrotactile frequency discrimination, regions in the secondary somatosensory (S2), medial premotor

cortex (MPC), ventral premotor cortex (VPC), and DLPFC were reported as being involved in decision formation [27]–[30]. As with the LIP, responses in the MPC and VPC were also shown to form a decision by computing the difference between the responses of S2 neurons that code for each of the two frequencies used for stimulation [29], [30]. Furthermore, it is believed that some S2 neurons have the capacity to combine past and present sensory information to form a decision [28] and that short-term memory representation in DLPFC might also play a role in decision-making [27], [31].

B. Perceptual Decision-Making in Humans

The principles that have emerged from the electrophysiology work in animals have inspired scientists to study the neural correlates of perceptual decision-making in humans using noninvasive neuroimaging techniques such as the fMRI and the EEG.

Heekeren *et al.* used fMRI to study perceptual decision-making during a face-versus-house discrimination task [32]. Consistent with previous neuroimaging studies [33]–[37], they identified face and house selective regions in the ventral temporal cortex that, much like MT in a motion discrimination task, provide the sensory evidence upon which a decision is later computed. Specifically, they showed that these regions were modulated by the amount of degradation in the face and house stimuli, respectively. Inspired by the work of Romo *et al.*, recent fMRI studies have also used vibrotactile frequency tasks to identify regions that encode the sensory evidence needed for a somatosensory decision. Consistent with electrophysiological data in monkeys, they found that the somatosensory cortex and S1 in particular encode the incoming sensory information [38], [39].

As summarized above, animal work has provided evidence that the brain combines this sensory information using a difference-based comparator operation. Heekeren *et al.* directly tested whether a comparison operation is also at work in the human brain using the face-versus-house discrimination task outlined above [32]. They proposed that higher level decision areas should show the greatest activity on easy trials (clear images of faces and houses) rather than on hard ones (degraded images) and, in addition, they should correlate with the difference between the signals from the face and house selective regions that represent the sensory evidence. The only region that exhibited this pattern of activity was the posterior portion of the left DLPFC. Therefore, even for complex object categories, the brain appears to integrate the outputs of different pools of selectively tuned neurons in order to make perceptual decisions.

Another recent fMRI study also showed how the fMRI signal can be used to examine the process of accumulation of sensory evidence [40]. Pictures were revealed gradually over the course of several seconds, and participants indicated the time of recognition with a button press. There was a gradual buildup in fMRI signal peaking in correspondence with the time of recognition in inferior temporal, frontal, and parietal regions, suggesting that these regions might be involved in accumulating sensory evidence.

Contrary to electrophysiology studies in monkeys that have thus far claimed that "to see and decide is, in effect, to plan a motor-response" [41], Heekeren *et al.* also investigated whether

decisions may be transformed into actions independent of the motor effectors used to indicate the choice [42]. They used the same motion discrimination task used in the monkey studies and asked participants to indicate their choice with either a button press or an eye movement. They found that the area in the left DLPFC they identified earlier [32] exhibited decision-related activity (as highlighted above) independent of the motor system used to indicate the choice. They concluded "that humans may have evolved a more abstract decision-making network, thereby allowing a more flexible link between decision and action" [32].

Despite fMRI's excellent spatial resolution, the slow scanning rates and the low-pass nature of the hemodynamic signal place significant constraints on its capacity to provide high temporal resolution information. To overcome this limitation, advanced multivariate analysis techniques have been developed to also study the temporal characteristics of perceptual decision-making in humans, as we highlight in the next sections.

III. MULTIVARIATE PROCESSING AND NEUROIMAGING

A. Capability of Various Neuroimaging Modalities

All noninvasive neuroimaging modalities make only indirect measures of neuronal activity. fMRI records the local hemodynamic response (blood oxygenation level) within the brain following increased metabolic activity of neurons. This can be recorded with millimeter spatial resolution over the entire brain, which results in thousands of simultaneously recorded voxels. Hemodynamic response is sluggish (1–5 s), and so activity is typically sampled only every 1 or 2 s.[1]

EEG measures electrical potentials that build up at the surface of the scull resulting from large-scale ionic currents in the brain. These currents originate primarily at the dendrites of neurons reflecting neuronal input activity. Surface electrical potentials are easy to record and capture primarily cortical activity with centimeter spatial resolution. Currents from deeper structures decay significantly and are too broadly distributed on the scull to be localized with much accuracy. However, EEG signals do track fast changes of neural activity and are thus recorded with millisecond temporal resolution. Useful signal may be found at frequencies of up to 80 Hz, which is much higher than the less than 1 Hz temporal resolution of fMRI.

Magnetic encephalography (MEG) detects static magnetic fields generated by the same ionic currents that underlie the EEG signal. MEG therefore has a similar temporal resolution. However, because the brain is engulfed by the cerebrospinal fluid, direct currents measured in EEG are partially shunted, while magnetic field lines cross unaltered, leading to a somewhat better localization of cortical current sources when using MEG. As a corollary, fast synchronous neuronal activity that has a shorter length scale is easier to detect and localize, and so the useful frequencies in MEG may be as high as 100 Hz. The disadvantage of MEG is its high cost, which is comparable to MRI with its much higher spatial resolution.

[1]Because measuring magnetic resonance is very costly and requires significant infrastructure, some have proposed to image hemodynamic activity using near-infrared light, which does penetrate the scull. While this functional near-infrared technology is more accessible, its spatial resolution and penetration depth are limited by light diffusion, and its temporal resolution remains limited by the slow hemodynamic response.
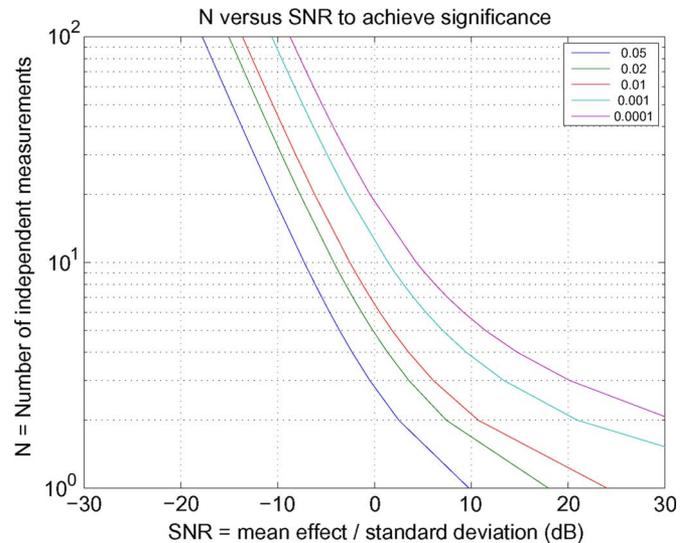
Fig. 1. Power analysis: number of independent measures N required at a given SNR to achieve statistical significant difference in mean activity between two experimental conditions. Curves for various significance levels (p-values) are computed assuming Normal distributed data. Both fMRI and EEG require about the same number of trials to show comparable significance levels, thus suggesting that their SNR is comparable. Multivariate analysis increases N, in effect, by using multiple samples, thus requiring fewer trials for averaging. Machine-learning techniques are used to determine over which samples to average and with which weighting.

B. Decoding Problem is Similar for Different Modalities

While the various functional neuroimaging modalities may have different spatial and temporal resolutions, when considering activity on a second by second basis, they all provide a comparable number of separate measurements. fMRI measures perhaps once a second a volume of approximately $50 \times 50 \times 20$ voxels. EEG or MEG measure about 100 channels at perhaps 500 samples per second. In either case, one is left with approximately 50 000 measurements per second. We will refer to these in the following simply as "samples" reflecting a sampling of brain activity either in space (fMRI voxels) or space–time (EEG electrodes and samples). Interestingly, the amount of noise in the two modalities are also comparable. Typically, for either modality, trials have to be repeated several times before significant differences between two experimental conditions can be seen (maybe, $N = 100$). A simple power analysis indicates that the signal-to-noise ratio (SNR) in the individual dimensions is therefore about -13 dB (see Fig. 1 for other values of N).

Rarely do individual samples carry sufficient information to distinguish between different states of the brain. To recover significant activity, conventional analysis has relied on averaging over many repeated trials of the same stimulus and task conditions under the assumption that the relevant neuronal activity is reproducible from one trial to the next. While this approach has been highly successful in the past, we now know that it severely underestimates the amount of information that can be extracted from these simultaneous recordings. The reason is that activity may be distributed over very many samples and may not be strong enough in any one sample to allow a distinction between different neural processes. Instead, when carefully combining multiple samples with multivariate analysis techniques, one can

identify neural activity that correlates well with observed behavioral responses despite rather low SNR.

Given the different nature of fMRI and EEG, one would think that the analysis techniques to accomplish this differ drastically between the two modalities. Surely the correlation structure of neighboring samples and the trial-to-trial variability differ significantly, and this should affect how one is to combine these high-dimensional data. However, despite the significant difference of fMRI and EEG, there has been a remarkable convergence in the techniques that have been used to extract behavioral correlates. The successful multivariate analysis methods are based on modern machine-learning techniques, which aim to uncover a pattern in high-dimensional spaces based on a small set of examples. The linear decoding techniques often amount to cleverly averaging over the large dimensional space, thus accomplishing what trial-averaging typically does, i.e., use a large number of independent measurements to increase SNR (see Fig. 1).

