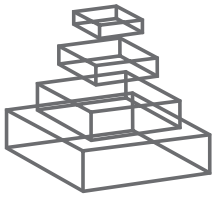
frontiers
RESEARCH TOPICS

SINGLE-TRIAL ANALYSES OF BEHAVIOURAL AND NEUROIMAGING DATA IN PERCEPTION AND DECISION-MAKING

Hosted by
Paul Sajda, Guillaume A. Rousselet and Cyril R. Pernet

frontiers in PSYCHOLOGY



frontiers

Cover image provided by Ibbl sarl, Lausanne CH

ISSN 1664-8714
ISBN 978-2-88919-023-2
DOI 10.3389/978-2-88919-023-2

ABOUT FRONTIERS

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

FRONTIERS JOURNAL SERIES

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing.
All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

DEDICATION TO QUALITY

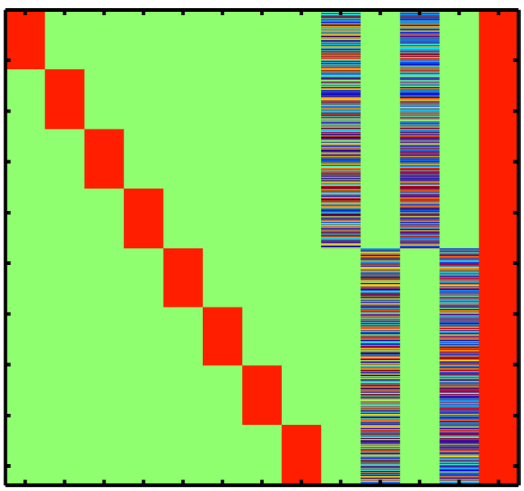
Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.
By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

WHAT ARE FRONTIERS RESEARCH TOPICS?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area!
Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

SINGLE-TRIAL ANALYSES OF BEHAVIOURAL AND NEUROIMAGING DATA IN PERCEPTION AND DECISION-MAKING

Hosted By
Paul Sajda, Columbia University, USA
Guillaume A. Rousselet, University of Glasgow, UK
Cyril R. Pernet, University of Edinburgh, UK



The cognitive psychology of perception and decision-making is at a cross-road. Most studies still employ categorical designs, a priori classified stimuli and perform statistical evaluations across subjects. However, a shift has been observed in recent years towards parametric designs in which the information content of stimuli is systematically manipulated to study the single-trial dynamics of behaviour (reaction times, eye movements) and brain activity (EEG, MEG, fMRI). By using the information contained in the variance of individual trials, the single-trial approach goes beyond the activity of the average brain: it reveals the specificity of information processing in individual subjects, across tasks and stimulus space, revealing both inter-individual commonalties and differences. This Research Topic provides theoretical and empirical support for the study of single-trial data.

Topics of particular interest include:

1. description of the richness of information in single-trials and how it can be successfully extracted;
2. statistical issues related to measures of central tendency, control for multiple comparisons, multivariate approaches, hierarchical modelling and characterization of individual differences;

3. how manipulation of the stimulus space can allow for a direct mapping of stimulus properties onto brain activity to infer dynamics of information processing and information content of brain states;
4. how results from different brain imaging techniques can be integrated at the single-trial level.

Image Caption: Example of a design matrix for an ANCOVA used to study how single-trial ERP amplitude is modulated by image characteristics and task constraints.

Image Credits: Rousselet, G. A., Gaspar, C. M., Wieczorek, K. P., and Pernet, C. R. (2011). Modeling single-trial ERP reveals modulation of bottom-up face visual processing by top-down task constraints (in some subjects). Front. Psychol. 2:137.

Table of Contents

frontiers in
PSYCHOLOGY

EDITORIAL
published: 08 November 2011
doi: 10.3389/fpsyg.2011.00322

Single-trial analyses: why bother?

Cyril R. Pernet[1]*, Paul Sajda[2] and Guillaume A. Rousselet[3]

[1] *Brain Research Imaging Centre, SINAPSE Collaboration, University of Edinburgh, Edinburgh, UK*
[2] *Department of Biomedical Engineering, Columbia University, New York, NY, USA*
[3] *Centre for Cognitive Neuroimaging, Institute of Neuroscience and Psychology, University of Glasgow, Glasgow, UK*
**Correspondence: cyril.pernet@ed.ac.uk*

Neuroimaging techniques have been traditionally used to demonstrate differences between means calculated across conditions or groups of subjects. However, as illustrated by the articles in this research topic, by studying the variability across trials, single-trial analyses can in some situations allow us to go beyond this kind of imaging to the mean. Indeed, single-trial analyses can provide additional information that is unobservable if we collapse the data to a mean. For example, single-trial analyses can help us provide a systematic mapping between (i) brain activity and stimulus information space (Schyns, 2010; Rousselet et al., 2011), (ii) brain activity and subject's behavioral variability (Ratcliff et al., 2009), and (iii) brain activity measured using different imaging techniques, e.g., fMRI and EEG (Goldman et al., 2009; deBettencourt et al., 2011). Importantly, using certain parametric experimental designs, single-trial analyses can give us access to brain mechanisms, by allowing us to specify the information content of brain activity and its transformation (Schyns, 2010; Rousselet and Pernet, 2011).

Single-trial analyses refer to methods that consider the variance within subjects. Two broad families of methods can be distinguished: univariate methods extract information among trials in space, time, or both; multivariate methods extract information across space, time, or both, in individual trials. Single-trial analyses may thus be used for behavioral experiments (e.g., Etchells et al., 2011) and neuroimaging experiments (e.g., Cohen and Cavanagh, 2011; Macdonald et al., 2011; Milne, 2011; Rousselet et al., 2011; Touryan et al., 2011; Wutte et al., 2011). Single-trial analyses of neuroimaging data have seen their use increase since the late 1960s, starting with Donchin (1969). Despite this long tradition and several advantages over group analyses, single-trial analyses remain nevertheless marginal.

The simplest form of univariate single-trial analysis is a regression over all of the trials in single subjects, to measure the relationship between, e.g., the signal amplitude and a parameterized stimulus space. This approach is often referred to as parametric design in fMRI. In this Research Topic, Rousselet et al. (2011) showed that a similar approach can be used in EEG to quantify brain responses to stimulus information in individual subjects, and characterize the probability of observing a mapping between stimulus information and EEG amplitude, thus going beyond the study of the average brain. Cohen and Cavanagh (2011) also demonstrated that the single-trial parametric approach can be extended to time–frequency decompositions of power and phase. Variance among trials also contains information about subjects or cognitive states. For instance, Milne (2011) established that children with autism have significantly more variance in the latency of their P1 response to Gabor patches than control participants. Macdonald et al. (2011) found that pre-stimulus alpha power is correlated with subjects' judgment of attentional state (see also VanRullen et al., 2011 for a review of evidence linking alpha oscillations to perception and attention).

Multivariate methods are often used to characterize the spatial–temporal variance in each trial in order to derive pattern classifiers (see however Friston et al., 1996 for a more traditional use). For instance, Touryan et al. (2011) used the variance in space and time to train a discriminant function to classify, in real time, brain activity related to familiar and unfamiliar faces. In their experiment, although the group ERPs differed between familiar and unfamiliar faces over frontal and parietal electrodes, the classification revealed that only the parietal response allowed the discrimination of the stimulus category on a single-trial basis. This result illustrates that group averaging may be misleading, presenting an abstract signal that cannot be found in individual subjects (see also Gaspar et al., 2011). Wutte et al. (2011) also used a pattern classification technique (support vector machine) to read-out motion direction from areas V1 and MT+ using fMRI. Although the spatial variance reflected the direction of perceived stimuli, individual perceptual thresholds were predicted by the relative variance in amplitude between activation and rest trials, thus illustrating the complementarity of univariate and multivariate methods.

In addition to a unique window on brain mechanisms, single-trial analyses also allow researchers to interpret individual differences by quantifying effects within and between subjects, providing a richer data description mandatory to build efficient models of perception and decision-making. It is often said that single-trial analyses require either too many trials, or dense coverage (dense arrays in MEEG or fast TR in fMRI), or both. It is true that in order to obtain good signal-to-noise ratio (regression over trials) many trials are necessary and, in order to obtain good patterns ("weighting" across electrodes/voxels, time intervals, frequency intervals), dense coverage is mandatory. Many trials are nevertheless also mandatory for an average to be a meaningful measure (Rousselet et al., 2008), just as dense coverage is necessary to ensure that minima or maxima located between sampled time points in fMRI or between channels in MEEG are not overlooked. There are a growing number of user friendly toolboxes available to perform single-trial analyses (e.g., Parra et al., 2005; Hanke et al., 2009; Delorme et al., 2011; Hartmann et al., 2011; Oostenveld et al., 2011; Pernet et al., 2011). We encourage everyone interested in understanding how the stimulus space and behavioral response map onto brain activity to use these tools rather than merely amass binary results showing group differences in brain activity among conditions (Rousselet and Pernet, 2011).

REFERENCES

Cohen, M. X., and Cavanagh, J. F. (2011). Single-trial regression elucidates the role of prefrontal theta oscillations in response conflict. *Front. Psychol.* 2:30. doi: 10.3389/fpsyg.2011.00030

deBettencourt, M., Goldman, R., Brown, T., and Sajda, P. (2011). Adaptive thresholding for improving sensitivity in single-trial simultaneous EEG/fMRI. *Front. Psychol.* 2:91. doi: 10.3389/fpsyg.2011.00091

Delorme, A., Mullen, T., Kothe, C., AkalinAcar, Z., Bigdely-Shamlo, N., Vankov, A., and Makeig, S. (2011). EEGLAB, SIFT, NFT, BCILAB, and ERICA: new tools for advanced EEG processing. *Comput. Intell. Neurosci.* 2011, 130714.

Donchin, E. (1969). Discriminant analysis in average evoked response studies: the study of single-trial data. *Electroencephalogr. Clin. Neurophysiol.* 27, 311–314.

Etchells, P. J., Benton, C. P., Ludwig, C. J. H., and Gilchrist, I. D. (2011). Testing a simplified method for measuring velocity integration in saccades using a manipulation of target contrast. *Front. Psychol.* 2:115. doi: 10.3389/fpsyg.2011.00115

Friston, K. J., Stephan, K. M., Heather, J. D., Frith, C. D., Ioannides, A. A., Liu, L. C., Rugg, M. D., Vieth, J., Keber, H., Hunter, K., and Frackowiak, R. S. J. (1996). A multivariate analysis of evoked responses in EEG and MEG data. *Neuroimage* 3, 167–174.

Gaspar, C. M., Rousselet, G. A., and Pernet, C. R. (2011). Reliability of ERP and single-trial analyses. *Neuroimage* 58, 620–629.

Goldman, R. I., Wei, C. Y., Philiastides, M. G., Gerson, A. D., Friedman, D., Brown, T. R., and Sajda, P. (2009). Single-trial discrimination for integrating simultaneous EEG and fMRI: identifying cortical areas contributing to trial-to-trial variability in the auditory oddball task. *Neuroimage* 47, 136–147.

Hanke, M., Halchenko, Y. O., Sederberg, P. B., Hanson, S. J., Haxby, J. V., and Pollmann, S. (2009). PyMVPA: a python toolbox for multivariate pattern analysis of fMRI data. *Neuroinformatics* 7, 37–53.

Hartmann, T., Schulz, H., and Weisz, N. (2011). Probing of brain states in real-time: introducing the console environment. *Front. Psychol.* 2:36. doi: 10.3389/fpsyg.2011.00036

Macdonald, J. S. P., Mathan, S., and Yeung, N. (2011). Trial-by-trial variations in subjective attentional state are reflected in ongoing prestimulus EEG alpha oscillations. *Front. Psychol.* 2:82. doi: 10.3389/fpsyg.2011.00082

Milne, E. (2011). Increased intra-participant variability in children with autistic spectrum disorders: evidence from single-trial analysis of evoked EEG. *Front. Psychol.* 2:51. doi: 10.3389/fpsyg.2011.00051

Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J. M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell Neurosci.* 2011, 156869.

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

Pernet, C. R., Chauveau, N., Gaspar, C., and Rousselet, G. A. (2011). LIMO EEG: a toolbox for hierarchical linearmodeling of electroencephalographiuc data. *Comput. Intell. Neurosci.* 2011, 831409.

Ratcliff, R., Philiastides, M. G., and Sajda, P. (2009). Quality of evidence for perceptual decision making is indexed by trial-to-trial variability of the EEG. *Proc. Natl. Acad. Sci. U.S.A.* 106, 6539–6544.

Rousselet, G. A., Gaspar, C. M., Wieczorek, K. P., and Pernet, C. R. (2011). Modeling single-trial ERP reveals modulation of bottom-up face visual processing by top-down task constraints (in some subjects). *Front. Psychol.* 2:137. doi: 10.3389/fpsyg.2011.00137

Rousselet, G. A., Husk, J. S., Bennett, P. J., and Sekuler, A. B. (2008). Time course and robustness of ERP object and face differences. *J. Vis.* 8, 3.1–18.

Rousselet, G. A., and Pernet, C. R. (2011). Quantifying the time course of visual object processing using ERPs: it's time to up the game. *Front. Psychol.* 2:107. doi: 10.3389/fpsyg.2011.00107

Schyns, P. G. (2010). Grand challenges in perception science: modeling the future. *Front. Psychol.* 1:10. doi: 10.3389/fpsyg.2010.00010

Touryan, J., Gibson, L., Horne, J. H., and Weber, P. (2011). Real-time measurement of face recognition in rapid serial visual presentation. *Front. Psychol.* 2:42. doi:10.3389/fpsyg.2011.00042

VanRullen, R., Busch, N. A., Drewes, J., and Dubois, J. (2011). Ongoing EEG phase as a trial-by-trial predictor of perceptual and attentional variability. *Front. Psychol.* 2:60. doi: 10.3389/fpsyg.2011.00060

Wutte, M. G., Smith, M. T., Flanagin, V., and Wolbers, T. (2011). Physiological signal variability in hMT+ reflects performance on a direction discrimination task. *Front. Psychol.* 2:185. doi: 10.3389/fpsyg.2011.00185

Received: 17 August 2011; accepted: 20 October 2011; published online: 08 November 2011.

Citation: Pernet CR, Sajda P and Rousselet GA (2011) Single-trial analyses: why bother? Front. Psychology ***2****:322. doi: 10.3389/fpsyg.2011.00322*

This article was submitted to Frontiers in Perception Science, a specialty of Frontiers in Psychology.

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 26 May 2011
doi: 10.3389/fpsyg.2011.00115

Testing a simplified method for measuring velocity integration in saccades using a manipulation of target contrast

Peter J. Etchells*, Christopher P. Benton, Casimir J. H. Ludwig and Iain D. Gilchrist

School of Experimental Psychology, University of Bristol, Bristol, UK

Edited by:
Paul Sajda, Columbia University, USA

Reviewed by:
Roberto Caldara, University of Glasgow, UK
Piers D. L. Howe, Harvard Medical School, USA

****Correspondence:***
Peter J. Etchells, School of Experimental Psychology, University of Bristol, 12A Priory Road, Clifton, Bristol BS8 1TU, UK.
e-mail: peter.etchells@bristol.ac.uk

A growing number of studies in vision research employ analyses of how perturbations in visual stimuli influence behavior on single trials. Recently, we have developed a method along such lines to assess the time course over which object velocity information is extracted on a trial-by-trial basis in order to produce an accurate intercepting saccade to a moving target. Here, we present a simplified version of this methodology, and use it to investigate how changes in stimulus contrast affect the temporal velocity integration window used when generating saccades to moving targets. Observers generated saccades to one of two moving targets which were presented at high (80%) or low (7.5%) contrast. In 50% of trials, target velocity stepped up or down after a variable interval after the saccadic go signal. The extent to which the saccade endpoint can be accounted for as a weighted combination of the pre- or post-step velocities allows for identification of the temporal velocity integration window. Our results show that the temporal integration window takes longer to peak in the low when compared to high contrast condition. By enabling the assessment of how information such as changes in velocity can be used in the programming of a saccadic eye movement on single trials, this study describes and tests a novel methodology with which to look at the internal processing mechanisms that transform sensory visual inputs into oculomotor outputs.