What is most remarkable is that the successful techniques for decoding neuroimaging signals are very similar to the techniques that have been proposed in machine learning for the goal of classifying complex dynamical patterns. The basic notion of what is known as reservoir computing [43], [44] is to represent a stimulus in a high dimensional space capturing many of its nonlinear features but then use a simple weighted sum of these features to classify the pattern into a few categories. It appears that the specific choice of nonlinear features is not very important as long as there is a large number of distinct features to assure that all relevant invariances are captured and that the representation is diverse enough to distinguish between different stimuli [45].

The neuroimaging signals we measure surely are a highly nonlinear representation of the perceptual stimuli or decision processes we aim to decode. The relevant parameters for decoding are then the dimensionality of the problem, the amount of noise, and the invariance and diversity in the feature space. In that regard, the challenge of extracting neural correlates is strikingly similar for the different neuroimaging modalities: The goal is to find in an approximately 10^4-dimensional space neural activity that correlates with a few possible decision outcomes (often just two) based on about 10^2 exemplars of signals with at best $-10 \, \mathrm{dB}$ in SNR. The only potential difference in the modalities is then the invariance and diversity of neural signals. Only here is where domain knowledge of the different modalities is required.

C. Single-Trial Analysis and the Decoding Problem

The original emphasis on single-trial analysis of neuroimaging data comes from the ambitious goal of "reading the brain," that is, interpreting neural signals at the speed in which they develop so as to translate thought into action [46]. Aside from the potential practical benefits of such technology,[2] there is an important benefit to the study of perceptual and cognitive processing.

[2] Human studies have used EEG signal to extract neural activity in real-time with sufficient accuracy to control a cursor on the screen in one or two dimensions [47], [48] or to control the grip of a robotic hand [49]. Animal studies using electrodes implanted in various cortical areas have demonstrated two- and three-dimensional control of a robotic arm [50]. The obvious goal of this work is to develop a brain-controlled communication or prosthetic device to overcome the impairments resulting from spinal-cord injuries (for an overview, see [51]). Our team has used single-trial analysis to augment the performance of unimpaired subjects at a perceptual decision-making task, specifically, visual search [52].

At a minimum, by combining multiple measures to an aggregate measure with higher SNR, one may be able to discover signals that have thus far eluded detection. Equivalently, it may be possible to obtain similar effect sizes with fewer trials, thus allowing the experimenter to segment the data into multiple response categories, e.g., distinguish fast responses from slow responses to identify the correlates of response-time variability [53]. Indeed, even when subjects are given an identical task with an identical stimulus, they never repeat the exact same behavior at the exact same speed and with the same neuronal signals. Uncovering the systematic covariation of neural signals with variations in behavioral response (speed, accuracy, strength, etc.) may shed light on internal states such as arousal, attention, motivation, etc., which are difficult to control experimentally. In addition, in some experiments, the number of stimulus dimensions may be too large to systematically explore. Quantifying the correlation of single-trial activity with a large set of stimulus features may uncover relationships that would otherwise be impossible to explore due to a limited sample size. Finally, single-trial activity could be used to correlate distinct neuroimaging modalities to each other [54] and even to identify common signals in the activity detected for different subjects [55].

The methods that have been developed for single-trial analysis can be grouped into two categories: 1) methods in which the presented stimulus is used to predict the resulting neural activity—we call this the forward model; and 2) methods that try to determine from the neural signal what stimulus was presented to the subject or how the subject responds. We call this direct decoding and note that forward models are often also inverted to similarly decode brain activity. These two approaches will be discussed next, followed by a discussion of the central problem of regularization.

D. Direct Decoding of Discrete States

The most straightforward method of "reading the brain" is to try to distinguish between two externally observable events based on some appropriate sum of the observed brain activity. For instance, say a subject on a given trial is presented with a target stimulus—call this the positive class. The negative class is the case in which the subject was presented a stimulus without a target. Assume that one records for the jth trial in sample i the activity x_{ij}. To decide on the class, a linear classifier takes a weighted sum of this activity x_{ij}

$$y_j = \sum_{i=1}^{D} w_i x_{ij}. \tag{1}$$

If this sum is larger than some threshold $y_j > \theta$, then this trial j is considered to belong to the positive class; otherwise it is considered to belong to the negative class. The magnitude and sign of the weights indicates how much each sample should contribute to the decision and in which direction. The goal is to find weights w_i such that the classification into positive and negative class is correct for most trials.

Essentially, the goal is to determine from the magnitude of the measured brain activity whether the subject did or did not perceive the target. If we are able to do so, then we can argue that we have found the neural correlate of perception, or, equivalently,

we can argue that we can "read" the brain signals to determine whether the subject did or did not perceive a target. The same could be done with signals associated with decision-making by trying to correctly classify what choice a subjects is about to make, or by predicting whether a response to a task will be correct or incorrect.

Given the $D \approx 10^4$ dimensions of the samples, it is easy to find a suitable set of weights that give the correct classification on, say, $N = 100$ "training" exemplars. The challenge lies in finding weights that will perform well on "testing" examples that have not been used during "learning," i.e., weights that will generalize well. There are a variety of learning algorithm for this binary classification problem, which vary in their ability to generalize despite the small number of training exemplars and in their ability to tolerate noisy exemplars. Support-vector machines (SVMs) perform well when the number of degrees of freedom is large compared to the number of training exemplars—as in the present case with $D > N$. These algorithms focus on exemplars that are closest to the dividing boundary between the two classes—these are called the support vectors. However, without modifications, SVMs do not perform well in the presence of significant noise, as the boundary elements will invariably be misleading (for most brain-reading applications SNR < -20 dB). Alternatively, one can use logistic regression (LR), which is naturally suited to noisy data, as the transition from one class to the other is assumed to be gradual following a logistic function. On the other hand, in the case of perfectly separable data, LR has to be modified to prevent a transition boundary that is infinitely sharp. Regularization can be used to trade off between perfect separability, for which SVM dominates, with low SNR for which LR dominates (see Section III-G). In our applications, we have sometimes also used Fisher linear discriminants (FLDs), which aim to increase SNR [56]. While FLD is simple to compute, it is not very robust since the required covariance estimates are dominated by outliers, which are common in neuroimaging data.

Direct decoding using linear classification as outlined above has been used with great success in fMRI, and several reviews are already available [46], [57]. The following are just a few examples of the type of activity that could be extracted: Brain activity of a listener was used to detect the identity of one of two speakers [58]. The geometry of a visual stimulus can be determined from visual cortex activity by predicting areas of high and low contrast on the image [59]. Decoders have also been used to distinguish spatial memories by distinguishing between two locations to which a subject navigated within a virtual environment despite identical visual appearance of the locations [60]. Concepts evoked by simple line drawings could be categorized without the use of visual cortex based on activity broadly distributed over the entire brain, and this categorization generalized well across subjects [61]. Free choice decisions could be detected several seconds ahead of an overt behavioral response [62], and even ahead of the conscious realization by the subject that he/she had reached a decision [63].

The voxels that contribute significant discriminant activity in these classifications indicate which brain areas are involved in a specific task. The results from these studies have changed many existing notions of localization of brain activity. In many instances, it has become clear that processing involved larger areas that are overlapping, suggesting a higher degree of multitasking than previously thought.

We have used linear classification of EEG signals on a single-trial basis to analyze rapid perceptual decisions. Linear classifiers using about 1 s of data can determine with high accuracy (>90%) when a subject has seen an object of interest in an image presented only 50–100 ms [64]. We have been able to distinguish between processing of images that contain a face from images with other objects [65]. We have also used linear classification to detect when a subject perceived to have made a mistake in a rapid perceptual decision task [66]. In Section IV, we will discuss in detail how linear classification in EEG has been used to identify the time course and spatial distribution of activity associated with perceptual decision-making.

E. Linear Versus Nonlinear Features in EEG/MEG

It is surprising that so much can be accomplished by linearly combining the raw magnitude of fMRI or EEG signals. Basic pre-possessing methods, of course, are beneficial.[3] However, in EEG and MEG, the time axis does require special consideration, resulting in the use of nonlinear features. In these modalities, the temporal accuracy of the signal is better than the temporal variability of the underlying neural processing. Thus, aligning the time samples from one trial to the next is problematic. This can be addressed to some degree by locking activity to different external events such as the time of stimulus presentation or the time of a behavioral response (see Section IV for an example of this). Alternatively, one can aim to find a signal that is not locked in time. By this, we mean that the signal has no consistent sign but instead oscillates at some frequency with arbitrary phase. Therefore, instead of analyzing magnitude, a linear property of the signal, one will analyze power, a second-order feature of the signal.

Powers—or, more often, log-powers—are computed using Fourier or wavelet analysis. Essentially, the samples in time are converted into time-frequency samples, resulting typically in the same number of independent degrees of freedom. Traditional analysis has identified various relevant frequency bands such as the alpha band (8–12 Hz) and the gamma band (25–50 Hz), among others. In EEG, the power in the alpha band is very pronounced, and it has been found to decrease with the active engagement in a task and increase with active inhibition of unattended stimuli [67]. Single-trial classification of power in the alpha band has been used successfully to control the movement of a pointer on the screen in two dimensions (left versus right and up versus down) [48]. Gamma-band activity has often been associated with perceptual processing [68], but gamma oscillations are often transient and more localized in space, and thus harder to detect.[4]

---

[3]In fMRI and MEG, usually activity is first corrected to account for head motion and in fMRI activity is sometimes filtered in time to account for the slow hemodynamic response. For EEG, the data are filtered to remove slow drifts (<1 Hz) and power-line inductive noise (60/50 Hz).

[4]Fast oscillations are spatially more localized, presumably because fast firing can only be coherent for smaller neuronal populations. Thus the approach of spatial averaging as outlined here may not improve SNR significantly. Gamma activity is typically also more transient in time, and thus trial averaging will require many more samples unless the transient oscillations are induced reliably at a given point in time.

Other nonlinear features include autoregressive (AR) model parameters that capture the shape of the power spectrum. AR coefficients are popular in brain–computer interface (BCI) research, as they give a low-dimensional parametrization of the entire spectrum without committing to any preconceived notion of a relevant frequency band. AR coefficients have often been used to classify EEG signals into more than two categories by comparing the coefficients to prototypes from various categories [49]. EEG signals have also been classified into more than two categories (in this case, spoken words) by reducing the time course of the evoked EEG activity into its strongest sinusoidal components [69]. Other popular nonlinear features are various coherence measures, power modulation, and more fanciful nonlinear features, such as various entropy and information measures, fractal dimension, Lyapunov exponents, etc. While some of these parametrizations are rather popular, in particular AR coefficients, they have not found the universal applicability of simple power measures of various frequency bands.

F. Decoding Using a Forward Model

Direct decoding as outlined above is often limited by the number of categories they can recover—in most cases, just two. To study cognition, it may be necessary to identify more than binary categories, and certainly in perception, continuous stimulus properties are of great importance. An alternative approach to direct decoding is to build a "forward model," that is, a model that predicts brain activity that will be evoked by a given stimulus. One can then compare the predicted activity to that previously generated by various stimuli and identify the stimulus as the one that is closest to the prediction.

To identify a significant number of categories with this approach, one requires techniques that can predict brain signals with sufficient accuracy. In fMRI, this approach has allowed identification of thousands of categories [59]. However, this has required significant knowledge of the stimulus domain, which is perhaps not surprising, as one has to replicate the highly nonlinear mapping from the stimulus domain to the response of the brain. Typically, stimuli s_j have to be decomposed into relevant features—say, $\phi_m(s_j)$ represents the mth feature for the jth stimulus. In [70], for instance, images are decomposed into local signal power at multiple scales and multiple orientations. These features are then combined linearly to predict the activity of individual voxels in the brain

$$x_{ij} = \sum_m a_{im} \phi_m(s_j). \quad (2)$$

Aside from reconstructing the activity associated with images, this technique has also been used to predict brain activity associated with the meaning of nouns [71]. To do so, noun words were represented by 25 semantic features quantifying the relationship of the word to various senses such as "see," "hear," "taste" or actions such as "run," "push," "enter," and so on). In this case, the feature value quantified the average distance of the stimulus word to the feature word within a large corpus of written text. With this technique, it was possible to identify 60 different new nouns such as "airplane" or "celery." Evidently, the feature set used here is extremely complex and required

a strong hypothesis as to the relevant feature dimensions. The value of the prediction approach outlined here lies precisely in its ability to validate scientific hypotheses as to the relevant dimensions the brain uses to encode information.

G. Regularization and Generalization

A central challenge of the decoding and prediction problems in (1) and (2) is to constrain the large number of free coefficients w_i and a_{im}. Regularization refers to the general problem of constraining or reducing the effective number of degrees of freedom in a parametrization or regression task.[5] Without regularization, the degrees of freedom are severely underconstrained given the large size of the feature space, $D = 10^4$ or more, and the small number exemplars, typically $N = 10^2$ or less. Various techniques for regularization have been developed in the machine-learning community to address this problem, and they will only be broadly categorized here. All techniques essentially rely on various prior assumptions about the classification or prediction problem.

1) Finite Noise: Finding a good classifier or predictor often involves minimizing a cost function, which quantifies the errors in prediction or classification. Regularization often involves adding to this cost function an extra "penalty term," which gives preference to some parameter values over others. For instance, a quadratic penalty term $L_2 = \sum_i |w_i|^2$ will assign a cost to the nonzero coefficient, so that reducing the cost will reduce the coefficient to zero unless the training data provide sufficient evidence that a specific coefficient is really required. Penalized logistic regression [56] and ridge regression [72] use this technique. L_2 quantifies the norm of the weight vector; hence it is also called L_2 norm regularization. In LR, the L_2 norm regularization is particularly useful: LR models the transition from one class to the other as a sigmoid logistic function. The norm of the weight vector corresponds to the sharpness of this transition boundary. Thus, penalizing the norm limits the sharpness of the transition boundary. This, in effect, implements the assumption that the data carry some noise even in the case that the training samples are perfectly separable (which is always the case because $D > N$).

2) Smoothness: If one can assume that the activity in neighboring samples is similar, then it is reasonable to assume that the corresponding weights should be similar as well. It is easy to incorporate such neighborhood relationships into a quadratic penalty term by defining the corresponding expected covariance matrix for the weights as R_{ww} and then incorporating this covariance in the norm as $L_2 = w^T R_{ww}^{-1} w$ (see, e.g., [73]).

3) Sparseness: The problem with a quadratic penalty term, however, is that large coefficients are penalized more than small coefficients, and so the solutions will favor a distributed representation with many small values. Alternatively, one can use $L_1 = \sum_i |w_i|$ as penalty, which implicitly assumes that weights are sparse, i.e., most weights are zero or very small with a few

---

[5]Note the opposing uses of this word. In statistical testing, "degrees of freedom" refers to the number of independent measurements minus the number of parameters that have been fit to the data. Instead, in the context of parameter fitting and machine-learning, "degrees of freedom" stands for the number of free parameters that are to be constrained by the measurements. We prefer the latter use of the word.

very large values. Sparse priors and LASSO techniques [74] use this approach. This will encourage extraction of more localized activity and will ignore samples that do not contribute significant predictive/discriminant activity. Smoothness constraints are difficult to implement for sparse priors. Instead, in cases where different samples are expected to behave similarly, one can enforce constrain using sparse penalty terms that apply to groups of coefficients together (such as the group-LASSO technique [75] or elastic nets [76] with combine L1 and L2 norms). One should note that various feature selection algorithm exist that conceptually aim to do the same thing as sparse priors, i.e., select a few dimensions that will contribute significantly to the classification while giving all other features zero weight. We recommend to use sparse priors as a more systematic and modern approach to perform feature selection.

4) Factorized Parametrization: An alternative approach to regularization is to parametrize the weight vector with a smaller number of free parameters. For instance, one might assume that in EEG or MEG, the time and space axis are independent. If that is the case, then the weight vector can be factorized as a bilinear model with $T + K$ parameters instead of TK (where T is the number of samples in time and K the number of electrodes) [73]. If there are more than two dimensions to organize the weight vector (for instance, there may be an additional frequency dimension), one can use more complex factorizations such as the PARAFAC model [77]. These models can lead to a drastic reduction in the number of degrees of freedom, thus reflecting rather strong assumptions on the parameter space that have to be well justified.

5) Spatial Components: An additional approach to reducing the dimensionality of the problem is to combine spatial dimensions into a subspace and then perform linear classification or prediction within that subspace. The most notable of these approaches are common spatial pattern (CSP), independent component analysis (ICA), and principal component analysis (PCA). CSP combines electrodes linearly such that the resulting aggregate signals have maximal or minimal power for two conditions, respectively [56]. These powers are then used as features for classification. CSP therefore uses information about the two different classes to combine electrodes into some optimal spatial components. ICA and PCA, in contrast, combine signals linearly without class information based on other statistical criteria [56]. These techniques may be useful in neuroimaging studies that have an insufficient number of labels or simply lack labels for categorization.

IV. IDENTIFYING THE NEURAL CORRELATES OF DECISION-MAKING USING MULTIVARIATE SINGLE-TRIAL ANALYSIS OF EEG AND/OR FMRI

A. Single-Trial EEG Reveals Temporal Evolution of Decision-Making

One of the first studies to use single-trial analysis of the EEG to explore the temporal characteristics of perceptual decision-making in humans was by Philiastides and Sajda [65]. Motivated from previous work in primates [4], [7], we quantified the relationship between neural activity and behavioral output during a simple "face"/"car" decision-making task. The difficulty of the task was manipulated by changing the spatial phase coherence of the stimuli in a range that spanned psychophysical threshold.