Keywords: saccades, contrast, velocity integration, motion, prediction

INTRODUCTION

Saccadic eye movements serve to orient the fovea onto an object or region of interest within the visual environment. These movements are the result of a decision process that is typically based on the analysis of sensory information, and so offer an ideal route through which to assess how decision-making mechanisms may be implemented by sensorimotor circuits in the brain (Gold and Shadlen, 2001, 2007; Glimcher, 2001; Schall, 2003). In recent years, there has been growing interest in the development of methods with which to assess how perceptual signals inform eye movement decisions (Beutter et al., 2003; de Brouwer et al., 2002; Caspi et al., 2004; Ludwig et al., 2005, 2007; Bennett et al., 2007; Eckstein et al., 2007; Spering et al., 2007; Nummela et al., 2008; Tavassoli and Ringach, 2009; Etchells et al., 2010).

Although the questions under investigation in these various studies differed, as did the precise methods used, there is a common theme. In general, a visual stimulus is perturbed in some way or another (e.g., adding random luminance noise over time in Caspi et al., 2004 and Ludwig et al., 2005). Careful analysis of how this perturbation influences behavior on single trials then enables estimation of the spatial and/or temporal portions of the stimulus that preferentially drive decisions, through a variety of techniques (e.g., reverse correlation or logistic regression approaches). Important new insights have been obtained with these methodologies. For instance, Caspi et al. (2004) were able to show that the uptake of visual information in a single fixation drove not only the immediately following eye movement decision, but also the one after that. Ludwig et al. (2005) showed that decisions were driven by visual information time-locked to display onset, rather than saccade onset. Indeed, these authors showed that only a remarkably short portion of the overall latency period was used to integrate the sensory evidence (see also Ludwig, 2009).

Recently, we have developed a related method to assess over what time interval object velocity information is extracted in order to accurately intercept a moving object with a saccade (Etchells et al., 2010). Targeting a moving object poses a challenging decision problem: sensory input and motor output delays, as well as the eye movement duration itself, will result in a delay between the decision being made to generate an eye movement and the actual completion of that movement (Kerzel and Gegenfurtner, 2003). Consequently, some decision has to be made regarding how far ahead of the "currently seen" object position a saccade is to land, given the continued object motion during movement programming and execution. Clearly, having an estimate of the object velocity is desirable for this purpose.

Our method to identify the epoch over which this information is extracted, follows the same logic as presented above (and is closely related to the double-step method used to infer over what epoch position information is extracted; Becker and Jürgens, 1979). Observers are presented with two targets moving at a particular velocity. A "go" signal indicates which object observers have to saccade to. At some point after the go signal, target velocity is perturbed: the objects abruptly speed up or slow down. The random variation from trial-to-trial in the timing of the speed step, coupled with the natural variability in saccade latency, can be used to build up a picture of how much time the saccadic system needs to be able to incorporate information about the second speed into the saccade program.

The landing position on each trial may be used to estimate the relative weights attributed to the pre- and post-step velocities, by comparing the observed endpoint with the predicted endpoints based on the two velocities. We then assess how these weights change as a function of time from saccade onset. For instance, if the velocity step occurs long before movement onset the observer will have had more time to base their decision on the post-step, veridical velocity. As will be explained in detail below, fitting these weights over time with a model provides an estimate of the time interval over which object velocity was extracted.

Our previous work suggests that the system used a temporal window with a duration of ~100 ms to estimate target velocity (Etchells, et al., 2010). The end of the window is positioned ~80 ms before the onset of the saccade. The latter period may be considered the saccadic dead-time, which is functionally defined as the period during which new visual information can no longer affect the saccade endpoint (Becker and Jürgens, 1979; Findlay and Harris, 1984; Aslin and Shea, 1987; Ludwig et al., 2007). The observed endpoint from each trial is converted into a relative weight associated with the post-step velocity. These weights are then fitted with some functional form.

Our model is not a process model that specifies the visual mechanisms involved in velocity estimation. However, there is a process interpretation of the model, which is illustrated in **Figure 1**. We assume that during the latency period object velocity is estimated by convolving the input velocities with some temporal filter (Benton and Curran, 2009) such as that seen in **Figure 1**. This operation is analogous to computing a running, weighted average of the input. The temporal integration performed by the filter necessarily results in a certain amount of blurring of the velocity information when the velocity is variable. As a result, the observed endpoints may not simply reflect either the pre-step velocity or the post-step velocity, but may be driven by intermediate velocity estimates. The prediction period shown in the figure is assumed to consist of the dead-time and the saccade duration itself.

In the present study, our aims were twofold. First, we sought to validate this process interpretation of the model using a straightforward manipulation of the input which is known to profoundly affect the visual system: a variation in contrast. Second, we aimed to simplify the method of fitting the model to make it more user-friendly. The model presented by Etchells et al. (2010) included specification and estimation of a multitude of noise sources that, together, produced variability in the velocity weights (e.g., variability in saccade duration, which is correlated with variability in saccade amplitude). In this article we describe and test a significant simplification, which essentially combines all noise sources together and eases the estimation of the critical parameters of interest: those that describe the velocity weighting function.

In the model presented in Etchells et al. (2010), observers were presented with targets that did not differ in contrast from trial-to-trial. In the current study, in order to test and demonstrate our simplified model, we examine the effects of changing stimulus contrast on velocity integration. The work in the current paper therefore presents (1) a methodological advance in the form of a simple technique for characterizing the incorporation of velocity information into saccadic planning, and (2) an empirical advance in the form of a quantification of the effects of changing contrast on velocity integration during saccade planning. A wealth of research over the past 50 years has given us detailed knowledge of how contrast affects the visual system (e.g., Mansfield, 1973; Breitmeyer, 1975; Harwerth and Levy, 1978; Plainis and Murray, 2000; Weiss et al., 2002; Murray and Plainis, 2003; Carpenter, 2004; Ludwig et al., 2004; Taylor et al., 2006; White et al., 2006) and its underlying neurophysiology (e.g., Enroth-Cugell and Robson, 1966; Pack et al., 2005; Krekelberg et al., 2006; Livingstone and Conway, 2006). Consequently, we can make some informed predictions about the effect that contrast will have on the generation of saccades to moving targets.

Weiss et al. (2002) suggest that at low-contrast, there is less precise information about the actual speed of a given stimulus. The greater level of uncertainty is represented by an increase in the spread of the likelihood function of target velocity estimates. In other words, reducing stimulus contrast corresponds to a decrease in the signal-to-noise ratio (SNR) of the velocity measurement. If the velocity weighting function we measure with our method maps onto the underlying temporal filter used to estimate velocity, we might reasonably expect the width of the filter to increase when the target contrast is low. By extending the amount of time during which the velocity signal is sampled and averaged, SNR is increased to obtain a more precise estimate of target velocity.

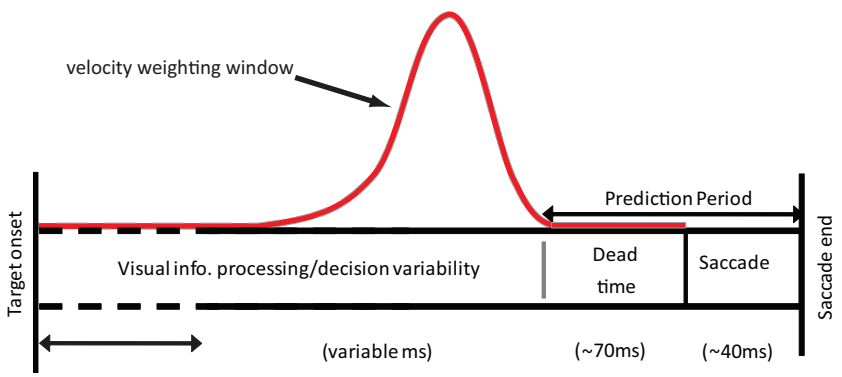


FIGURE 1 | The time course of a saccade from target detection to the end of the saccade. During this time, average velocity is estimated using a weighting window similar in nature to the one shown here (red solid line). The target is assumed to move at this averaged velocity during the prediction period. This information can then be used to determine where the target is likely to be located at the end of the saccade.

Alternatively, a reduction in contrast may result in an increase in the time it takes for the incoming velocity information to reach the integration mechanism, as a result of increasing neuronal conduction latencies (e.g., Kaplan and Shapley, 1982). For example, Maunsell et al. (1999) showed that, depending on the number of inputs summed, latency differences in the magnocellular pathway through the lateral geniculate nucleus (LGN) are likely to be on the order of 5–15 ms between high and low intensity stimuli, with high intensity stimuli producing faster responses. A change in the velocity of the lower contrast stimulus will therefore take a longer time to register, which would result in a delay in the velocity signal reaching the integration mechanism. In our methodology, the time-to-peak of the velocity weighting function reflects the time at which emphasis is shifted onto more recent velocity inputs. Therefore, when contrast is reduced, we might expect to see a delay in the point at which a velocity change is detected and incorporated into the final velocity estimate.

EXPERIMENTAL OVERVIEW

Observers were required to fixate a central diamond-shaped fixation stimulus on a computer screen whilst two Gaussian patches (SD = 0.32°) traversed horizontally across the screen, 6° above and below the midline. During the trial, the fixation point would change into either an upwards- or downwards-pointing arrow, indicating which patch the observers had to make a saccade to (see **Figure 2A**). On some trials, after a variable delay the speed of the patches would change. By looking at the relationship between saccade landing positions and the time of the speed changes, we can determine how the saccadic system weighs the two velocities over time. We examine the nature of this velocity integration function in two conditions: high and low-contrast.

OBSERVERS

Six observers were recruited from the students of the University of Bristol, UK (3 females, age range 24–30, mean age 26.0). All had self-reported normal or corrected-to-normal vision. Data were collected over the course of four sessions, performed on different days. The study was approved by the local ethics committee.

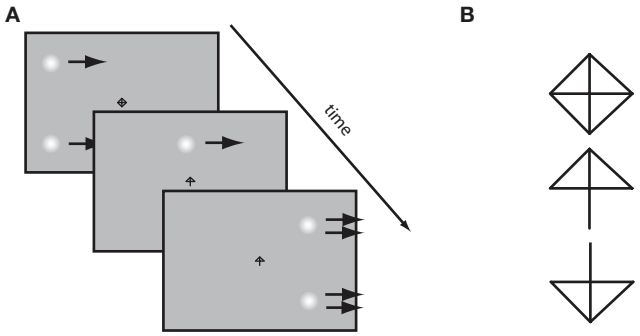


FIGURE 2 | (A) Outline of the time course for a rightwards, step up trial. Gaussian patches start moving rightward. After some interval of time the fixation stimulus changes to an arrow, signaling that a saccade to the top patch should be made. After this, the patches step up in speed (illustrated by the double arrows). Note that patches and fixation point are not to scale, for illustration purposes. **(B)** The fixation stimulus. Removal of either the bottom two or top two diagonal line segments results in an arrow indicating which patch to saccade to.

EYE MOVEMENT RECORDING

Stimuli were displayed on a 21-inch, gamma-corrected CRT monitor (LaCie Electron Blue) running at 75 Hz. The monitor resolution was 1152 × 864 pixels, and the screen subtended 36° × 24° of visual angle. An Eyelink 1000 system (SR Research, Mississauga, ON, Canada) was used to record and monitor eye movements. This is an infrared tracking system that uses the pupil center in conjunction with corneal reflection to sample eye position at 1000 Hz. For each data sample, a dedicated parser algorithm (SR Research, Mississauga, ON, Canada) computes the instantaneous velocity and acceleration of the eye. These are then compared to threshold criteria for velocity (30°/s) and acceleration (8000°/s^2). If either is above threshold, the eye movement is classified as a saccade. Visual inspection of a random selection of saccades confirmed that the automatic algorithm placed the on and offsets of the saccades appropriately, without including any apparent contributions from the potential pursuit component that may have followed the saccade. Head position was stabilized at a viewing distance of 57 cm via the use of the Eyelink 1000 built-in chin rest. Observers viewed the display monocularly using their dominant eye, and eye dominance was measured using the hole-in-the-card technique (Seijas et al., 2007). The experimental software was programmed in MATLAB using the Psychophysics Toolbox (Brainard, 1997) and Eyelink Toolbox extensions (Cornelissen et al., 2002).

DESIGN AND PROCEDURE

Observers performed 28 experimental blocks, each containing 128 trials. Prior to each block, a nine-point calibration procedure was performed in which observers were asked to fixate a black cross that appeared randomly on a 3 × 3 grid. The fixation stimulus measured 0.3° × 0.3°, and the calibration grid subtended 31° × 19° of visual angle. On a given trial, observers were instructed to fixate a central stimulus which took the form of a black diamond containing a cross (see Walker et al., 2000, Experiment 2). Observers were instructed which patch to make a saccade to by the removal of either the top two or bottom two diagonal segments, respectively forming either a downwards or upwards arrow (see **Figure 2B**).

In 50% of the trials within each block, the Gaussian patches would start moving at a constant speed of 18°/s and remain at this speed for the duration of the trial. In 25% of the trials, the patches would step up from 18 to 30°/s at a variable time after the change in fixation stimulus. In the remaining 25% of the trials, the patches would step down to 6°/s. The patches were shown for at least 445 ms before the change in fixation stimulus would occur. After this time, an exponential, i.e., "non-aging," foreperiod (Nickerson and Burnham, 1969; Oswal et al., 2007) was used to determine the time of the fixation change. A non-aging foreperiod can be described as one in which the probability of a target appearing in the next time interval decreases exponentially over time. This results in an observer's expectation remaining constant over the course of a trial, which avoids portions of the observer's response being attributable to something other than the visual information in the stimulus. The mean of this exponential distribution was

128 ms. A second non-aging foreperiod, with a mean of 100 ms, was used to determine the time of the speed step. In both cases, a maximum cumulative probability of 95% was used to truncate the distribution, in order to prevent the generation of extremely long foreperiods that would take the patterns off the screen.

Within each block the contrast of the patches remained the same, with contrasts being randomized between blocks. Half of the blocks were presented at a high contrast and half at low-contrast. Contrast was defined as $L_{max} - L_o/L_o$ where L_o indicates background luminance. This gives contrast values of 80% in the high condition (L_{max} = 65.4 cd/m^2, L_o = 36.4 cd/m^2) and 7.5% in the low condition (L_{max} = 39.1 cd/m^2, L_o = 36.4 cd/m^2).

DATA ANALYSIS

Only data describing the first saccade in each trial were considered. As the minimum distance between the fixation point and target was 6° (when the target was located directly above or below fixation), trials in which the amplitude of the first saccade was less than 4° were rejected, as were trials in which no saccade was generated. Saccade endpoints were recorded, along with the target location at saccade termination and saccade amplitude. The primary analysis concerned the horizontal component of each saccade.

RESULTS

For each contrast condition, we first want to determine the extent to which the first orienting saccade is influenced by the pre- and post-step velocities. We do this by measuring, for each trial, the horizontal error between where the saccade landed, and where the target was at the end of the saccade. This saccade endpoint error is then plotted as a function of the time between the velocity step and saccade onset, which we term *D*. This is aligned on saccade onset, such that *D* = 0 ms corresponds to the start of the saccade, and *D* = 500 ms corresponds to the velocity step occurring 500 ms before saccade onset.