Using the LR approach (see Section III-C) for different time windows and phase coherence levels, we identified two EEG components whose single-trial amplitudes discriminated maximally between faces and cars [Fig. 2(a)]. The early component was consistent with the well-known face-selective N170 [78]–[82], and its temporal onset appeared to be unaffected by task difficulty. The late component (>300 ms after the stimulus) systematically shifted later in time and became more persistent (i.e., broader) as a function of task difficulty. Both of these components indexed decision accuracy in that a high magnitude discriminator output value indicated an easy trial, while values near zero indicated more difficult decisions.

To compare neuronal to psychophysical performance, we constructed neurometric functions by considering the data in the early and late time windows. Specifically, receiver operating characteristic (ROC) analysis was used to quantify the discriminator's performance at each phase coherence level. The area under the ROC curves was plotted against the corresponding phase coherence levels to construct the neurometric functions. We showed that the neurometric curves were good predictors of behavioral performance as captured by the psychometric function [Fig. 2(b)]. Neurometric functions from the late component were a better match to the psychophysical data than those from the early one. Choice probability analysis [7] also revealed that the late component was a significant predictor of the content of the subject's final decision.

Situated somewhere between the early and late components (≈220 ms poststimulus), there was a third component, the strength of which systematically increased with increasing task difficulty. This component was a good predictor of the onset time of the late component. We originally speculated that this component reflects a top-down influence of attention on decision-making rather than a mere bottom-up processing of the stimulus evidence. To substantiate this claim, we ran a variant of the original behavioral paradigm, where the same stimuli were colored red or green and the subjects were either cued to perform a "green"/"red" discrimination or the original "face"/"car" discrimination [83]. This manipulation allowed us to vary the difficulty of the task while leaving the stimulus evidence unchanged in order to treat confounding factors relating to early, bottom-up, processing of the stimulus. For example, for images with the same overall phase coherence level, we compared the single-trial amplitudes of the difficulty component during challenging face/car decisions with those of trivial color decisions. We found that the trial-to-trial amplitudes of the difficulty component were significantly reduced when the subjects were simply discriminating the color of the stimulus.

This variant of the experiment has also yielded additional evidence on the role of the early and late components. We showed that while the early component remained unaffected by task demands (in that the single-trial amplitudes of this component to face versus car stimuli remained unchanged during color discrimination), the late component was largely diminished when subjects were making a color decision. More recently, we demonstrated that the late component can also
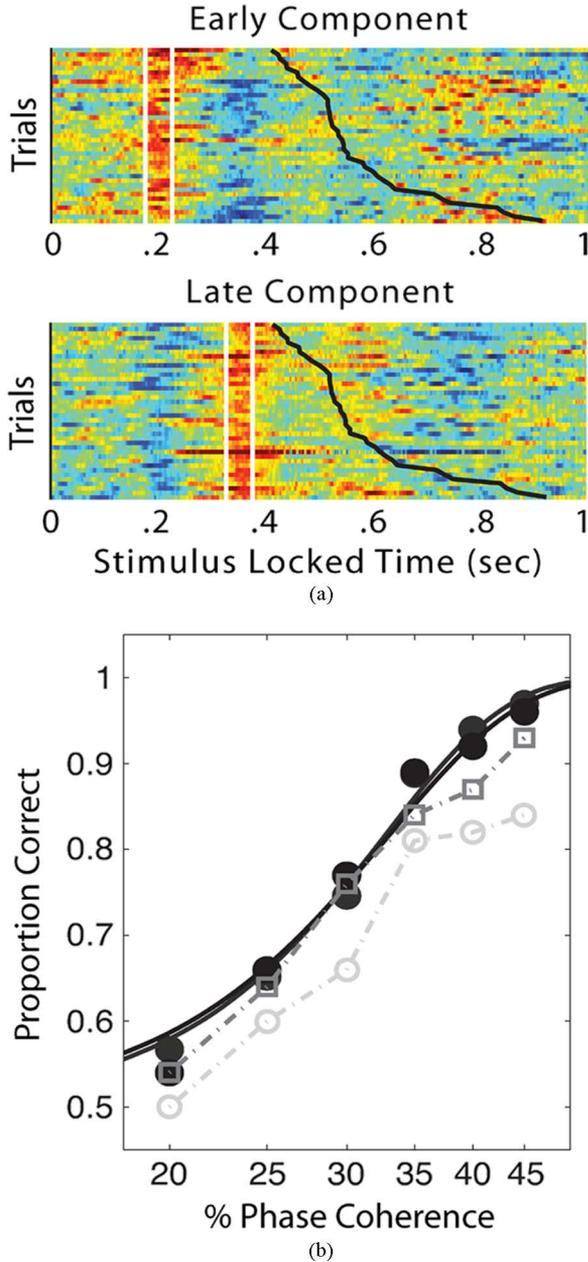
Fig. 2. Single-trial EEG components correlating with decision accuracy on our "face"/"car" categorization task. (a) Discriminant component maps for the early (≈ 170 ms) and late (≈ 300 ms) components. Trials are aligned to the onset of visual stimulation (time 0 ms) and sorted by response time (black sigmoidal curves). Only face trials are shown here. Each row in these maps represents the discriminator output for a single face trial with the mean of all car trials subtracted (i.e., $\mathbf{y}_{\text{face}\_i} - \bar{\mathbf{y}}_{\text{cars}}$). (b) Comparison of one subject's psychometric function (gray, solid line) with neurometric functions obtained from the early component (light gray dotted line), the late component (dark gray dotted line), and a combination of the two (black solid line).

be used to index the quality of evidence used in the decision process as quantified by the drift rate parameter in a diffusion model simulation [84]. Specifically, the results showed that dividing the data from nominally identical stimuli on the basis of the late-component single-trial amplitudes (but not those of the early one) produced differences in the drift-rate estimates. Though trial-to-trial variability in drift rate is often assumed, prior to this work, there has been no independent way to measure its neurophysiological correlates.

B. EEG-Informed fMRI Reveals Spatiotemporal Characteristics of Decision-Making

Despite significant progress made in understanding perceptual decision-making in humans using EEG and fMRI in isolation, the spatial localization restrictions of EEG and the temporal resolution constraints of fMRI suggest that only a combination of these modalities can provide a full spatiotemporal characterization of this process.

One way this fusion could be achieved is by simultaneous EEG/fMRI measurements, where trial-to-trial variability in EEG components can be used to construct parametrically modulated fMRI regressors. Major technical challenges for simultaneous acquisition include 1) removal of large magnetic field gradients and radio-frequency (RF) pulses used to produce the MR images from the EEG [85], 2) special EEG amplifier design to remove the dc components without allowing the gradients to saturate the input stage [86], 3) novel EEG electrode design to minimize artifact formation [85], [87], 4) removal of cardiac-related artifacts (ballistocardiogram) [85], [88], and 5) removal of motion artifacts in the EEG, which are usually amplified when subjects are placed in an MR scanner [89]. Our group has overcome most of the technical difficulties outlined above and has been able to develop a truly simultaneous EEG and fMRI recording system [54], [90]–[94], which includes novel signal processing for artifact removal [95] and a discriminant-based multivariate analysis framework for integrating single-trial variability of EEG with fMRI [54].

Specifically, we perform single-trial discrimination to identify task-relevant components $\mathbf{y}_\tau$, with τ indexing the time window of interest [Fig. 3(a)]. The discriminator output will have dimensions $T \times N$, where T is the total number of training samples and N the total number of trials. Finally, to achieve more robust single-trial estimates for $\mathbf{y}_\tau$, we typically average across all training samples to obtain

$$\bar{\mathbf{y}}_{\tau\_i} = \frac{1}{T} \sum_{j=1}^{T} \mathbf{y}_{\tau\_ij} \tag{3}$$

where i is used to index trials and j training samples. Hypothetical single-trial discriminator amplitudes can be seen in Fig. 3(b). $\bar{\mathbf{y}}_\tau$ can then be used to construct parametric fMRI regressors for each component of interest [Fig. 3(c)]. These in turn are convolved with a prototypical hemodynamic response function [Fig. 3(d)] in order to be used as fMRI data predictors in the context of a general linear model (GLM). Identifying the brain regions that correlate with each of these regressors can enable a more comprehensive characterization of the cortical network involved in different neurocognitive processes. We have already evaluated our system for yielding unique and meaningful single-trial activations in a simple auditory oddball task [54] and are currently in the process of using the system for perceptual decision-making paradigms.

In the absence of simultaneous EEG/fMRI measurements, an EEG-informed fMRI approach can be used instead. In this case, the neural characterization proceeds in two steps. Initially, an EEG study is performed to identify components of interest. The

(a)

(b)

(c)

(d)
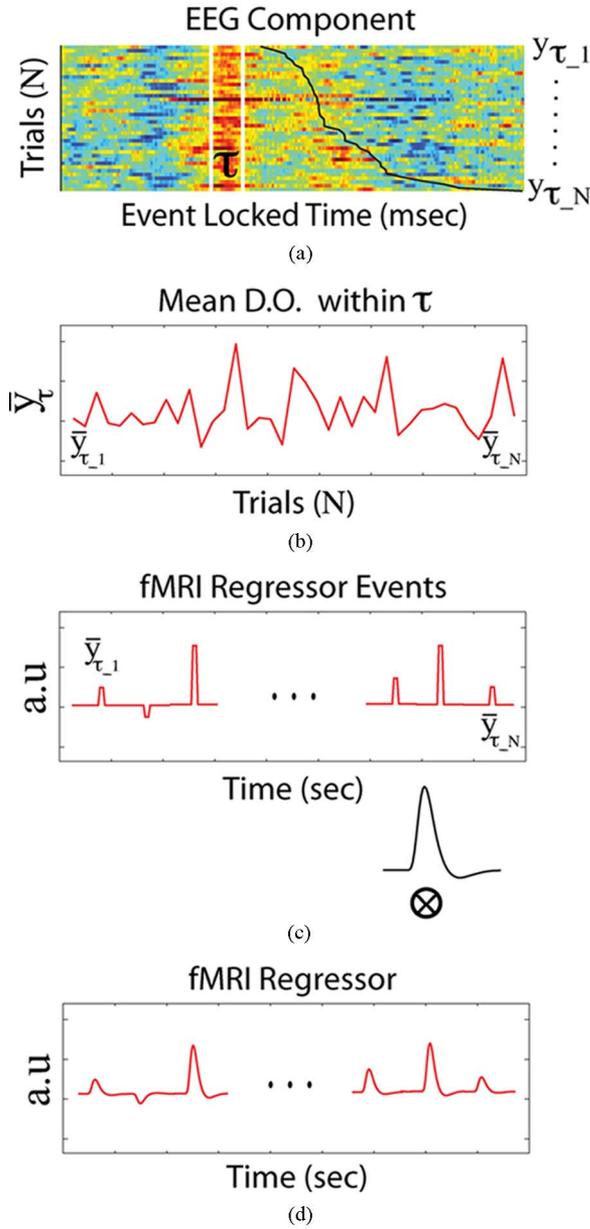
Fig. 3. EEG-derived fMRI regressors. (a), (b) The discriminator output (D.O.) is to used to acquire single-trial information for several temporally distinct EEG components. (c) The onset time and duration of each of the regressor events are determined by the onset time (τ) and duration (δ) of the EEG components. The amplitude of each regressor event will be based on the output of the linear discriminator $\mathbf{y}_\tau$ as defined in (3). (d) EEG-derived regressors are convolved with a canonical hemodynamic response function prior to the GLM analysis.

discriminator output associated with each component and each experimental condition is then averaged across trials

$$\bar{y}_\tau^c = \frac{1}{N}\frac{1}{T}\sum_{i=1}^{N}\sum_{j=1}^{T}\mathbf{y}_{\tau\text{-}ij}^c \qquad (4)$$

where once again τ is used to index the component (time) of interest and c is used to index the different experimental conditions. In a second step, the same experiment is repeated while fMRI data are collected. The average discriminator output per component and experimental condition obtained from (4) will
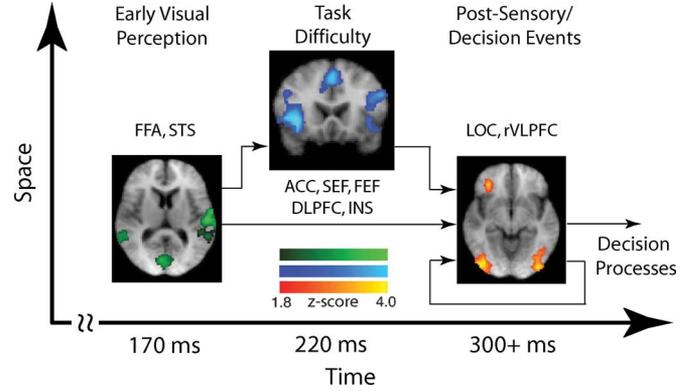
Fig. 4. Spatiotemporal processing timing diagram resulting from an EEG-informed fMRI analysis. Color bar represents z-scores for each signature (corrected for multiple comparisons). (Adapted from [102].)

now be used to model the fMRI data. Importantly, $\bar{y}_\tau^c$ is now a scalar—that is, in the absence of single-trial information during the fMRI session, all like trials will be modeled in the same way. Though intertrial variability is ultimately concealed in this formulation, important information regarding the localization of each of the EEG components, that would otherwise be unattainable using EEG or fMRI alone, can now be obtained.

To demonstrate the efficacy of this approach, we applied this methodology for the perceptual decision-making work presented in the previous section, where different EEG components were found to respond uniquely to different experimental manipulations/conditions. That is, all EEG-derived regressors from (4) were uncorrelated, a requirement that every sensible fMRI experimental design should satisfy.

As highlighted in the previous section, the strength of our early EEG component was proportional to the stimulus evidence (i.e., stronger for easy than hard trials), and it remained unchanged during the face/car and color discriminations. The late EEG component also responded proportionally to the stimulus evidence during the face/car discrimination, but it was stronger across all difficulty levels relative to the early one. Unlike the early component, however, it was virtually eliminated during the color discrimination. In contrast to both the early and late components, the strength of the difficulty component was inversely proportional to the amount of stimulus evidence (i.e., stronger for hard than easy trials).

As a result of these observations, we repeated this experiment in the scanner; and using (4), we constructed three parametric fMRI regressors, one for each of the early, difficulty, and late components. To modulate the heights of the corresponding regressor events, we estimated the relative strengths of our components with respect to the difficulty (i.e., low [L] versus high [H] coherence) and the type of task (i.e., face versus car [FC] or red versus green [RG]) (i.e., $\bar{y}_\tau^{FC,L}$, $\bar{y}_\tau^{FC,H}$, $\bar{y}_\tau^{RG,L}$, $\bar{y}_\tau^{RG,H}$, $\tau = \{\text{early,difficulty,late}\}$).

Fig. 4 summarizes our results. For the early component, we identified significant correlations with activity in areas implicated in early visual processing of objects/faces such as the fusiform face area (FFA), the superior temporal sulcus (STS). Both the FFA and STS have previously been implicated in early

visual processing of faces using neuroimaging [34], [96]–[98] and field potentials recorded directly from the cortical surface [99]–[101]. These results implicate the early component in early visual processing of the incoming sensory evidence.

The difficulty component, correlated with activity in brain regions that are typically associated with the human attentional network such as the supplementary and frontal eye fields, the anterior cingulate cortex, the DLPFC, and the anterior insula [1], [32]. These observations support the hypothesis that there exists an attentional control system that exerts top-down influence on decision-making.

Finally, we found activations correlating with the late component in the lateral occipital complex (LOC) and in the right ventrolateral prefrontal cortex (rVLPFC). Aside from its involvement in object categorization [103]–[107], the LOC has been implicated in "perceptual persistence" [108], [109], a process in which a percept assembled by lower visual areas is allowed to remain in the visual system, via feedback pathways, as a form of iconic memory [110]–[112].

The brief stimulus durations used in these experiments suggest that perceptual persistence is a likely mechanism by which rapid object decision-making takes place. That is, for brief presentations, the accumulation of evidence is not based on the decaying stimulus traces themselves but rather on a durable representation of the stimulus retained in short-term memory. This interpretation explains why the late component was a better predictor of overall behavioral performance than the early one, why it correlated strongly with drift rate in a diffusion model simulation, and why it disappeared when a demanding face versus car discrimination was no longer required (e.g., during color discrimination).

Taken together, these results reaffirm that the EEG-informed fMRI approach is a promising new tool in mapping out the spatiotemporal characteristics of different neurocognitive processes, such as perceptual decision-making, in humans.

V. Conclusion

Though perceptual decision-making is perhaps rather simplistic relative to the complex decisions we make everyday, understanding the neural processes governing even the most simple decisions will shed light on how we make decisions which involve context, reward and value. It is clear, however, that a marriage of tools and methods from both neuroscience and engineering is required in that this undertaking involves analysis of a tremendous amount of high-dimensional data, systems for acquiring the data in real-time, and sophisticated and insightful paradigms for manipulating our behavior and interpreting the results. In summary, we are just at the beginning of having the necessary tools and techniques to develop a deep understanding of how we make a decision—for example, on whether to read this paper or not.

References

[1] H. R. Heekeren, S. Marrett, and L. Ungerleider, "The neural systems that mediate human perceptual decision making," *Nat. Rev. Neurosci.*, vol. 9, no. 6, pp. 467–479, Jun. 2008.

[2] J. I. Gold and M. Shadlen, "The neural basis of decision making," *Annu. Rev. Neurosci.*, vol. 30, pp. 535–574, 2007.

[3] R. Romo and E. Salinas, "Flutter discrimination: Neural codes, perception, memory and decision making," *Nat. Rev. Neurosci.*, vol. 4, no. 3, pp. 203–218, 2003.

[4] K. H. Britten, M. N. Shadlen, W. T. Newsome, and J. A. Movshon, "The analysis of visual motion: A comparison of neuronal and psychophysical performance," *J. Neurosci.*, vol. 12, no. 12, pp. 4745–4765, 1992.