Figure 3 shows, for a single observer, this landing position error as a function of *D* for both high and low-contrast conditions.

As values of *D* increase from zero (saccade onset), the general pattern of data in the high contrast condition indicates a gradual increase in the amount of error between the saccade endpoint and the target location at saccade end, up until a time of around 160–170 ms prior to saccade onset. After this time, landing position error decreases back to zero at *D* = 200–300 ms in the majority of trials. The initial increase in error reflects an over-reliance on the pre-step speed. As *D* increases, the observers will have seen the target traveling for a longer time at the post-step speed, and therefore will begin to rely more heavily on the veridical velocity. This pattern, whilst similar in the low-contrast condition, shows relatively fewer trials in which the landing position error returns to zero. In those cases where it does, there is an increase in the time it takes for the error to do so, coupled with much greater variability. We now turn to a description of how we estimate the relative weighting of the two velocities, and how these weights can be used to identify the period over which velocity is integrated.

For a given observer, we first calculate two errors for each saccade *n*. ε_n^{new} is the difference between saccade landing position and the actual position of the target at the end of the saccade. This error is simply the distance from the zero-error abscissa in **Figure 3**.

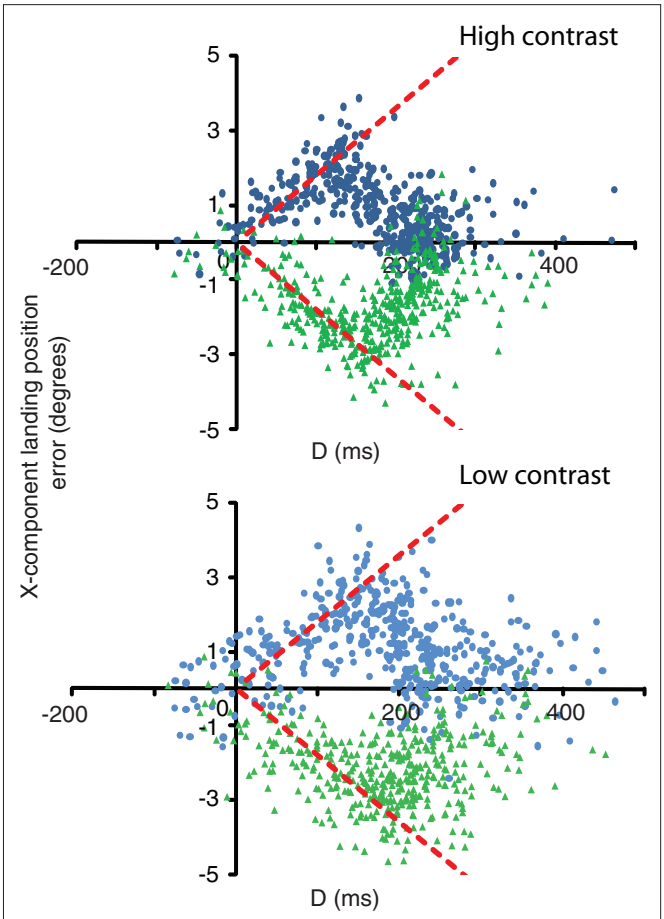


FIGURE 3 | *D* (time between velocity change and saccade onset) versus the *x*-component error in saccade landing positions for a single observer (observer 1). Data are separated out by contrast condition. Green triangles denote speed step up trials; blue circles denote speed step down trials. Predicted behavior based solely on the post-step speed corresponds to zero-error (i.e., the *x*-axis). Predicted behavior based solely on the pre-step speed is shown by the red dashed lines.

ε_n^{old} is the difference between the saccade landing position and the point where the target would have been had it not changed speed. The error that would result from completely following the pre-step velocity is shown by the dashed lines in **Figure 3**. Note that these predictions depend on the saccade duration and will therefore vary slightly from trial-to-trial. The dashed lines are drawn on the basis of the average saccade duration, just for the purpose of illustration.

With these two error terms in place, we determine a relative weighting that the saccadic system places on the post-step velocity, ρ_n, given by:

$$\rho_n = \frac{\varepsilon_n^{old}}{\varepsilon_n^{old} - \varepsilon_n^{new}} \quad (1)$$

Values of ρ_n range between 0 and 1, with ρ_n = 0 equivalent to the saccadic system solely basing its response on the pre-step speed, and ρ_n = 1 equivalent to the system solely utilizing the post-step speed. **Figure 4** illustrates (in 15 roughly equal bins) how these weights vary as a function of *D*. As expected, just before saccade

onset, the system has not had time to include the new velocity in its movement program, corresponding to a post-step velocity weight of 0. As time to saccade onset increases, more emphasis is placed on the post-step velocity, eventually reaching values close to 1. It is clear that the transition is gradual, which may be attributed to the varying portions of the pre- and post-step velocities falling under the temporal filter.

More formally, if we think of the velocity change as a step function falling within some temporal filter $f(t)$, then ρ_n corresponds to the area under $f(t)$ that falls after the velocity change. Therefore, the plot of ρ_n at a range of values of D gives us the integral of the temporal filter. In order to obtain an estimate of the filter, we fit a reasonable function to these data, and then take its derivative. Note that our actual model fits are based on the data from individual trials, not the binned data which are shown in **Figure 4** for illustration purposes only.

We opted for maximum likelihood parameter estimation, because this allows us to perform likelihood-based hypothesis testing of the effects of our experimental manipulation on the various properties of the estimated filter (see below; Burnham and Anderson, 2002; Wagenmakers, 2007). Maximum likelihood requires specification of the probability distribution from which the data points are drawn. This is not straightforward, because it is difficult to know (a) what sources of noise contribute to the variability in the post-step velocity weights and (b) how these noise sources are distributed at any given value for D.

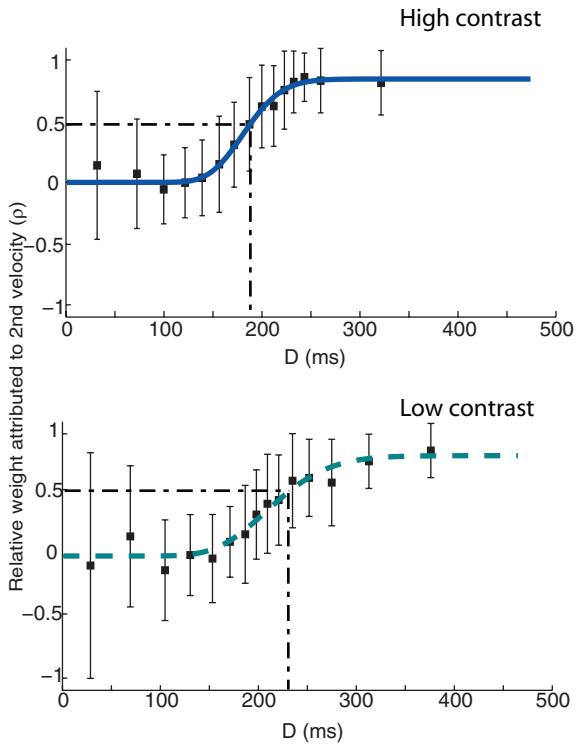


FIGURE 4 | ρ_n (the relative weight attributed to the post-step velocity) as a function of *D* (time between velocity change and saccade onset) for a single observer. Dark blue solid line in the top panel denotes the fit of a cumulative Gamma function to the high contrast condition data. Light blue dashed line in the lower panel denotes a Gamma function fit to the low-contrast data. Error bars denote the SD. The dash–dot line in both plots denotes the time at which $\sigma_n = 0.5$, i.e., when both velocities are being weighted equally. Note that while the data are presented here in approximately 15 equal bins, the actual Gamma fit was conducted on the full data set.

In our previous work (Etchells et al., 2010) we assumed that the two types of endpoint error, ε^{old} and ε^{new}, were both Gaussian distributed variables. With the weights defined as a ratio, their probability distribution was described as a ratio of Gaussian densities (Marsaglia, 2006). The parameters of the individual Gaussian components depended on a number of variables that were of minor theoretical interest, such as the mean saccade duration, variability in saccade duration, and variability in saccade landing position. An added complication was that these three quantities could, and to some extent did, vary as a function of D. To limit the number of free parameters we used a kernel estimator for the values of these quantities across the entire range of D and incorporated these estimates in the full expression of the probability distribution of ρ.

We appreciate that this procedure is rather cumbersome, for what appears to be a relatively straightforward and lawful pattern of data. We were therefore keen to develop a simplified method and assess to what extent the estimated velocity weights would be affected by the simplification. In the simplified model, we assume that ρ_n is Gaussian distributed, with a mean that varies as a function of D according to some functional form (see below). This is the variation that is of primary theoretical interest.

In standard maximum likelihood regression, the SD describes the residuals around the (predicted) mean. It is generally assumed to remain constant and left a free parameter. However, in our case the variability around the (mean) weights clearly varies as a function of D, which can be seen in the size of the error bars in **Figure 4**. It is important to include this variation: the fit should be most heavily constrained by those data points that were estimated with greater accuracy (i.e., for larger values of D). To make the dependence on D explicit, we write:

$$\rho_n \sim \mathcal{N}[\bar{\rho}(D), \sigma(D)] \tag{2}$$

We were reluctant to introduce additional free parameters to describe the relation between the SD and time from saccade onset. After all, it is not immediately obvious what function best describes this relation and, more importantly, this relation is not of primary theoretical interest. For this reason, we estimated the SD from the observed values of ρ_n using a Gaussian kernel, for values of D ranging from 1 to 500 ms. The bandwidth of this smoothing window was set for each observer separately, to whatever bandwidth best captured the distribution of D sampled for that observer. We reasoned that as the variable of interest is sampled as a function of D, a bandwidth for the optimal sampling of D would provide a reasonable window for smoothing the variability in the weights. Specifically, we computed a weighted SD, for every value of D in 1-ms increments:

$$\rho(D) = \sqrt{\frac{\sum_{i=1}^{N} \omega_i (\rho_i - \mu(D))^2}{\sum_{i=1}^{N} \omega_i}} \tag{3}$$

Here the vector of weights, ω, is the Gaussian smoothing function sampled at 1-ms intervals and μ is the Gaussian-weighted mean.

To capture $\overline{\rho}(D)$ we initially chose a scaled cumulative Gamma function. The Gamma function was chosen to accommodate both symmetric and asymmetrical filters, and has frequently been used to describe temporal filters (Watson, 1986; Smith, 1995). The smooth curves in **Figure 4** show the fits of the simplified model. These curves are characterized by three free parameters, namely *a* (the upper asymptote – the lower bound was set to zero), *k* (shape), and θ (scale). The Nelder–Mead Simplex method (Nelder and Mead, 1965) was used in order to find the set of best-fitting parameters.

It is clear that these functions describe the data well. For each of the 12 data sets reported in the paper (six observers and two contrast levels), we computed the correlation between the predicted weights estimated using the original and simplified fitting methods, for the observed values of *D*. In all cases the correlation was greater than 0.99. As such, the drastically simplified model results in very similar velocity weighting functions as the more complete (and complex) model developed in our previous work.

Having established the viability of the simplified model, we now turn to the empirical question of interest: what is the effect of contrast on the velocity weighting function? For the data shown in **Figure 4**, it appears that it takes the system much longer to incorporate the post-step speed into the saccade landing position calculation, as indicated by the more gradual rise to $\rho_n = 1$ in the low-contrast condition. For clarity, the dash–dot line in **Figure 4** illustrates the point at which the two velocities are being weighed equally (i.e., when $\rho_n = 0.5$), showing a shift from ~190 to ~230 ms between the high and low-contrast conditions.

Figure 5 shows, for each observer, the derivatives of the weight versus *D* functions for both contrast conditions. The solid black line shows the filter plot for the high contrast data, and the dashed line shows the filter plot for the low-contrast data. The corresponding shaded regions denote the 95% confidence intervals. These were calculated by producing 1000 bootstrap replications of the fit parameters, using the percentile method (Efron and Tibshirani, 1993).

The data show that the filter peaks shift toward larger values of *D* for the low-contrast condition ($M = 207$ ms, SEM = 1.5 ms) as compared to the high contrast condition ($M = 188$ ms, SEM = 2.5 ms), for every single observer. There also appears to be a slight increase in the width of the filter as contrast is reduced ($M = 79$ ms, SEM = 6.7 ms in the high contrast condition, $M = 88$ ms, SEM = 7.7 ms in the low-contrast condition). These effects are not concomitant with an increase in saccade latency in the step conditions – in the high contrast condition, the mean saccade latency across all observers was 271 ms (SEM = 8.3 ms), compared to a mean saccade latency of 274 ms (SEM = 6.5 ms) in the low-contrast condition – an increase of only 3 ms. The

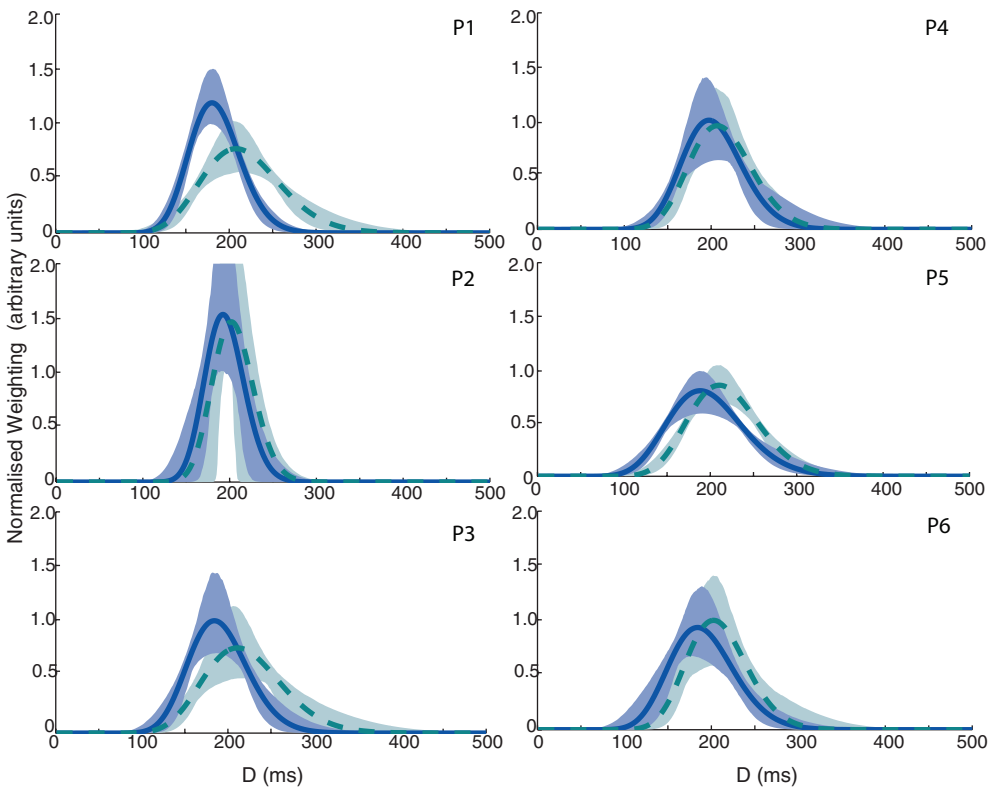


FIGURE 5 | Velocity integration filters for all six observers in Experiment 1. Solid dark blue line denotes the plot based on the high contrast data; dashed light blue line denotes the plot for the low-contrast data. Shaded areas indicate 95% confidence limits based on 10,000 bootstrap replications (dark blue and light green for high and low-contrast conditions, respectively).

saccade latency should correspond to any changes in the saccadic go signal (i.e., the fixation stimulus change). However, if as a result of the reduction in contrast, it is harder to localize the target (for example, for the purposes of the final position grab), then we might reasonably expect that the saccade latency would increase. The fact that we do not see this is important, as it shows that peak shifts that we see are not simply a result of "stretching" that occurs due to a general increase in the time it takes to detect or generate a saccade to a lower contrast target.