[5] K. H. Britten, M. N. Shadlen, W. T. Newsome, and J. A. Movshon, "Responses of neurons in macaque MT to stochastic motion signals," *Vis. Neurosci.*, vol. 10, no. 12, pp. 4745–4765, 1993.

[6] M. N. Shadlen and W. T. Newsome, "Motion perception: Seeing and deciding," *Proc Nat. Acad. Sci.*, vol. 93, pp. 628–633, 1996.

[7] K. H. Britten, W. T. Newsome, M. N. Shadlen, S. Celebrini, and J. A. Movshon, "A relationship between behavioral choice and visual responses of neurons in macaque MT," *Vis. Neurosci.*, vol. 14, pp. 87–100, 1996.

[8] C. D. Salzman, K. H. Britten, and W. T. Newsome, "Cortical microstimulation influences perceptual judgements of motion direction," *Nature*, vol. 346, no. 6280, pp. 174–177, Jul. 1990.

[9] C. D. Salzman, C. M. Murasugi, K. H. Britten, and W. T. Newsome, "Microstimulation in visual area mt: Effects on direction discrimination performance," *J. Neurosci.*, vol. 12, no. 6, pp. 2331–2355, Jun. 1992.

[10] J. Ditterich, M. E. Mazurek, and M. Shadlen, "Microstimulation of visual cortex affects the speed of perceptual decisions," *Nat. Neurosci.*, vol. 6, no. 8, pp. 891–898, Aug. 2003.

[11] C. Keysers, D.-K. Xiao, P. Foldiak, and D. I. Perrett, "The speed of sight," *J. Cogn. Neurosci.*, vol. 13, no. 1, pp. 90–101, Jan. 2001.

[12] S. R. Afraz, R. Kiani, and H. Esteky, "Microstimulation of inferotemporal cortex influences face categorization," *Nature*, vol. 442, no. 7103, pp. 692–695, August 2006.

[13] E. Salinas, A. Hernandez, A. Zainos, and R. Romo, "Periodicity and firing rate as candidate neural codes for the frequency of vibrotactile stimuli," *J. Neurosci.*, vol. 20, no. 14, pp. 5503–5515, Jul. 2000.

[14] R. Romo, A. Hernandez, A. Zainos, and E. Salinas, "Somatosensory discrimination based on cortical microstimulation," *Nature*, vol. 392, no. 6674, pp. 387–390, March 1998.

[15] R. Romo, A. Hernandez, A. Zainos, C. D. Brody, and L. Lemus, "Sensing without touching: Psychophysical perfrormance based on cortical microstimulation," *Neuron*, vol. 26, no. 1, pp. 273–278, Apr. 2000.

[16] R. Ratcliff, "A theory of memory retrieval," *Psychol. Rev.*, vol. 85, no. 2, pp. 59–108, 1978.

[17] R. Ratcliff and F. Tuerlinckx, "Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameters variability," *Psychon. Bull. Rev.*, vol. 9, no. 3, pp. 438–481, 2002.

[18] M. E. Mazurek, J. D. Roitman, J. Ditterich, and M. N. Shadlen, "A role for neural integrators in perceptual decision making," *Cereb. Cortex*, vol. 13, no. 11, pp. 1257–1269, Nov. 2003.

[19] R. Bogacz, "Optimal decision making theories: Linking neurobiology with behaviour," *Trends. Cogn. Sci.*, vol. 11, no. 3, pp. 118–125, Mar. 2007.

[20] S. Grossberg and P. K. Pilly, "Temporal dynamics of decision-making during motion perception in the visual cortex," *Vision Res.*, vol. 48, no. 12, pp. 1345–1373, Jun. 2008.

[21] M. N. Shadlen and W. T. Newsome, "Neural basis of perceptual decision making in the parietal cortex (area LIP) of the rhesus monkey," *J. Neurophysiol.*, vol. 86, no. 4, pp. 1916–1936, Oct. 2001.

[22] J. D. Roitman and M. N. Shadlen, "Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task," *J. Neurosci.*, vol. 22, no. 21, pp. 9475–9489, Nov. 2002.

[23] A. C. Huk and M. N. Shadlen, "Neural activity in the macaque parietal cortex reflects temporal integration of visual motion signals during perceptual decision making," *J. Neurosci.*, vol. 25, no. 45, pp. 10420–10436, Nov. 2005.

[24] T. D. Hanks, J. Ditterich, and M. Shadlen, "Microstimulation of macaque area lip affects decision-making in a motion discrimination task," *Nat. Neurosci.*, vol. 9, no. 5, pp. 682–689, May 2006.

[25] A. K. Churchland, R. Kiani, and M. N. Shadlen, "Decision-making with multiple alternatives," *Nat. Neurosci.*, vol. 11, no. 6, pp. 693–702, Jun. 2008.

[26] J. N. Kim and M. N. Shadlen, "Neural correlates of decision making in the dorsolateral prefrontal cortex of the macaque," *Nat. Neurosci.*, vol. 2, no. 2, pp. 176–185, Feb. 1999.

[27] R. Romo, C. Brody, A. Hernandez, and L. Lemus, "Neuronal correlates of parametric working memory in prefrontal cortex," *Nature*, vol. 399, no. 6735, pp. 470–473, June 1999.

[28] R. Romo, A. Hernandez, A. Zainos, L. Lemus, and C. D. Brody, "Neuronal correlates of decision-making in secondary somatosensory cortex," *Nat. Neurosci.*, vol. 5, no. 11, pp. 1217–1224, Nov. 2002.

[29] R. Romo, A. Hernandez, and A. Zainos, "Neuronal correlates of decision-making in ventral premotor cortex," *Neuron*, vol. 41, no. 1, pp. 165–173, Jan. 2004.

[30] A. Hernandez, A. Zainos, and R. Romo, "Temporal evolution of a decision-making process in medial premotor cortex," *Neuron*, vol. 33, no. 6, pp. 959–972, Mar. 2002.

[31] C. D. Brody, A. Hernandez, A. Zainos, and R. Romo, "Timing and neural encoding of somatosensory parametric working memory in macaque prefrontal cortex," *Cereb. Cortex*, vol. 13, no. 11, pp. 1196–1207, Nov. 2003.

[32] H. R. Heekeren, S. Marrett, P. A. Bandettini, and L. G. Ungerleider, "A general mechanism for perceptual decision-making in the human brain," *Nature*, vol. 431, pp. 859–862, 2004.

[33] J. V. Haxby, "The functional organization of human extrastriate cortex: A PET-rCBF study of selective attention to faces and locations," *J. Neurosci.*, vol. 14, pp. 6336–6353, 1994.

[34] N. Kanwisher, J. McDermott, and M. M. Chun, "The fusiform face area: A module in human extrastriate cortex specialized for face perception," *J. Neurosci.*, vol. 17, pp. 4302–4311, 1997.

[35] G. McCarthy, A. Puce, J. Gore, and T. Allison, "Face-specific processing in the human fusiform gyrus," *J. Cogn. Neurosci.*, vol. 9, pp. 605–611, 1997b.

[36] R. Epstein and N. Kanwisher, "A cortical representation of the local visual environment," *Nature*, vol. 392, pp. 598–601, 1998.

[37] A. Ishai, L. G. Ungerleider, A. Martin, J. L. Schouten, and J. V. Haxby, "Distributed representation of objects in the human ventral visual pathway," *Proc. Nat. Acad. Sci.*, vol. 96, pp. 9379–9384, 1999.

[38] B. Pleger, C. C. Ruff, F. Blankenburg, S. Bestmann, K. Wiech, K. E. Stephan, A. Capilla, K. J. Friston, and R. J. Dolan, "Neural coding of tactile decisions in the human prefrontal cortex," *J. Neurosci.*, vol. 26, no. 48, pp. 12596–12601, Nov. 2006.

[39] C. Preuschhof, H. R. Heekeren, B. Taskin, T. Schubert, and A. Villringer, "Neural correlates of vibrotactile working memory in the human brain," *J. Neurosci.*, vol. 26, pp. 13231–13239, 2006.

[40] E. J. Ploran, S. M. Nelson, K. Velanova, D. I. Donaldson, S. E. Petersen, and M. E. Wheeler, "Evidence accumulation and the moment of recognition: Dissociating perceptual recognition processes using fmri," *J. Neurosci.*, vol. 27, no. 44, pp. 11912–11924, Oct. 2007.

[41] A. E. Rorie and W. T. Newsome, "A general mechanism for perceptual decision-making in the human brain?," *Trends Cogn. Sci.*, vol. 9, pp. 41–43, 2005.

[42] H. R. Heekeren, S. Marrett, D. A. Ruff, P. A. Bandettini, and L. G. Ungerleider, "Involvement of human left dorsolateral prefrontal cortex in perceptual decision making is independent of response modality," *Proc. Nat. Acad. Sci.*, vol. 103, no. 26, pp. 10023–10028, 2006.