To assess the effects of contrast more formally, we adopted a model selection approach (Burnham and Anderson, 2002; Wagenmakers, 2007). Indeed, this was the motivation for estimating the model parameters using maximum likelihood. For this purpose, we defined a number of competing models that represent different hypotheses about the effect(s) of contrast. The likelihoods of these models constitute an index of the amount of evidence provided by the data for the different hypotheses.

The following four competing models were defined. If contrast had no effect on the width or peak location of the filter, we would expect a set of four parameters to suffice – a single peak location parameters, a width parameter, and two separate asymptotes (this is hereafter known as the baseline model). At the other extreme, contrast could affect every possible aspect of the filter, necessitating two separate sets of three parameters to account for the data. We refer to this model as the saturated model. In between these two extremes fall two reduced models of critical interest: (1) a five-parameter peak location model accommodates the data from both contrast conditions with a common width, but allows the location and asymptote to vary with contrast; (2) a five-parameter width model which assumes different widths and asymptotes for the low and high contrasts. In other words, we force the location and/or widths to be the same.

One problem with the Gamma function is that its parameters do not correspond to experimentally interesting parameters such as filter width and peak position. This makes it very difficult to, for example, test the competing models that we have outlined above. Therefore, for the purpose of this test, we chose to fit our data with scaled cumulative Gaussian curves, instead of the Gamma functions used earlier. As can be seen in **Figure 5**, the identified temporal filters are relatively symmetrical. Seeing as this is the case, a Gaussian function is attractive because its two parameters are independent and correspond directly to the location and width parameters of the filter. In comparison, the shape and scale parameters of the Gamma interact to jointly determine its location and width. The inability of the Gaussian to accommodate the slight asymmetries in the filters, did result in a decreased goodness-of-fit, as we will show below. However, the critical components of the filter, the width and location, were numerically very close under both models. Moreover, the effect(s) of contrast on these two components was also very similar when estimated with Gaussian or Gamma functions. For this reason, we see the reduction in goodness-of-fit as a price worth paying for the greater utility of the Gaussian function in terms of parameter interpretation.

Finally, the four models defined above differ in the number of free parameters. In evaluating their likelihoods, it is desirable to take this variation in complexity into account. A Bayesian information criterion (BIC) adjusts the log-likelihood of a model, given the observed data and best-fitting parameters, according to the number of free parameters (Schwarz, 1978; Wagenmakers, 2007). In particular:

$$\mathrm{BIC} = -2L + k\ln(N) \tag{4}$$

where L is the maximum log-likelihood, k is the number of free parameters of a model, and N is the number of observed data points. As is readily apparent from Eq. 4, the BIC balances goodness-of-fit with parsimony. Models with smaller BICs are more competitive, as those with a greater number of free parameters (which should produce a better fit) are penalized.

Table 1 shows the values of the BICs for all four models and observers, the summed totals across the sample, as well as the BICs for the fits of the saturated Gamma functions for each observer. The Gamma function is the clear overall winner with the lowest BIC – the BIC difference with the saturated Gaussian model is 41, which corresponds to a corrected likelihood-ratio, or Bayes factor, of greater than 1000 (Wagenmakers, 2007). Thus it is clear that fitting the data with a Gaussian function results in a less desirable fit of the data; however, the Gaussian still offers a considerable advantage in terms of the interpretability of the parameters. Moreover, inspection of a plot of the weightings based on the best-fitting Gaussian and Gamma (both full and simplified versions) models shows that all three functions fit the data reasonably well (see **Figure 6**).

With respect to the Gaussian model comparisons then, the five-parameter location model is the winning model by a clear margin. In other words, allowing a peak shift, but not a width shift, provides the best description of the data. The BIC difference with the nearest competitor – the saturated Gaussian model – is ~37, which again corresponds to a Bayes factor, of greater than 1000. The direct comparison between the width and location models also came out strongly in favor of the location model (a combined BIC difference of almost 100). In both circumstances, the size of the Bayes factor corresponds to an effective p value of <0.01 (Wagenmakers, 2007). In conclusion, it seems unlikely that allowing the width of the integration filters to vary with contrast improves our model fits in any meaningful way – it is the location of the peak of the filter that is important. Across all six observers, this shift in μ corresponded to a peak shift of 18 ms as contrast is reduced, comparable to the shift found for the Gamma fits.

Table 1 | Bayesian Information Criterions (and summed BICs) for all six observers for each of the four Gaussian comparison models.

	Baseline	Location	Width	Saturated	Gamma
OBSERVER					
1	1256.1	1221.9	1262.9	1224.2	1220.4
2	3717.0	3723.2	3724.3	3730.6	3703.6
3	1343.7	1333.9	1350.2	1340.2	1339.2
4	1230.8	1234.2	1238.1	1241.1	1239.7
5	1137.9	1116.1	1144.5	1123.4	1114.7
6	1438.3	1438.4	1445.1	1445.2	1444.9
Σ	10123.8	10067.7	10165.1	10104.7	10062.5

The model with the lowest BIC is considered the most preferable. Note that for the Gamma model, a saturated model is used, in which parameters are fit separately for the high and low-contrast data.

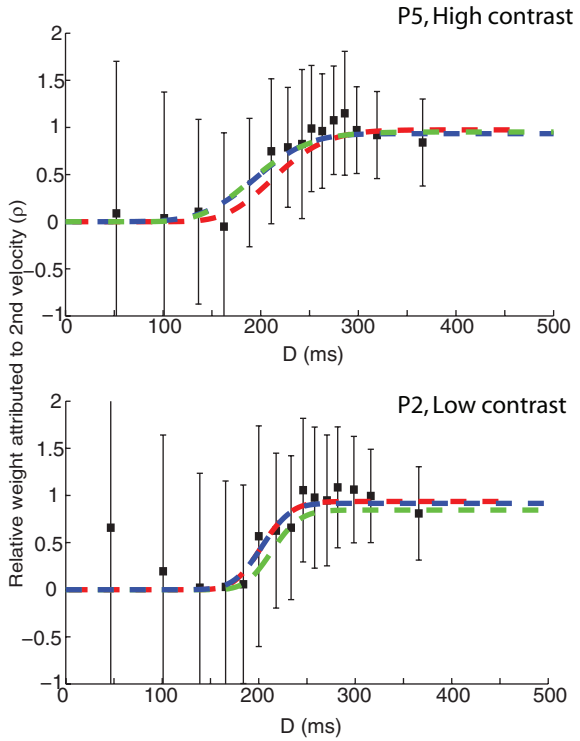


FIGURE 6 | Comparison of the weighting function fits based on the original Gamma model (green), the simplified Gamma model (red), and a Gaussian (blue). The top panel shows the participant data set with the highest correlation between the model fits (P5).The bottom panel shows the participant data set with the lowest correlation between the model fits (P2).

DISCUSSION

Sensorimotor decisions in a dynamic world necessarily involve an element of prediction. Behavior needs to be adapted to the future characteristics of the targeted object (e.g., its position). Careful analysis of variable single-trial behavior in response to a perturbation of the sensory input affords insight over what interval the relevant object characteristics are estimated. In the present study, we simplified and adjusted our method (Etchells et al., 2010) for identifying the temporal filter that is used by observers to estimate the velocity of a moving object. This estimate then guides the observers' prediction about the future location of the object, taking into account the interval between the decision to move and the completion of that movement.

Single cell recording in the primate brain and VEP studies in humans generally suggest an increase in conduction latency with lower contrast (e.g., Shapley and Victor, 1978; Kuba and Kubova, 1992; Kubova et al., 1995; Bach and Ullrich, 1997). Moreover, neuronal conduction latencies generally increase in higher cortical areas (for example, in MT and beyond) as successive stages of processing are added (Raiguel et al., 1999), and this effect will be amplified by a reduction in contrast. Behavioral studies have often measured reaction times, either of manual button presses (e.g., Mansfield, 1973; Breitmeyer, 1975; Harwerth and Levy, 1978; Plainis and Murray, 2000; Murray and Plainis, 2003)or saccades (Carpenter, 2004; Ludwig et al., 2004; Taylor et al., 2006; White et al., 2006). Generally speaking, reductions in contrast diminish the behavioral responsiveness to changes in the stimulus.

However, to get from the brain and responses of single cells to general behavior requires postulating internal information processing mechanisms. For the case of intercepting moving targets, any reduced *behavioral* responsiveness to velocity changes at low-contrast could come about either through a widening of the temporal filter used to estimate instantaneous velocity, and/or through a delay in the input to the integration mechanism. It was by no means obvious *a priori* which mechanism(s) would be responsible for diminished visual–saccadic performance. Our analysis method allowed us to distinguish between these possibilities. We obtained strong evidence in favor of a shift of the filter toward longer latencies, most likely produced by the longer input delays reviewed above. We found no evidence in favor of a consistent increase in the width of the filters at the lower contrast level.

Whilst our method cannot tell us where velocity integration for saccadic planning resides in the brain, it would not be unreasonable to suggest area MT as a likely candidate. Area MT is important for motion integration (Born and Bradley, 2005), and is known to project to eye movement related structures such as the superior colliculus (Ungerleider et al., 1984) and frontal eye fields (Tian and Lynch, 1996; Leigh and Zee, 2006). As noted above, lowering the contrast of moving visual stimuli results in an increase in the time that it takes for signals to reach MT (e.g., Kubova et al., 1995; Bach and Ullrich, 1997). It would be a matter for future research, perhaps using imaging techniques, in order to determine the likely cortical location of this mechanism.

Using a closely related approach, Tavassoli and Ringach (2009) identified the temporal filter driving eye velocity during smooth pursuit. Their pursuit stimulus was perturbed with Gaussian velocity noise. By correlating the velocity of the noisy stimulus with the eye velocity at different lags, they were able to estimate the latency with which the system responds to a perturbation, as well as the interval over which velocity information was integrated. For a large reduction in target contrast comparable to that of the current study (although note that their lowest contrast level was lower than ours), the filters showed a shift in the time-to-peak on the order of ~30 ms. In addition, they also found a moderate increase in the width of the filters of ~20. Given the evidence for shared visual processing between the pursuit and saccadic systems (Liston and Krauzlis, 2005; Orban de Xivry and Lefèvre, 2007), it would be reasonable to suggest that any neuronal effects of target speed estimation on smooth pursuit might also be reflected in the saccadic system. Indeed, one of the major inputs shared between the two systems appears to be target velocity (e.g., Newsome et al., 1985; Gellman and Carl, 1991; de Brouwer et al., 2002).

Given that lowering the contrast of a stimulus results in a decrease in the SNR (e.g., Weiss et al., 2002), we expected that a widening of the temporal filters might be a useful way in which to counteract this decrease. One reason that we did not see this decrease is that the contrast level we used in the low-contrast condition was simply not close enough to threshold for it to cause real issues with the SNR. Additionally, widening the integration epoch in order to boost the SNR of the internal velocity estimate may only be adaptive in situations where the velocity remains

(approximately) constant. In realistic terms, moving objects are often subject to changes in velocity. In such an environment, widening the temporal filter might not be useful, as it would hinder the fidelity with which the observer could track changes in velocity, and ultimately result in less accurate estimates of how fast a target is moving. In our present study, observers are presented with a mix of constant velocity and changing velocity trials. Importantly, the velocity change was relatively large (e.g., compared to the zero-mean noise used in Tavassoli and Ringach, 2009). It may be that the cost of widening the filter under these conditions was simply too large. The necessity for rapidity of response might well have outweighed any need for increasing the SNR. Note that if this explanation is correct, it suggests a degree of flexibility in the system so that adaptation of the integration period depends on the wider context in which the system operates.

The methodology that we present here sits well within the broader context of research assessing perceptual and decision-making behavior on single trials. Using a relatively simple experimental paradigm (i.e., tracking primary orienting eye movements to a moving target), quite complex and temporally precise knowledge about how visual information is used within the saccadic latency period can be gathered. By varying target contrast and the timing of the saccadic go signal from trial-to-trial, we are able to show not only that the integration of visual (in this case, velocity) information can be modified during the latency period, but we are also able to show *how* it is modified. By assessing these results within the context of similar studies (e.g., Caspi et al., 2004; Ludwig et al., 2005), we can build up a coherent picture about the time course of how eye movement decisions are made using simple behavioral paradigms. This allows for a great deal of flexibility in the methods and models that are implemented, and allows for simple modifications in order to answer further questions about what exactly occurs during the saccadic latency period. For example, the use of non-constant velocities in the present experiment would aid in the assessment of how the integration filters deal with acceleration. In turn, this may provide an understanding as to how acceleration is represented within areas such as MT (e.g., Schlack et al., 2007, 2008), and also how sampling mechanisms are used in the predict drive for ocular pursuit (see Bennett et al., 2007). Along more general lines, such approaches may help to further our understanding of how the smooth pursuit and saccadic systems interact and coordinate with each other.

CONCLUSION

The simplified methodology presented in this paper provides a novel addition to a growing toolbox for the behavioral study of how information on single trials may be integrated and used in eye movement programming. Our approach allows for maximum likelihood estimates of the parameters of the temporal filter that best accounts for the interval over which the saccadic system samples the velocity of a moving target. The likelihoods provide a solid basis upon which to assess the significance of experimentally targeted variables, such as contrast in this study. More generally, we believe this methodology more uniquely constrains internal processing mechanisms that transform sensory inputs into motor outputs, linking brain and behavior.

REFERENCES

Aslin, R. N., and Shea, S. L. (1987). The amplitude and angle of saccades to double-step target displacements. *Vision Res.* 27, 1925–1942.

Bach, M., and Ullrich, D. (1997). Contrast dependency of motion-onset and pattern-reversal VEPs: interaction of stimulus type, recording site and response component. *Vision Res.* 37, 1845–1849.

Becker, W., and Jürgens, R. (1979). An analysis of the saccadic system by means of double-step stimuli. *Vision Res.* 19, 967–983.

Bennett, S. J., Orban de Xivry, J. J., Barnes, G. R., and Lefèvre, P. (2007). Target acceleration can be extracted and represented within the predictive drive to ocular pursuit. *J. Neurophysiol.* 98, 1405–1414.

Benton, C. P., and Curran, W. (2009). The dependence of perceived speed upon signal intensity. *Vision Res.* 49, 284–286.

Beutter, B. R., Eckstein, M. P., and Stone, L. S. (2003). Saccadic and perceptual performance in visual search tasks. I. Contrast detection and discrimination. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* 20, 1341–1355.

Born, R. T., and Bradley, D. C. (2005). Structure and function of visual area MT. *Annu. Rev. Neurosci.* 28, 157–189.

Brainard, D. H. (1997). The psychophysics toolbox. *Spat. Vis.* 10, 433–436.

Breitmeyer, B. G. (1975). Simple reaction time as a measure of the temporal response properties of transient and sustained channels. *Vision Res.* 15, 1411–1412.

Burnham, K. P., and Anderson, D. R. (2002). *Model Selection and Inference: A Practical Information-Theoretic Approach*. New York: Springer.

Carpenter, R. H. S. (2004). Contrast, probability and saccadic latency: evidence for independence of detection and decision. *Curr. Biol.* 14, 1576–1580.

Caspi, A., Beutter, B. R., and Eckstein, M. P. (2004). The time course of visual information accrual guiding eye movement decisions. *Proc. Natl. Acad. Sci. U.S.A.* 101, 13086–13090.

Cornelissen, F. W., Peters, E. M., and Palmer, J. (2002). The eyelink toolbox: eye tracking with MATLAB and the psychophysics toolbox. *Behav. Res. Methods Instrum. Comput.* 34, 613–617.

de Brouwer, S., Missal, M., Barnes, G., and Lefèvre, P. (2002). Quantitative analysis of catch-up saccades during sustained pursuit. *J. Neurophysiol.* 87, 1772–1780.

Eckstein, M. P., Beutter, B. R., Pham, B. T., Shimozaki, S. S., and Stone, L. S. (2007). Similar neural representations of the target for saccades and perception during search. *J. Neurosci.* 27, 1266–1270.

Efron, B., and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap. Monographs on Statistics and Applied Probability*. Boca Raton, FL: Chapman & Hall/CRC.

Enroth-Cugell, C., and Robson, J. G. (1966). The contrast sensitivity of retinal ganglion cells of the cat. *J. Physiol.* 187, 517–552.

Etchells, P. J., Benton, C. P., Ludwig, C. J. H., and Gilchrist, I. D. (2010). The target velocity integration function for saccades. *J. Vis.* 10, 1–14.

Findlay, J. M., and Harris, L. R. (1984). "Small saccades to double-stepped targets moving in two dimensions," in *Theoretical and Applied Aspects of Eye Movement Research*, eds A. G. Gale and F. Johnson (Amsterdam: Elsevier), 71–78.

Gellman, R. S., and Carl, J. R. (1991). Motion processing for saccadic eye movements in humans. *Exp. Brain Res.* 84, 660–667.

Glimcher, P. W. (2001). Making choices: the neurophysiology of visual-saccadic decision making. *Trends Neurosci.* 24, 654–659.

Gold, J. I., and Shadlen, M. N. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends Cogn. Sci.* 5, 10–16.

Gold, J. I., and Shadlen, M. N. (2007). The neural basis of decision making. *Annu. Rev. Neurosci.* 30, 535–574.

Harwerth, R. S., and Levy, D. M. (1978). Reaction time as a measure of suprathreshold grating detection. *Vision Res.* 18, 1579–1586.

Kaplan, E., and Shapley, R. M. (1982). X and Y cells in the lateral geniculate nucleus of macaque monkeys. *J. Physiol.* 330, 125–143.

Kerzel, D., and Gegenfurtner, K. R. (2003). Neuronal processing delays are compensated in the sensorimotor branch of the visual system. *Curr. Biol.* 13, 1975–1978.

Krekelberg, B., van Wezel, R. J., and Albright, T. D. (2006). Interactions between speed and contrast tuning in the middle temporal area: implications for the neural code for speed. *J. Neurosci.* 26, 8988–8998.

Kuba, M., and Kubova, Z. (1992). Visual evoked potentials specific for motion-onset. *Doc. Ophthalmol.* 80, 83–89.

Kubova, Z., Kuba, M., Spekreijse, H., and Blakemore, C. (1995). Contrast dependence of motion-onset and pattern-reversal evoked potentials. *Vision Res.* 35, 197–205.

Leigh, R. J., and Zee, D. S. (2006). *The Neurology of Eye Movements*, 4th Edn. Oxford: OUP.

Liston, D., and Krauzlis, R. J. (2005). Shared decision signal explains performance and timing of pursuit and saccadic eye movements. *J. Vis.* 5, 678–689.

Livingstone, M. S., and Conway, B. R. (2006). Contrast affects speed tuning, space-time slant, and receptive-field organisation of simple cells in macaque V1. *J. Neurophysiol.* 97, 849–857.

Ludwig, C. J. H. (2009). Temporal integration of sensory evidence for saccade target selection. *Vision Res.* 49, 2764–2773.

Ludwig, C. J. H., Gilchrist, I. D., and McSorley, E. (2004). The influence of spatial frequency and contrast on saccade latencies. *Vision Res.* 44, 2597–2604.

Ludwig, C. J. H., Gilchrist, I. D., McSorley, E., and Baddeley, R. J. (2005). The temporal impulse response underlying saccadic decisions. *J. Neurosci.* 25, 9907–9912.

Ludwig, C. J. H., Mildinhall, J. W., and Gilchrist, I. D. (2007). A population coding account for systematic variation in saccadic dead time. *J. Neurophysiol.* 97, 795–805.

Mansfield, R. J. W. (1973). Latency functions in human vision. *Vision Res.* 13, 2219–2234.

Marsaglia, G. (2006). Ratios of normal variables. *J. Stat. Softw.* 16, 1–10.

Maunsell, J. H. R., Ghose, G. M., Assad, J. A., McAdams, C. J., Boudreau, C. E., and Noerager, B. D. (1999). Visual response latencies of magnocellular and parvocellular LGN neurons in macaque monkeys. *Vis. Neurosci.* 16, 1–14.

Murray, I. J., and Plainis, S. (2003). Contrast coding and magno/parvo segregation revealed in reaction time studies. *Vision Res.* 43, 2707–2719.

Nelder, J. A., and Mead, R. (1965). A simplex-method for function minimization. *Comput. J.* 7, 308–313.

Newsome, W. T., Wurtz, R. H., Dursteler, M. R., and Mikami, A. (1985). Deficits in visual-motion processing following ibotenic acid lesions of the middle temporal visual area of the macaque monkey. *J. Neurosci.* 5, 825–840.

Nickerson, R. S., and Burnham, D. W. (1969). Response times with nonaging foreperiods. *J. Exp. Psychol.* 79, 452–457.

Nummela, S. U., Lovejoy, L. P., and Krauzlis, R. J. (2008). Saccade selection when reward probability is dynamically manipulated using Markov chains. *Exp. Brain Res.* 187, 321–330.

Orban de Xivry, J. J., and Lefèvre, P. (2007). Saccades and pursuit: two outcomes of a single sensorimotor process. *J. Physiol.* 584, 11–23.

Oswal, A., Ogden, M., and Carpenter, R. H. S. (2007). The time course of stimulus expectation in a saccadic decision task. *J. Neurophysiol.* 97, 2722–2730.

Pack, C. C., Hunter, J. N., and Born, R. T. (2005). Contrast dependence of suppressive influences in cortical area MT of alert macaque. *J. Neurophysiol.* 93, 1809–1815.

Plainis, S., and Murray, I. J. (2000). Neurophysiological interpretation of human visual reaction times: effect of contrast, spatial frequency and luminance. *Neuropsychologia* 38, 1555–1564.

Raiguel, S. E., Xiao, D. K., Marcar, V. L., and Orban, G. A. (1999). Response latency of macaque area MT/V5 neurons and its relationship to stimulus parameters. *J. Neurophysiol.* 82, 1944–1956.

Schall, J. D. (2003). Neural correlates of decision processes: neural and mental chronometry. *Curr. Opin. Neurobiol.* 13, 182–186.

Schlack, A., Krekelberg, B., and Albright, T. D. (2007). Recent history of stimulus speeds affects the speed tuning of neurons in area MT. *J. Neurosci.* 27, 11009–11018.

Schlack, A., Krekelberg, B., and Albright, T. D. (2008). Speed perception during acceleration and deceleration. *J. Vis.* 8, 1–11.

Schwarz, G. (1978). Estimating the dimensions of a model. *Ann. Stat.* 6, 461–464.

Seijas, O., Gómez de Liaño, P., Gómez de Liaño, R., Roberts, C. J., Piedrahita, E., and Diaz, E. (2007). Ocular dominance diagnosis and its influence in monovision. *Am. J. Ophthalmol.* 144, 209–216.

Shapley, R. M., and Victor, J. D. (1978). The effect of contrast on the transfer properties of cat retinal ganglion cells. *J. Physiol.* 285, 275–298.

Smith, P. L. (1995). Psychophysically principled models of visual simple reaction time. *Psychol. Rev.* 102, 567–593.

Spering, M., Kerzel, D., Braun, D. I., Hawken, M. J., and Gegenfurtner, K. R. (2005). Effects of contrast on smooth pursuit eye movements. *J. Vis.* 5, 455–465.

Tavassoli, A., and Ringach, D. L. (2009). Dynamics of smooth pursuit maintenance. *J. Neurophysiol.* 102, 110–118.

Taylor, M. J., Carpenter, R. H. S., and Anderson, A. J. (2006). A noisy transform predicts saccadic and manual reaction times to changes in contrast. *J. Physiol.* 573, 741–751.

Tian, J. R., and Lynch, J. C. (1996). Corticocortical input to the smooth and saccadic eye movement subregions of the frontal eye field in the Cebus monkey. *J. Neurophysiol.* 76, 2754–2771.

Ungerleider, L. G., Desimone, R., Galkin, T. W., and Mishkin, M. (1984). Subcortical projections of area MT in the macaque. *J. Comp. Neurol.* 223, 368–386.

Wagenmakers, E. J. (2007). A practical solution to the pervasive problems of p values. *Psychon. Bull. Rev.* 14, 779–804.

Walker, R., Walker, D. G., Husain, M., and Kennard, C. (2000). Control of voluntary and reflexive saccades. *Exp. Brain Res.* 130, 540–544.

Watson, A. B. (1986). "Temporal sensitivity," in *Handbook of Perception and Human Performance*, eds K. R. Boff, L. Kaufman, and J. P. Thomas (New York: Wiley), 6.1–6.43.

Weiss, Y., Simoncelli, E. P., and Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nat. Neurosci.* 5, 598–604.

White, B. J., Kerzel, D., and Gegenfurtner, K. R. (2006). The spatio-temporal tuning of the mechanisms in the control of saccadic eye movements. *Vision Res.* 46, 3886–3897.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 02 March 2011; paper pending published: 03 May 2011; accepted: 16 May 2011; published online: 26 May 2011.
Citation: Etchells PJ, Benton CP, Ludwig CJH and Gilchrist ID (2011) Testing a simplified method for measuring velocity integration in saccades using a manipulation of target contrast. Front. Psychology ***2****:115. doi: 10.3389/fpsyg.2011.00115*
This article was submitted to Frontiers in Perception Science, a specialty of Frontiers in Psychology.

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 28 February 2011
doi: 10.3389/fpsyg.2011.00030

Single-trial regression elucidates the role of prefrontal theta oscillations in response conflict

Michael X Cohen[1,2]* and James F. Cavanagh[3]

[1] *Department of Psychology, University of Amsterdam, Amsterdam, Netherlands*
[2] *Department of Physiology, University of Arizona, Tucson, AZ, USA*
[3] *Cognitive, Linguistic, and Psychological Sciences, Brown University, Providence, RI, USA*

Edited by:
Guillaume A. Rousselet, University of Glasgow, UK

Reviewed by:
Gregor Thut, University of Glasgow, UK
Ali Mazaheri, University of California, USA

****Correspondence:***
Michael X Cohen, Department of Psychology, University of Amsterdam, Roetersstraat 15, 6th Floor, 1018 WB Amsterdam, Netherlands
e-mail: mikexcohen@gmail.com

In most cognitive neuroscience experiments there are many behavioral and experimental dynamics, and many indices of brain activity, that vary from trial to trial. For example, in studies of response conflict, conflict is usually treated as a binary variable (i.e., response conflict exists or does not in any given trial), whereas some evidence and intuition suggests that conflict may vary in intensity from trial to trial. Here we demonstrate that single-trial multiple regression of time–frequency electrophysiological activity reveals neural mechanisms of cognitive control that are not apparent in cross-trial averages. We also introduce a novel extension to oscillation phase coherence and synchronization analyses, based on "weighted" phase modulation, that has advantages over standard coherence measures in terms of linking electrophysiological dynamics to trial-varying behavior and experimental variables. After replicating previous response conflict findings using trial-averaged data, we extend these findings using single-trial analytic methods to provide novel evidence for the role of medial frontal–lateral prefrontal theta-band synchronization in conflict-induced response time dynamics, including a role for lateral prefrontal theta-band activity in biasing response times according to perceptual conflict. Given that these methods shed new light on the prefrontal mechanisms of response conflict, they are also likely to be useful for investigating other neurocognitive processes.

Keywords: EEG, oscillations, response conflict, cognitive control, single trial, regression, theta, medial frontal cortex

THE "MANY-TO-MANY MAPPING" PROBLEM OF COGNITIVE NEUROSCIENCE

In this report, we demonstrate how single-trial multiple regression analyses can help elucidate brain-behavior relationships, specifically linking cortical electrophysiological oscillatory dynamics to cognitive control processes. The heart of the issue is attempting to resolve the "many-to-many" mapping problem: in many cognitive neuroscience experiments, there are many behavioral and many experimental dynamics, and many indices of electrophysiological brain activity (e.g., over time, space, frequency), that vary from trial to trial. Determining which measurements of behavioral dynamics correspond to which measurements of brain activity is difficult but of central importance to cognitive neuroscience.

Experimental or behavioral variables that vary from trial to trial are often ignored or amalgamated, thereby reducing "many" to "few." The logic behind trial averaging is that, at the single-trial level, brain measurement tools (EEG, MEG, fMRI) and the neurocognitive systems they measure contain more noise than signal; thus, by averaging data over many trials of the same or similar experiment condition, signal-to-noise ratio increases and randomly distributed variance averages out. This reasoning is irrefutable – the influence of noise decreases as a function of the number of trials $\left(\frac{\text{noise}}{\sqrt{\text{trails}}}\right)$, and some cross-trial variance is unrelated to the hypotheses under investigation. In other situations, however, hypotheses must or should be tested using data from single trials within subjects, for example when linking neural dynamics to response time (Weissman et al., 2006; Yarkoni et al., 2009), visual stimulus parameters (Rousselet et al., 2008; Scholte et al., 2009), decision-making (Philiastides and Sajda, 2007; Ratcliff et al., 2009), or other parameters that vary from trial to trial.

CORRELATING SINGLE-TRIAL BRAIN AND BEHAVIOR DYNAMICS

Although most cognitive neuroscience studies use cross-trial averaging, there are many examples of how single-trial analyses have yielded important insights into neurocognitive function. Single-trial analyses have the obvious advantage of finding linear relationships between brain activity and trial-varying behavior or experimental manipulations. Within the field of cognitive control, for example, single-trial analyses have linked online and reactive behavior adaptations to medial frontal EEG and fMRI responses (Debener et al., 2005; Mars et al., 2008; Boehler et al., 2010; Cavanagh et al., 2010). One limitation, however, is that with simple correlations, only one variable should be tested. Multiple simple correlations may be suboptimal because shared variance among variables can bias correlation coefficients.

A DYNAMIC SOLUTION: SINGLE-TRIAL MULTIPLE REGRESSION

The approach we advocate here is an extension of the single-trial correlation approach. Multiple regression has several important advantages over correlation. First, many independent variables can be entered into the regression, and variance due to different variables can be parceled out. Second, interaction terms between

variables can be constructed to estimate possible non-linear effects of combinations of experimental and behavior variables. Third, with multiple regression one can examine not only the slopes (i.e., linear relationship between two variables) but also the intercept of the model ("DC" or mean-offset of the relationship), which may be useful for dissociating random (i.e., idiosyncratic relationships with an independent variable) versus fixed (i.e., all subjects show a general increase in brain activity as a function of condition) effects. Forth, related extensions to multiple regression such as hierarchical linear modeling are amenable for more in-depth analyses into differences across groups of subjects.