[43] W. Maass, T. Natschlager, and H. Markram, "Real-time computing without stable states: A new framework for neural computation based on perturbations," *Neural Comput.*, vol. 14, no. 11, pp. 2531–2560, Nov. 2002.

[44] H. Jaeger, W. Maass, and J. Principe, "Special issue on echo state networks and liquid state machines," *Neural Netw.*, 2007.

[45] D. Buonomano and W. Maass, "State-dependent computations: Spatiotemporal processing in cortical networks," *Nature Rev. Neurosci.*, vol. 10, no. 2, pp. 113–125, 2009.

[46] J.-D. Haynes and G. Rees, "Decoding mental states from brain activity in humans," *Nat. Rev. Neurosci.*, vol. 7, no. 7, pp. 523–534, Jul. 2006.

[47] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," *Clin. Neurophysiol.*, vol. 113, no. 6, pp. 767–791, Jun. 2002.

[48] J. R. Wolpaw and D. J. McFarland, "Control of a twodimensional movement signal by a noninvasive brain-computer interface in humans," *PNAS*, vol. 101, no. 51, pp. 17849–17854, Dec. 2004.

[49] G. Pfurtscheller and C. Neuper, "Motor imagery and direct brain-computer communication," *Proc. IEEE*, vol. 89, pp. 1123–1134, Jul. 2001.

[50] R. A. Andersen, J. W. Burdick, S. Musallam, B. Pesaran, and J. G. Cham, "Cognitive neural prosthetics," *Trends Cogn. Sci.*, vol. 8, no. 11, pp. 486–493, Nov. 2004.

[51] N. Birbaumer, "Brain-computer-interface research: Coming of age," *Clin. Neurophysiol.*, 2006.

[52] L. C. Parra, C. Christoforou, A. D. Gerson, M. Dyrholm, A. Luo, M. Wagner, M. G. Philiastides, and P. Sajda, "Spatiotemporal linear decoding of brain state: Application to performance augmentation in high-throughput tasks," *IEEE, Signal Process. Mag.*, vol. 25, pp. 95–115, Jan. 2008.

[53] A. D. Gerson, L. C. Parra, and P. Sajda, "Cortical origins of response time variability during rapid discrimination of visual objects," *Neuroimage*, vol. 28, no. 2, pp. 342–353, 2005.

[54] R. I. Goldman, C.-Y. Wei, M. G. Philiastides, A. D. Gerson, D. Friedman, T. R. Brown, and P. Sajda, "Single-trial discrimination for integrating simultaneous EEG and fMRI: Identifying cortical areas contributing to trial-to-trial variability in the auditory oddball task," *Neuroimage*, vol. 47, no. 1, pp. 136–147, Aug. 2009.

[55] U. Hasson, Y. Nir, I. Levy, G. Fuhrmann, and R. Malach, "Intersubject synchronization of cortical activity during natural vision," *Science*, vol. 303, no. 5664, pp. 1634–1640, Mar. 2004.

[56] L. C. Parra, C. D. Spence, A. D. Gerson, and P. Sajda, "Recipes for the linear analysis of EEG," *Neuroimage*, vol. 28, no. 2, pp. 326–341, 2005.

[57] K. A. Norman, S. M. Polyn, G. J. Detre, and J. V. Haxby, "Beyond mind-reading: Multi-voxel pattern analysis of fmri data," *Trends Cogn. Sci.*, vol. 10, no. 9, pp. 424–430, Sep. 2006.

[58] E. Formisano, F. De Martino, and G. Valente, "Multivariate analysis of fmri time series: Classification and regression of brain responses using machine learning," *Magn. Res. Imag.*, vol. 26, no. 7, pp. 921–934, Sep. 2008.

[59] Y. Miyawaki, H. Uchida, O. Yamashita, M. A. Sato, Y. Morito, H. C. Tanabe, N. Sadato, and Y. Kamitani, "Visual image reconstruction from human brain activity using a combination of multiscale local image decoders," *Neuron*, vol. 60, no. 5, pp. 915–929, Dec. 2008.

[60] D. Hassabis, C. Chu, G. Rees, N. Weiskopf, P. D. Molyneux, and E. A. Maguire, "Decoding neuronal ensembles in the human hippocampus," *Curr. Biol.*, vol. 19, no. 7, pp. 546–554, Apr. 2009.

[61] S. V. Shinkareva, R. A. Mason, V. L. Malave, W. Wang, T. M. Mitchell, and M. A. Just, "Using fMRI brain activation to identify cognitive states associated with perception of tools and dwellings," *PLoS One*, vol. 3, no. 1, p. e1394, 2008.

[62] J.-D. Haynes, K. Sakai, G. Rees, S. Gilbert, C. Frith, and R. E. Passingham, "Reading hidden intentions in the human brain," *Curr. Biol.*, vol. 17, no. 4, pp. 323–328, Feb. 2007.

[63] C. S. Soon, M. Brass, H.-J. Heinze, and J.-D. Haynes, "Unconscious determinants of free decisions in the human brain," *Nat. Neurosci.*, vol. 11, no. 5, pp. 543–545, May 2008.

[64] A. D. Gerson, L. C. Parra, and P. Sajda, "Cortically-coupled computer vision for rapid image search," *IEEE Trans. Neural Syst. Rehab. Eng.*, vol. 14, pp. 174–179, Jun. 2006.

[65] M. G. Philiastides and P. Sajda, "Temporal characterization of the neural correlates of perceptual decision making in the human brain," *Cereb. Cortex*, vol. 16, no. 4, pp. 509–518, 2006.

[66] L. Parra, C. Alvino, A. Tang, B. Pearlmutter, N. Young, A. Osman, and P. Sajda, "Linear spatial integration for single-trial detection in encephalography," *Neuroimage*, vol. 17, pp. 223–230, 2002.

[67] W. Klimesch, P. Sauseng, and S. Hanslmayr, "Eeg alpha oscillations: The inhibition-timing hypothesis," *Brain Res. Rev.*, vol. 53, no. 1, pp. 63–88, Jan. 2007.

[68] E. Rodriguez, N. George, J. P. Lachaux, J. Martinerie, B. Renault, and F. J. Varela, "Perception's shadow: Long-distance synchronization of human brain activity," *Nature*, vol. 397, no. 6718, pp. 430–433, Feb. 1999.

[69] P. Suppes, Z. L. Lu, and B. Han, "Brain wave recognition of words," *Proc. Nat. Acad. Sci. USA*, vol. 94, no. 26, pp. 14965–14969, Dec. 1997.

[70] K. N. Kay, T. Naselaris, R. J. Prenger, and J. L. Gallant, "Identifying natural images from human brain activity," *Nature*, vol. 452, pp. 352–356, 2008.

[71] T. M. Mitchell, S. V. Shinkareva, A. Carlson, K.-M. Chang, V. L. Malave, R. A. Mason, and M. A. Just, "Predicting human brain activity associated with the meanings of nouns," *Science*, vol. 320, no. 5880, pp. 1191–1195, May 2008.

[72] A. E. Hoerl and R. W. Kennard, "Ridge regression: Applications to nonorthogonal problems," *Technometrics*, vol. 12, no. 1, pp. 69–82, Feb. 1970.

[73] M. Dyrholm, C. Christoforos, and L. C. Parra, "Bilinear discriminant component analysis," *J. Machine Learn. Res.*, vol. 8, pp. 1097–1111, 2007.

[74] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Royal Statist. Soc. B*, vol. 58, no. 1, p. 2670288, 1996.

[75] L. Meier, S. van de Geer, and P. Buhlmann, "The group lasso for logistic regression," *J. Royal Statist. Soc. B*, vol. 70, no. 1, pp. 53–71, 2008.

[76] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," *J. Royal Statist. Soc. B*, vol. 67, no. 2, pp. 301–320, 2005.

[77] M. Morup, L. K. Hansen, C. S. Hermann, J. Parnas, and S. M. Arnfred, "Parallel factor analysis as an exploratory tool for wavelet transformed event-related eeg," *NeuroImage*, vol. 29, no. 3, pp. 938–947, 2006.

[78] D. A. Jeffreys, "Evoked studies of face and object processing," *Vis. Cogn.*, vol. 3, pp. 1–38, 1996.

[79] S. Bentin, T. Allison, A. Puce, A. Perez, and G. McCarthy, "Electrophysiological studies of face perception in humans," *J. Cogn. Neurosci.*, vol. 8, pp. 551–565, 1996.

[80] E. Halgren, T. Raij, K. Marinkovic, V. Jousmaki, and R. Hari, "Cognitive response profile of the human fusiform face area as determined by MEG," *Cereb. Cortex*, vol. 10, pp. 69–81, 2000.

[81] J. Liu, M. Higuchi, A. Marantz, and N. Kanwisher, "The selectivity of the occipitotemporal M170 for faces," *Neuroreport*, vol. 11, pp. 337–341, 2000.

[82] B. Rossion, C. Joyce, G. W. Cottrell, and M. J. Tarr, "Early laterization and orientation tuning for face, word, object processing in the visual cortex," *Neuroimage*, vol. 20, pp. 1609–1624, 2003.