Single-trial multiple regression has been applied to EEG data. For example, multiple regression of visual stimulus properties demonstrated early responses to noise and feature processing during face viewing (Rousselet et al., 2008), which is affected by aging (Rousselet et al., 2009). Eichele et al. (2010) recently used single-trial multiple regression to remove variance due to factors less relevant to the hypotheses (e.g., P300 modulation). Regression is also one method for removing blink and other artifacts (Schwind and Dormann, 1986). In these studies, regressions were performed on time-domain EEG data. However, because EEG data are driven largely by oscillatory cortical processes, considerable information in EEG may be contained in the frequency domain and therefore lost in the time domain (Cohen, 2011b). Therefore, in this study we conducted single-trial multiple regression over time–frequency estimates of power, and we introduce a method to link non-linear phase distributions to trial-varying behavioral and experimental variables.

RESPONSE CONFLICT: A TRIAL-VARYING PHENOMENA THAT IS OFTEN AVERAGED OVER IN EXPERIMENTS

Response conflict occurs when multiple response options are activated, but only one should be selected according to task demands. Response conflict activates the cognitive control system, a set of high-level neurocognitive processes that monitor behavior and the environment for errors, potential errors, and negative performance feedback, and facilitate flexible and adaptive adjustments in behavior to improve future performance. It is thought that structures in the medial frontal cortex are a fulcrum for the cognitive control system, and work with other prefrontal and task-specific sensory and motor regions to support flexible behavior adaptation (Ridderinkhof et al., 2004). The cognitive control system is thought to wane when no response conflict is detected, and to wax when response conflict is detected. Consequently, a trial containing response conflict elicits a reactive activation of the cognitive control system only when the preceding trial contained no response conflict. This phenomenon is termed the "Gratton effect" (Gratton et al., 1992), has been observed in a variety of empirical studies (Egner, 2007), and can be captured by mathematical models (Botvinick et al., 2001). For this reason, researchers often separate trials according to the response conflict on the previous as well as on the current trial.

Activation of the cognitive control system (e.g., during response conflict) can be measured through a variety of dependent variables, including behavioral (reaction time and accuracy), hemodynamic, and electrophysiological. Relevant for the present study, response conflict (and other cognitive control situations such as errors or negative performance feedback) increases theta-band oscillatory activity over medial frontal cortex (Luu and Tucker, 2001; Cohen et al., 2008; Hanslmayr et al., 2008; Cavanagh et al., 2009; Cohen et al., 2009). This medial frontal theta has been proposed to reflect an electrophysiological mechanism for coordinating neural networks involved in monitoring behavior and the environment as well as facilitating task-specific adaptive changes in performance in conjunction with lateral prefrontal cortex and sensory-motor areas.

In typical cognitive control experiments, response conflict is induced either at the response, the perceptual, or the semantic level. Conflict at the response level is elicited by priming two competing responses when only one is correct. At the stimulus-level, conflict can be induced by making the stimuli ambiguous (Szmalec et al., 2008) or low in luminance (Yeung et al., 2007). These examples reflect the common treatment of conflict as a discrete variable, such that conflict is assumed to be present or absent on any given trial. For example, these stimulus-level conflict studies used dim vs. bright stimuli, or high vs. low ambiguous stimuli (Yeung et al., 2007; Szmalec et al., 2008). Although these experimental manipulations are categorical (i.e., trials either contain or do not contain response conflict), the effects of these manipulations on internally experienced conflict may not be discrete, but rather may vary from trial to trial (e.g., the Gratton effect).

It is clear that response conflict is not an all-or-none phenomenon in the brain. Given that the strength of conflict elicited by exogenous (i.e., parametric experimental manipulations not under subjects' control) and endogenous (i.e., fluctuations in internal cognitive processes) factors may vary from trial to trial, it is apparent that trial averaging provides a limited characterization of the neural mechanisms underlying cognitive control processes. Therefore, alternative analytical techniques are needed to link trial-varying behavioral dynamics to corresponding trial-varying neural dynamics. Thus, the purposes of this experiment were to use single-trial regression to isolate spatial–temporal–frequency characteristics predicted by (1) exogenously induced conflict via continuous (trial-varying luminance) and discrete (trial type) manipulations, and (2) endogenously experienced conflict, as measured through reaction time. Reaction times are often used as a dependent measure to quantify the behavioral effects of conflict (Gratton et al., 1992; Egner, 2007), but to the extent that they reflect internally experience conflict, reaction times can be used as an independent variable (Weissman et al., 2006; Forstmann et al., 2008). Here, EEG data were transformed into their time–frequency representation, and these estimates of time-, space-, and frequency-specific power were regressed against reaction time, stimulus luminance, and the interaction between the two. We also introduce a method to link these single-trial experimental dynamics to oscillation phase angle ("weighted" phase modulation), which has advantages over standard inter-trial phase coherence measures.

METHODS

SUBJECTS

Seventeen subjects from the University of Amsterdam community (aged: 18–31, two male) participated in exchange for course credit or 14 Euros. Subjects were self-reported free of neurological diseases and signed informed consent documents that were approved by the local ethics committee. Data from two subjects were removed

because one subject had excessive noise and EMG artifacts (over 30% of trials were rejected) and one subject did not complete the experiment because he felt ill. Thus, data from 15 subjects were included in the final analyses.

TASK

A modified flankers task was used, in which subjects responded to a central target letter while ignoring flanking letters. There were 1200 trials in three blocks. Each block contained two letters (M and N, E and F, and X and Y, counter-balanced across subjects); one letter required a left-hand response and the other required a right-hand response. "Congruent" trials contained the same flanking and target letter (e.g., XXXXX); "incongruent" trials contained different target and flanking letters (e.g., XXYXX). Thus, conflict was induced when the flanking letters were associated with a different response compared to the target letter. Subjects' eyes were approximately 90 cm from the monitor, making the letters 0.382° high and 0.2546° wide, with a 0.2546° blank space separating each letter. The target and flankers were presented for 100 ms on a white background. A 1200-ms inter-trial-interval separated trials. Luminance was randomly selected on each trial to be between 0 and 200 (RGB values that vary between 0 and 255). Thus, higher luminance values mean the stimuli were closer to the background color and therefore more difficult to see. Luminance was applied equally to target and flankers. This was done, rather than, e.g., manipulating luminance only of the flankers, for two reasons: (1) This manipulation is experimentally orthogonal to the conflict conditions and therefore is appropriate for multiple regression with interaction terms; (2) manipulating the luminance of all stimuli is more experimentally tractable because differential luminance might produce non-linear net luminance effects on striate cortex at the level measured by EEG. There are many ways to modulate stimulus-level conflict in such tasks, including luminance, size, eccentricity, onset times, etc.; we would expect to find overall similar patterns of results when using different stimulus-level manipulations.

As discussed in the Section "Introduction," response conflict effects depend on whether conflict was present in the previous trial. Therefore, trials were discretized into four conditions: cC, cI, iC, and iI ("cI" means that the previous trial was congruent and the current trial was incongruent). Trials containing errors, and the trials thereafter, were removed prior to analyses, as were the first trials following each rest break. After EEG trial rejection (described below), there were on average, respectively, 213, 307, 307, and 177 trials per condition (minimum/maximum across subjects, respectively: 185/244, 241/358, 240/357, 84/206).

EEG DATA COLLECTION

EEG data were acquired at 512 Hz from 64 channels placed according to the international 10–20 system and from both earlobes (used as reference). Offline, EEG data were high-pass filtered at 0.5 Hz and then epoched from −1.5 to +2 s surrounding each trial (to avoid edge artifacts resulting from wavelet filtering). All trials were visually inspected and those containing EMG or other artifacts not related to blinks were manually removed. Independent components analysis was computed using EEGLAB software (Delorme and Makeig, 2004), and components containing blink/oculomotor artifacts or other artifacts that could be clearly distinguished from brain-driven EEG signals were subtracted from the data.

EEG ANALYSES: TIME–FREQUENCY DECOMPOSITION

All analyses were performed in matlab. Single-trial data were first decomposed into their time–frequency representation by multiplying the power spectrum of the EEG (obtained from the fast-Fourier-transform) by the power spectrum of complex Morlet wavelets [$e^{i2\pi tf}e^{-t^2/(2\sigma^2)}$, where t is time, f is frequency, which increased from 1 to 40 Hz in 30 logarithmically spaced steps, and σ defines the width of each frequency band, set according to $4/(2\pi f)$], and then taking the inverse fast-Fourier-transform. From the resulting complex signal, an estimate of frequency band-specific power at each time point was defined as the squared magnitude of the result of the convolution Z (real$[z(t)]^2$ + imag$[z(t)]^2$), and an estimate of frequency band-specific phase at each time point was taken as the angle of the convolution result. Relatively long epochs were cut from the continuous EEG data (−1.5 to 2 s) to allow edge artifacts due to sudden transitions in signal values between trials to subside outside the window of interest. Taking long epochs and trimming edge artifacts is preferred over windowing, because the latter method attenuates real signal whereas the former does not. Power was normalized using a decibel (dB) transform (dB power = 10 × log 10[power/baseline]), where the baseline activity was taken as the average power at each frequency band, averaged across conditions, from −300 to −100 ms pre-stimulus. Conversion to a dB scale ensures that data across all frequencies, time points, electrodes, conditions, and subjects are in the same scale and thus comparable. Inter-trial phase coherence (the consistency of frequency band-specific phase angles over trials time-locked to the response) was computed according to $|\frac{1}{n}\times\sum_{t=1}^{n}e^{i[\phi_{jt}]}|$, where n is the number of trials, and ϕ_j are the phase angles of electrode j. Phase coherence varies from 0 (no phase consistency across trials) to 1 (oscillations take on identical phase values across trials; Lachaux et al., 1999; Delorme and Makeig, 2004). Frequency band-specific phase synchronization (functional connectivity) was computed according to $|\frac{1}{n}\times\sum_{t=1}^{n}e^{i[\phi_{jt}-\phi_{kt}]}|$, where n is the number of trials, and ϕ_j and ϕ_k are the phase angles of electrodes j and k. In all analyses and plots, data are time-locked to the response; thus time 0 in the figures corresponds to the button press.

EEG ANALYSES: SINGLE-TRIAL REGRESSION

Single-trial analyses were conducted separately for power and phase. For power analyses, a robust regression was computed that estimated parameters at each time–frequency–space point for the following linear equation: Y = INT + b_1RT + b_2LUM + b_3 RT × LUM + E. Y is the data vector (power values at each time–frequency point across trials), INT is the intercept, b_{1-3} are regression coefficients, E is unexplained variance, and RT and LUM are trial vectors of the subject's reaction time and the stimulus luminance on each trial. Reaction time and luminance data were z-scored so that the interaction term was not dominated by reaction time, which has values an order of magnitude larger than luminance (note that this means the intercept simply accounts for Power Law scaling across frequencies and therefore is not of interest here). Robust regression was used to minimize the

contribution of potential outliers via iterative reweighted least squares that minimizes the impact of outliers with large leverage (O'Leary, 1990). In this regard, robust regression has a significant advantage over trial averaging. Specifically, during standard trial-averaging, outliers may affect the averaged data. However, with robust regression, outliers are de-weighted and therefore have minimal effect on the overall result. Ultimately, this procedure resulted in a time × frequency × space (electrodes) × condition matrix of *b* values for each subject. The regression was conducted separately for each condition rather than including condition as a covariate because the four conditions are categorical. Because these *b* values are normally distributed under the null hypothesis, they can be entered into standard parametric statistics such as *t*-tests and repeated-measures ANOVA. Before averaging across subjects, *b* values were standardized by scaling the coefficients by their SDs; this ensured that the coefficients were in the same scale and thus directly comparable across time, frequency, electrodes, and subjects.

Single-trial phase values, however, cannot be entered into regression because the data are circular (e.g., radian phase values of −3.1 and 3.1 are closer to each other than are 0.1 and 1.0). Therefore, we used an alternative approach, based on the idea that under the null hypothesis of no relationship between, e.g., reaction time and phase values, reaction times across trials should be uniformly distributed across phase. The less uniform this distribution, the more evidence accumulates to reject the null hypothesis. Taking each reaction time–phase pair as a vector with the phase as the angle and reaction time as the length, the magnitude of the average vector can be taken as a modulation of reaction time by phase angle (under the null hypothesis of no relationship, the average vector length would be close to zero). Here, reaction time and luminance data were rank-transformed because this method cannot be used with negative-valued data. Two issues inherent in magnitude scaling and phase distribution require a non-parametric intervention prior to group-level statistical analyses. The first issue is that non-transformed magnitudes are difficult to interpret because they scale with the data (reaction time or luminance values), which was different across subjects (this can be contrasted with inter-trial phase coherence, described above, for which the average vector magnitudes have a maximum of 1.0 and are inherently interpretable). The second issue is that if phase values are non-uniformly distributed across trials (as would be expected if there is, e.g., stimulus-induced phase reset), the distribution of reaction time would be artificially non-uniformly distributed. To counteract both of these issues, we applied permutation testing, in which the observed reaction time and phase values were shuffled with respect to each other. This provides a data-driven test of the null hypothesis that there is no consistent relationship between reaction time and phase angle. Five hundred iterations with shuffled reaction time–phase pairings were performed at each point in time–frequency–electrode–condition space, thus creating a distribution of reaction time–phase modulations under the null hypothesis. Finally, the standardized distance between the observed modulation and the null distribution was taken as a *z* value corresponding to the probability of finding the observed reaction time–phase modulation by chance, given the measured data. These processing steps resulted in a time × frequency × space (electrodes) × condition matrix of *z* values for each subject, which, like the *b* values described for the power regression, can be tested using parametric statistics. This entire procedure was then redone for luminance and the reaction time–luminance interaction. Note that this method, like robust regression, minimizes the impact of outliers because the results are based on within-subject permutation testing of observed data.

INDEPENDENT COMPONENTS ANALYSES

In order to compare our results (using data recorded from the electrodes) with another approach often used for single-trial analyses (independent components that estimate temporally dissociable sources of variance), independent components were selected and their time courses were subjected to the same analyses described above. We selected three components for each subject, one representing medial frontal dynamics and two representing lateral prefrontal dynamics (centered on FCz, F5, and F6, which were previously used in cognitive control studies; Cavanagh et al., 2009, 2010). Component selection was done automatically based on maximal spatial correlation between the components and templates. The templates were Gaussians surrounding electrodes of interest. For two subjects, right prefrontal components were manually re-selected because the time course of the automatically selected component resembled blinks.

EEG STATISTICS

Group-level statistics were performed using ANOVAs. Data from each subject were taken from a 100-ms, 4-Hz window surrounding the condition-averaged peak time–frequency point. This provides a compromise between a subject-specific data-driven approach (each subject retains his or her unique time–frequency maxima) while addressing the large multiple comparison problem (there are 155,520 time–frequency–space pixels that could possibly be tested). In order to facilitate visualization of spatial topographies, *t*-tests at each electrode were performed and electrodes with non-significant results at an uncorrected two-tailed $p < 0.01$ were set to zero (green color). Similarly, black contour lines on the time–frequency plots indicate continuous significance at two-tailed $p < 0.01$, with a minimum of 300 ms and three frequency bands.

RESULTS

BEHAVIOR

As expected based on the Gratton effect, there was a significant previous × current conflict interaction (repeated-measures ANOVA, $F_{1,14} = 14.29$, $p = 0.002$). Reaction times were longest on cI trials, shortest on cC trials, and in between during iI and iC trials (**Figure 1A**). To examine the effects of stimulus luminance on reaction times, we correlated, for each subject, luminance and reaction time separately for each condition (**Figure 1B**), and then tested those correlation coefficients across subjects. Correlation coefficients were significantly greater than zero in all conditions except for cC (*p*-values: 0.53, <0.001, 0.002, 0.033 for cC, cI, iC, and iI conditions), although the interaction term in a previous × current conflict repeated-measures ANOVA was not significant ($F_{1,14} = 0.59$). These results indicate that stimulus luminance affected subjects' performance only when the current or previous trial contained conflict.

TRIAL-AVERAGED EEG RESULTS

The trial-averaged time–frequency–space characteristics of the data provide a useful comparison for the results from the single-trial analyses. In general, the task elicited increased theta-band activity over medial frontal electrodes, centered around FCz, and maximal just prior to the response (**Figure 2A**). Although theta and delta power were significantly increased during all trials compared to baseline (**Figure 2C**), there was no significant main effect of current trial conflict, nor was there a significant previous × current trial conflict interaction (2 × 2 repeated-measures ANOVA: all p's > 0.05). Inter-trial phase coherence (the consistency of frequency band-specific phase angles across trials) showed low-frequency phase coherence and no clear medial frontal spatial focus (**Figures 2B,D**).

SINGLE-TRIAL REGRESSIONS

Figure 3 shows standardized regression coefficients from the single-trial multiple regression of reaction time (**Figure 3A**), luminance (**Figure 3B**), and their interaction (**Figure 3C**). A robust main effect of reaction time can be seen in the theta range prior to the response, centered around frontal sites. These coefficients were significantly greater than zero in all conditions (**Figure 3A**), and there was a significant interaction between previous and current conflict ($F_{1,14}$ = 11.67, p = 0.004), which was driven by significantly larger coefficients in cI compared to cC trials (t_{14} = 3.48, p = 0.0037), and no significant difference between iI and iC trials (t_{14} = 1.19, p = 0.187). The relationship between brain activity and RT may reflect conflict and also other response-related processes; however, because basic motoric response-related processes are present during all responses, the difference between regression coefficients in cC and cI trials reflects conflict processing.

There was a main effect of luminance in low frequencies (delta range, ~1–3 Hz) at central electrodes, mainly in cC trials (**Figure 3B**).

There was an interaction between reaction time and luminance in the theta band over lateral prefrontal sites, mainly in cI trials (interaction at electrode F6: $F_{1,14}$ = 6.76, p = 0.021; **Figure 3C**). Consistent with the lack of behavior effect of luminance on reaction time during cC trials, there was no corresponding reaction time × luminance interaction on theta-band activity.

To illustrate the theta-RT relationship at the individual subject level, **Figure 4** shows scatterplots from each individual subject. In these plots, the time–frequency point with the maximum pre-stimulus theta power-RT slope (averaged across all four conditions) was selected.

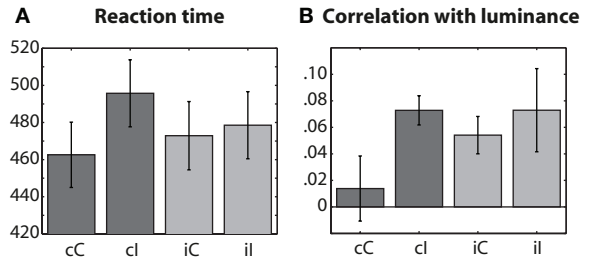


FIGURE 1 | Conditions (x-axis) are labeled according to previous (first, lowercase letter) and current trial (second, uppercase letter) conflict: C = congruent, I = incongruent. (A) Reaction times in ms. **(B)** Spearman's rank correlation coefficients. Error bars denote SEM.

SINGLE-TRIAL PHASE-BEHAVIOR COUPLING

In the next set of analyses, we examined how frequency band-specific phases might be modulated by reaction time, luminance, and their interaction. This is different from cross-trial phase coherence (plotted in **Figures 2B,D**) because this analysis tests the reliability of the relationship between phase angles and reaction time across trials, rather than the consistency of phase angles across trials. As seen in **Figure 5**, there was a robust main effect of reaction time in the pre-response theta range in all conditions. Similar to the effects observed in the power regressions, there was an interaction between current and previous conflict (at FCz: $F_{1,14}$ = 6.56, p = 0.023) which was driven by significantly larger coefficients in cI compared to cC trials (t_{14} = 5.36, p < 0.001), and no significant difference between iI and iC trials (t_{14} = 0.66, p = 0.516). There were no significant effects of luminance or reaction time–luminance interaction.

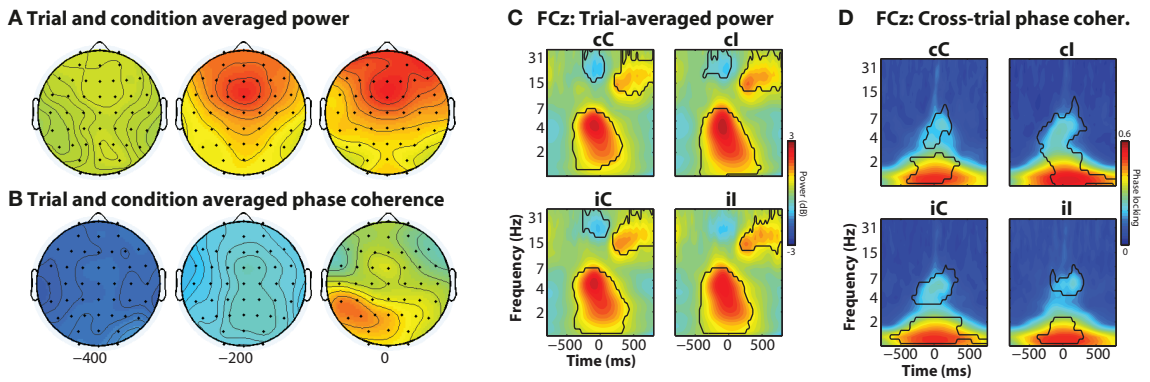


FIGURE 2 | Trial-averaged results, separated according to condition. (A,B) Condition-averaged topographical maps of power and phase coherence, which are the average of 150 ms surrounding the time point indicated below the maps. **(C,D)** Time–frequency plots of peri-response power and inter-trial phase coherence. Black areas enclose regions in which contiguous pixels were significantly different from inter-trial-interval baseline at p < 0.01 (two-tailed) for at least 300 ms and at least three consecutive frequencies.

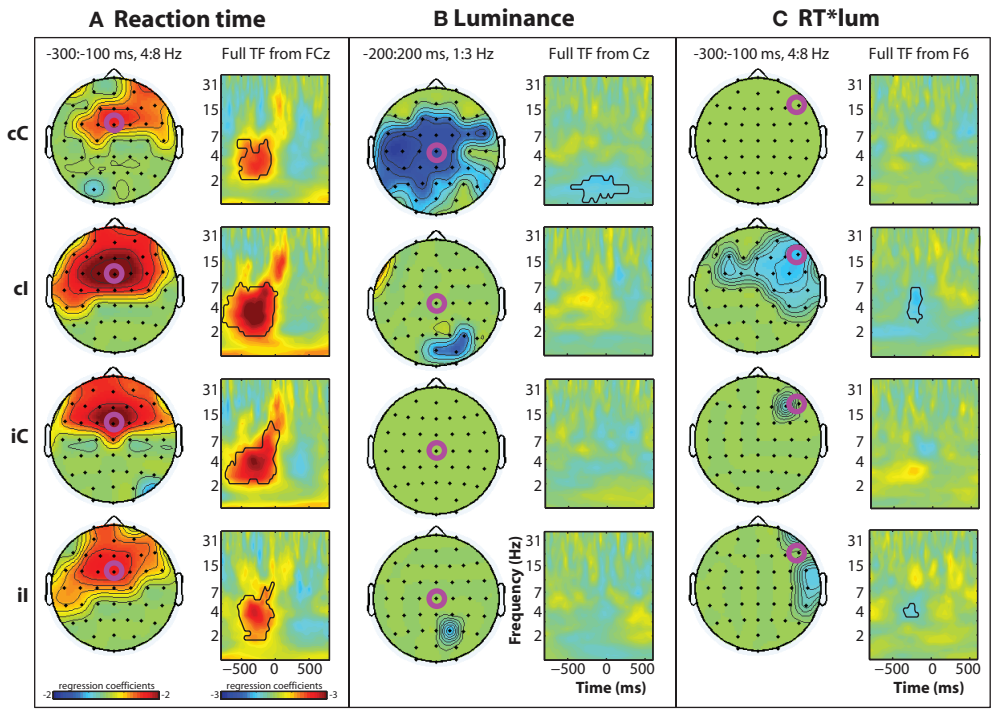


FIGURE 3 | Time–frequency plots of peri-response (time = 0) single-trial multiple regression coefficients, separated by condition (rows) and regression term (A–C). Black lines enclose significant regions, as in **Figure 2**. Time–frequency (TF) regression coefficient plots are taken from electrodes indicated by fuchsia circles.

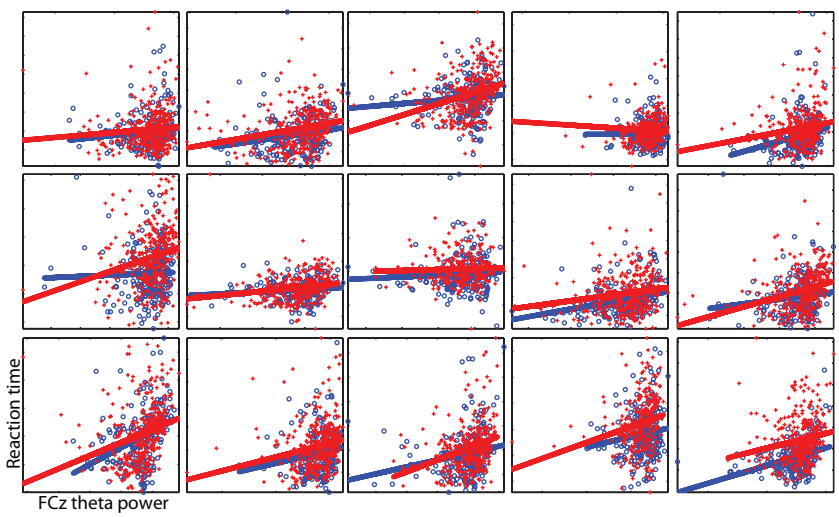


FIGURE 4 | Individual subject single-trial data illustrating the relationship between theta power (*x*-axis) and RT (*y*-axis). Axes differ per subject and are unlabeled. Each point is a single trial; lines reflect the best linear fit. Red pluses and lines are taken from cI trials, blue circles and lines from cC trials. These data correspond to the regression from a single time–frequency point for each subject. To select this subject-specific time–frequency point, we averaged the RT regression coefficients from all conditions and selected the point with the largest coefficient within the range of 1–12 Hz, –400 to 200 ms (the average across subjects was –233 ms and 5.1 Hz). Note that this selection procedure is not based on maximizing differences between cI and cC trials.

SINGLE-TRIAL PHASE SYNCHRONIZATION (FUNCTIONAL CONNECTIVITY)-BEHAVIOR COUPLING

We (Cavanagh et al., 2009) and others (Hanslmayr et al., 2008) have suggested that electrophysiological oscillatory synchronization between MFC and lateral prefrontal cortex may be a substrate of communication that supports cognitive control processes. Although inter-site synchronization increases most strongly during errors, synchronization is also observed immediately prior to

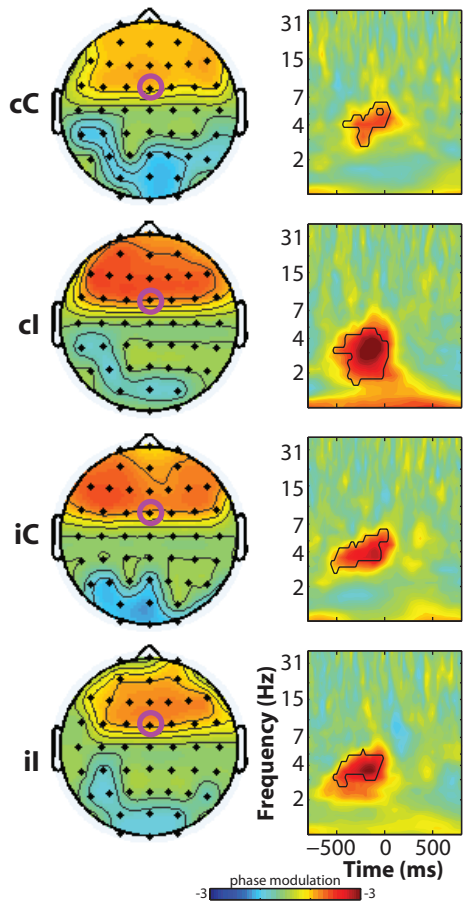


FIGURE 5 | Time–frequency plots of peri-response single-trial phase modulation, separated by condition. Black lines enclose significant regions, as in **Figure 2**. Note the differences between these phase modulation plots – where the dominant effects occur in the theta band over frontal regions – and the trial-averaged phase coherence in **Figure 2**, which shows a dominant delta-band effect without a medial frontal focus.

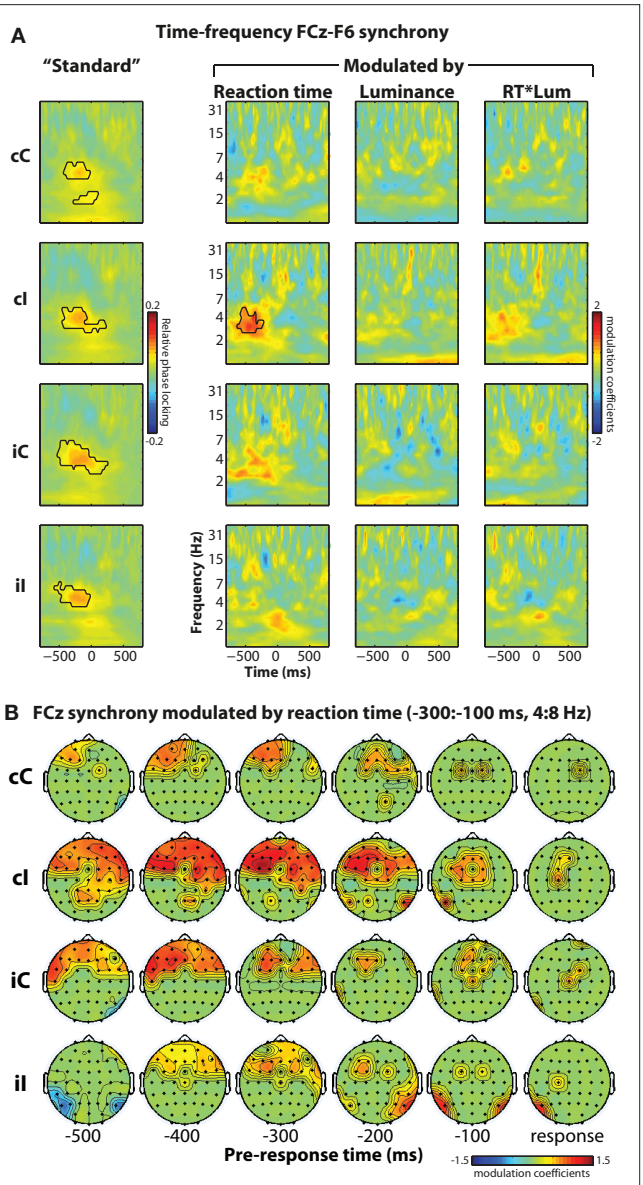


FIGURE 6 | Phase synchronization between electrode FCz (over MFC) and lateral prefrontal sites including F6 is present generally prior to the response, and, during high conflict situations, increases in strength with increasing reaction time. (A) Time–frequency plots of FCz–F6 "standard" (no modulation) phase synchronization (left column) and synchronization modulated by reaction time, luminance, and their interaction (right three columns). **(B)** Topographical maps of reaction time-modulated synchronization with FCz over time (columns) and condition (rows). Statistical thresholding is the same as in previous figures.

correct responses. Consistent with previous results, here we found increased theta-band synchronization between FCz and lateral frontal sites including F6 (this is the electrode pair analyzed in Cavanagh et al., 2009, 2010) prior to the response. The left-most column of **Figure 6A** shows that pre-response synchronization was significantly greater than that during the inter-trial interval (used here as a baseline), but was not different among conditions (all p's > 0.175).