[83] M. G. Philiastides, R. Ratcliff, and P. Sajda, "Neural representation of task difficulty and decision-making during perceptual categorization: A timing diagram," *J. Neurosci.*, vol. 26, no. 35, pp. 8965–8975, Aug. 2006.

[84] R. Ratcliff, M. G. Philiastides, and P. Sajda, "Quality of evidence for perceptual decision making is indexed by trial-to-trial variability of the EEG," *Proc. Nat. Acad. Sci.*, vol. 106, pp. 6539–6544, 2009.

[85] J. R. Ives and S. Warach *et al.*, "Monitoring the patient's EEG during echo planar MRI," *Electroencephalogr. Clin. Neurophysiol.*, vol. 87, no. 6, pp. 417–420, 1993.

[86] F. R. Huang-Hellinger and H. C. Breiter *et al.*, "Simultaneous functional magnetic resonance imaging and electrophysiological recording," *Human Brain Map.*, vol. 3, no. 1, pp. 13–25, 1995.

[87] N. K. Logothetis and J. Pauls *et al.*, "Neurophysiological investigation of the basis of the fMRI signal," *Nature*, vol. 412, no. 6843, pp. 150–157, 2001.

[88] R. M. Muri and J. Felblinger *et al.*, "Recording of electrical brain activity in a magnetic resonance environment: Distorting effects of the static magnetic field," *Magn. Res. Med.*, vol. 39, no. 1, pp. 18–22, 1998.

[89] R. A. Hill and K. H. Chiappa *et al.*, "EEG during MR imaging: Differentiation of movement artifact from paroxysmal cortical activity," *Neurology*, vol. 45, no. 10, pp. 1942–1943, 1995.

[90] R. I. Goldman and M. S. Cohen *et al.*, "Combining EEG and functional MRI: Cleaning up the electrical signals," in *Proc. Int. Soc. Magn. Res. Med.*, Denver, CO, 2000.

[91] R. I. Goldman and J. M. Stern *et al.*, "Acquiring simultaneous EEG and functional MRI," *Clin. Neurophysiol.*, vol. 111, no. 11, pp. 1974–1980, 2000.

[92] M. S. Cohen and R. I. Goldman *et al.*, "Simultaneous EEG and fMRI made easy," in *Proc. Org. Human Brain Map.*, Brighton, U.K., 2001.

[93] R. I. Goldman and M. S. Cohen, "Simultaneous EEG and fMRI of alpha rhythm," *Neuroreport*, vol. 13, no. 18, pp. 2487–2492, 2002.

[94] P. Sajda, R. I. Goldman, M. G. Philiastides, A. D. Gerson, and T. R. Brown, "A system for single-trial analysis of simultaneously acquired EEG and fMRI," in *Proc. 3rd Int. IEEE/EMBS Conf. Neural Eng.*, May 2–5, 2007, pp. 287–290.

[95] M. Dyrholm, R. Goldman, P. Sajda, and T. R. Brown, "Removal of BCG artifacts using a non-Kirchhoffian overcomplete representation," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 2, pp. 200–204, 2009.

[96] A. Puce, T. Allison, S. Bentin, J. C. Gore, and G. McCarthy, "Temporal cortex activation in humans viewing eye and mouth movements," *J. Neurosci.*, vol. 18, no. 6, pp. 2188–2199, 1998.

[97] J. V. Haxby, E. A. Hoffman, and M. I. Gobbini, "The distributed human neural system for face perception," *Trends Cogn. Sci.*, vol. 4, no. 6, pp. 223–233, 2000.

[98] E. Hoffman and J. V. Haxby, "Distinct representation of eye gaze and identity in the distributed human neural system for face perception," *Nat. Neurosci.*, vol. 3, pp. 80–84, 2000.

[99] T. Allison, A. Puce, D. D. Spencer, and G. McCarthy, "Electrophysiological studies of human face perception: Potentials generated in occipitotemporal cortex by face and nonface stimuli," *Cereb. Cortex*, vol. 9, pp. 415–430, 1999.

[100] A. Puce, T. Allison, and G. McCarthy, "Electrophysiological studies of human face perception—III. Effects of top-down processing on face-specific potentials," *Cereb. Cortex*, vol. 9, pp. 445–458, 1999.

[101] A. Puce, T. Allison, and G. McCarthy, "Electrophysiological studies of human face perception—II. Response properties of face specific potentials generated in occipitotemporal cortex," *Cereb. Cortex*, vol. 9, pp. 431–444, 1999.

[102] M. G. Philiastides and P. Sajda, "EEG-informed fMRI reveals spatiotemporal characteristics of perceptual decision making," *J. Neurosci.*, vol. 27, no. 48, pp. 13082–13091, Nov. 2007.

[103] T. W. James, G. K. Humphrey, J. S. Gati, R. S. Menon, and M. A. Goodale, "The effects of visual object priming on brain activation before and after recognition," *Curr. Biol.*, vol. 10, pp. 1017–1024, 2000.

[104] T. W. James, G. K. Humphrey, J. S. Gati, R. S. Menon, and M. A. Goodale, "Differential effects of viewpoint on object-driven activation in dorsal and ventral streams," *Neuron*, vol. 35, pp. 793–801, 2002.

[105] K. Grill-Spector, T. Kushnir, T. Hendler, and R. Malach, "The dynamics of object-selective activation correlate with recognition performance in humans," *Nat. Neurosci.*, vol. 3, pp. 837–843, 2000.

[106] K. Grill-Spector, Z. Kourtzi, and N. Kanwisher, "The lateral occipital complex and its role in object recognition," *Vis. Res.*, vol. 41, pp. 1409–1422, 2001.

[107] K. Grill-Spector, N. Knouf, and N. Kanwisher, "The fusiform face area subserves face perception, not generic within-category identification," *Nat. Neurosci.*, vol. 7, no. 5, pp. 555–562, 2004.

[108] S. Ferber, G. K. Humphrey, and T. Vilis, "The lateral occipital complex subserves the perceptual persistence of motion-defined groupings," *Cereb. Cortex*, vol. 35, pp. 793–801, 2002.

[109] M. E. Large, A. Aldcroft, and T. Vilis, "Perceptual continuity and the emergence of perceptual persistence in the ventral visual pathway," *J. Neurophysiol.*, vol. 93, pp. 3453–3462, 2005.

[110] V. Di Lollo, "Temporal characteristics of iconic memory," *Nature*, vol. 267, pp. 241–243, 1977.

[111] M. Coltheart, "The persistences of vision," *Philos. Trans. Royal Soc. B*, vol. 290, pp. 57–69, 1980.

[112] R. VanRullen and C. Koch, "Visual selective behavior can be triggered by a feed-forward process," *J. Cogn. Neurosci.*, vol. 15, no. 2, pp. 209–217, 2003.

Paul Sajda (S'91–M'93–SM'06) received the B.S. degree in electrical engineering from the Massachusetts Institute of Technology (MIT), Cambridge, in 1989 and the M.S. and Ph.D. degrees in bioengineering from the University of Pennsylvania (UPenn), Philadelphia, in 1992 and 1994, respectively.

In 1994, he joined the David Sarnoff Research Center, where he went on to become Head of the Adaptive Image and Signal Processing Group. He is currently an Associate Professor of biomedical engineering and radiology at Columbia University, New York, where he is Director of the Laboratory for Intelligent Imaging and Neural Computing. His research focuses on neural engineering, neuroimaging, computational neural modeling, and machine learning applied to image understanding.

Dr. Sajda is a Fellow of the American Institute of Medical and Biological Engineering. He has received several awards for his research, including a National Science Foundation CAREER Award (2002), the Sarnoff Technical Achievement Award (1996), The Pollack Award for outstanding dissertation research in bioengineering (UPenn, 1994), The Flexner Award for outstanding thesis research in the neurosciences (UPenn, 1993), and the Adler Award for outstanding undergraduate research in electrical engineering (MIT, 1989). He is an Associate Editor of the IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING and a member of the IEEE Technical Committee on Neuroengineering.

Marios G. Philiastides received the B.S. degree in biomedical engineering from Rutgers—The State University, New Brunswick, NJ, in 2001, the M.S. degree in electrical engineering from Stanford University, Stanford, CA, in 2003, and the Ph.D. degree in biomedical engineering from Columbia University, New York, in 2007.

He is currently a Research Scientist with the Max Planck Institute for Human Development, Berlin, Germany. His research focuses on neuroimaging and computational modeling of visual and cognitive function with emphasis on perceptual and reward-based decision-making.

Lucas C. Parra (M'99–SM'07) received the Ph.D. degree in physics from the Ludwig-Maximilian University, Germany, in 1996.

He is a Professor of biomedical engineering at The City College, City University of New York. Previously, he was Head of the Adaptive Image and Signal Processing Group, Sarnoff Corporation (1997–2003), and a Member of the Machine learning and the Imaging departments at Siemens Corporate Research (1995–1997). His areas of expertise include machine learning, acoustic array processing, emission tomography, and electroencephalography. His current research in biomedical signal processing focuses on functional brain imaging and computational models of the central nervous system.