This standard measure of functional connectivity, however, may be difficult to link to precise cognitive dynamics, because the connectivity may reflect a combination of several processes including response preparation, stimulus evaluation, attention, orientation, etc. Thus, we extend this connectivity measure to examine whether synchronization is modulated by an experimental variable (e.g., reaction time). The idea is that process-specific connectivity should be modulated by reaction time, whereas more general inducers of connectivity (e.g., orienting attention, general response preparation) should not be. Similar to the phase analysis above, this method tests the relationship between relative phase angles between seed (FCz) and target (F6 or other) electrode pairs and reaction time. The right three columns in **Figure 6A** show time–frequency plots of modulated phase synchronization between FCz (MFC) and F6 (lateral prefrontal cortex). This "response-modulated" functional connectivity increased prior to the response, and was strongest during cI trials (interaction term: $F_{1,14} = 5.48$, $p = 0.035$). This indicates that on a trial-by-trial basis, longer reaction times during high conflict situations were preceded by enhanced electrophysiological

connectivity between medial frontal and lateral prefrontal regions. Topographical maps in **Figure 6B** show that this effect was localized primarily to anterior and lateral prefrontal sites.

SINGLE-TRIAL ANALYSES BASED ON INDEPENDENT COMPONENTS ANALYSIS

Previous reports suggest advantages to conducting single-trial analyses using data from independent components (Debener et al., 2005, 2007; Eichele et al., 2009). In some cases, independent components and other methods (e.g., stable topographical maps based on clustering) provide converging results (De Lucia et al., 2010). Because independent components analysis results in a set of electrode weights that maximize temporally independent processes, this analysis might improve signal-to-noise. However, because each component is a weighting of all electrodes, specific maps must be selected for each subject. For example, in cognitive control experiments, maps are often selected based on a medial frontal topographical distribution (Debener et al., 2005; Eichele et al., 2010; Wessel and Ullsperger, 2010), although more sophisticated approaches are available based on higher-dimensional clustering (Onton et al., 2005). Therefore, in the interest of comparison with other approaches for analyzing single-trial data, we performed robust regression on the time–frequency representation of independent components. As described in the methods, components for each subject were selected according to a medial frontal and left/right lateral frontal topography (**Figure 7**).

The trial-averaged time–frequency representation of the components time courses are shown in **Figure 8**. The medial frontal component showed an increase in theta power that peaked just prior to the response, similar to results from electrode FCz (**Figure 2**).

Single-trial power regression analyses revealed similar effects as with data from FCz for reaction time (**Figure 9**), though generally less robust (compare to **Figure 3A**). A more striking difference was the lack of luminance × reaction time interaction at the right frontal component.

DISCUSSION

THE MEDIAL FRONTAL CORTEX, THETA OSCILLATIONS, AND COGNITIVE CONTROL

The trial-average results presented in **Figure 2** add to a growing body of work linking theta-band oscillatory dynamics recorded over medial frontal sites to action monitoring, cognitive control, and reinforcement learning (Trujillo and Allen, 2007; Hanslmayr et al., 2008; Marco-Pallares et al., 2008; Cavanagh et al., 2009; Christie and Tata, 2009; Mazaheri et al., 2009). Together, these findings suggest that medial frontal theta is a candidate mechanism for information processing and transfer during conflict, error, and negative performance feedback. Spatial filtering methods such as current source density, independent components analysis, and dipole modeling suggest that these theta dynamics originate in the pre-supplementary motor area or anterior cingulate cortex (Miltner et al., 2003; Debener et al., 2005; Vocat et al., 2008). This is confirmed by direct recordings in humans (Wang et al., 2005; Cohen et al., 2008) and functional MRI studies (van Veen et al., 2001; Mathalon et al., 2003).

Although results of these cross-trial averaging analyses generally link medial frontal theta to response conflict, the single-trial regression analyses provide more behaviorally relevant insights into the theta dynamics that may support cognitive control. Indeed, several

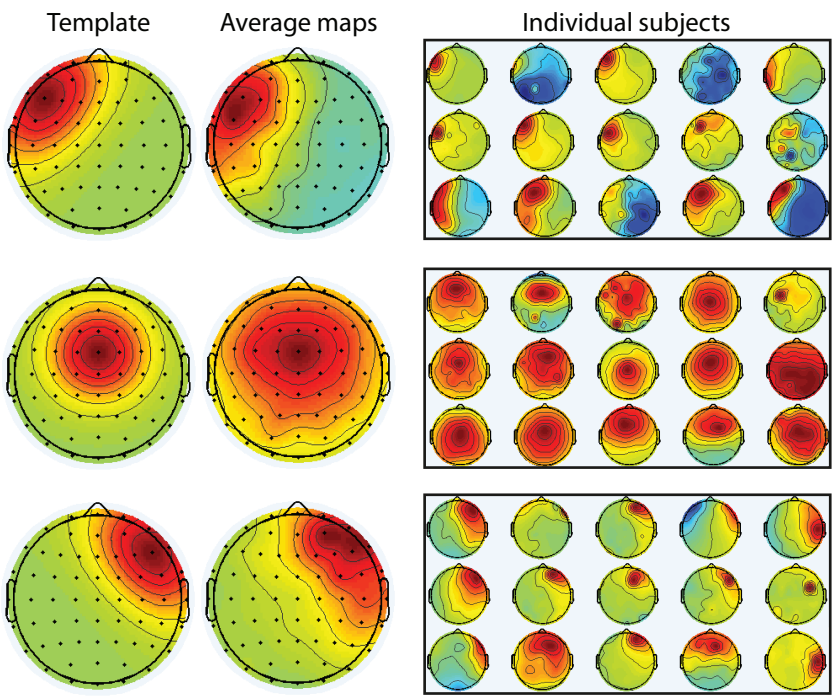


FIGURE 7 | Selection of independent components for all subjects. Components were selected based on spatial correlation with a priori specified templates (left column). The average component across subjects was similar to the templates (middle column). Individual maps from all 15 subjects are shown in the right-most column.

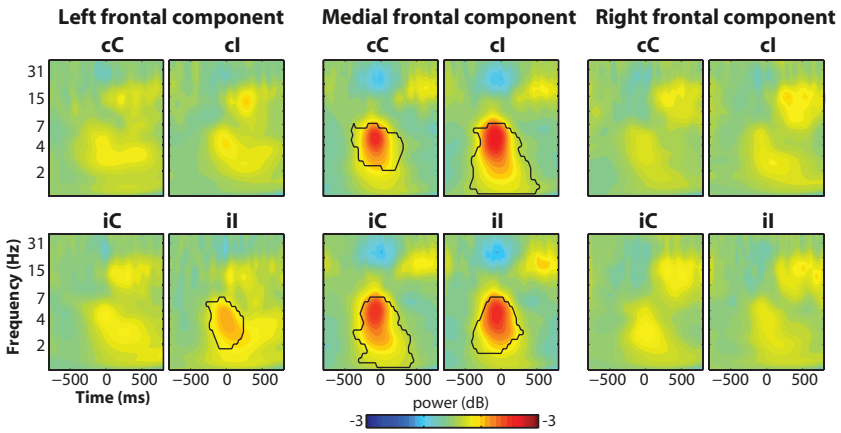


FIGURE 8 | Trial-averaged power results from the independent components (same as Figure 2C but using components instead of electrode-specific data).

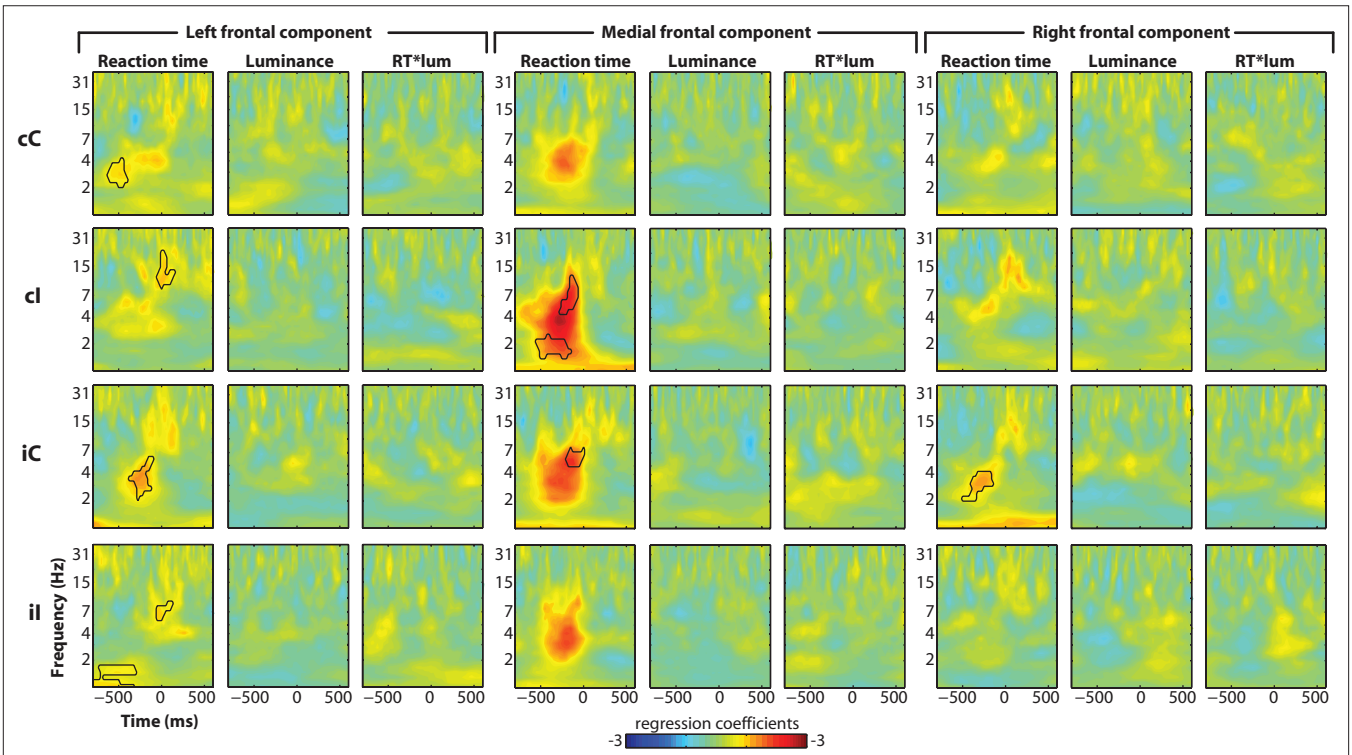


FIGURE 9 | Single-trial multiple regression analyses from independent components. The analysis was identical to that presented in **Figure 3** but using components time courses instead of electrode-specific data.

findings emerged only in the single-trial analyses (discussed in more detail below): (1) single-trial regression demonstrated that pre-response theta is involved in conflict-modulated response selection over a longer time period than is apparent in trial-averaged theta; (2) single-trial "weighted" phase modulation demonstrated that pre-response theta phase predicted endogenous conflict as reflected by reaction time, whereas trial-averaged phase coherence showed only a low-frequency general phase alignment; (3) single-trial phase synchrony modulation demonstrated that medial–lateral prefrontal phase synchronization was significantly modulated by upcoming reaction time in situations of high response conflict, whereas trial-averaged phase synchronization suggested that connectivity was unrelated to conflict.

SINGLE-TRIAL MULTIPLE REGRESSION OVER TIME AND FREQUENCY

The dissociation in timing between trial-averaged theta – which had a narrow peak 100 ms (power) and 122 ms (cross-trial phase coherence) prior to the response – and the single-trial theta response time regression coefficients – which started earlier and peaked 176 ms before the response (see **Figure 9**), suggests two

distinct but nearly temporally overlapping roles of medial frontal theta dynamics. On the one hand, peri-response theta is involved in general response-locked dynamics; on the other hand, it is the pre-response theta that is tightly coupled to response time. Indeed, medial frontal activity correlating with trial-to-trial RT has been reported previously with EEG independent components analysis (Delorme et al., 2007), and with a re-analysis of five fMRI datasets (Yarkoni et al., 2009). Further, the dissociation between the extended pre-response time course of the theta power regression and the theta phase modulation analysis (compare red and black lines in **Figure 10**), in combination with the weak pre-response phase coherence (**Figure 2D**), suggests that reaction time dynamics are more closely associated with non-phase-locked oscillatory activity compared to phase-locked transients. In other words, the neural processes within the medial frontal cortex that influence reaction time during conflict might reflect amplitude modulations of ongoing oscillations, rather than a sudden resetting of activity.

The interaction of luminance and reaction time predicting theta dynamics over lateral prefrontal cortex implicates this region in mediating stimulus-induced conflict. This is consistent with previous findings linking top–down control over visual information to lateral prefrontal functioning (Zanto et al., 2010), and in perceptual conflict (van Veen et al., 2001). Further, right lateral prefrontal cortex has been suggested to play a particularly prominent role in top–down control (Aron et al., 2004). More generally, this highlights two strengths of the single-trial multiple regression approach: (1) Trial-varying stimulus luminance drawn from a random distribution would normally be considered an experimental confound; here, this confound becomes an asset that reveals the involvement of the lateral prefrontal cortex in regulating decision time according to stimulus difficulty. (2) Theta activity is not often localized over lateral prefrontal sites in trial averages (**Figure 2A**, and also Cavanagh et al., 2009; Cohen, 2011a; Nigbur et al., under review), but shows robust trial-by-trial modulations with experiment dynamics. This finding, together with increased conflict-related synchronization with medial frontal sites, demonstrates that lateral prefrontal theta is indeed involved in cognitive control processes, but this is difficult to infer because trial-averaged theta may not increase significantly compared to pre-stimulus baseline activity.

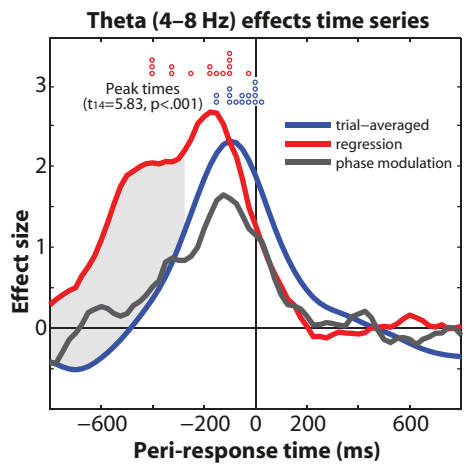


FIGURE 10 | Comparison of theta-band effects (trial-averaged power, power regression with reaction time, phase modulation by reaction time). Gray shading indicates that the regression coefficients are significantly greater than trial-averaged power at $p < 0.01$, minimum 300 ms continuous cluster threshold. Red and blue circles indicate subject-specific waveform peak times.

It is not clear why stimulus luminance had a main effect on low-frequency oscillations only during cC trials. We speculate that because these were the easiest trials, subjects may have had more cognitive resources to devote to low-level stimulus properties. However, this was not explicitly tested, nor did subjects spontaneously mention this.

SINGLE-TRIAL PHASE AND PHASE SYNCHRONY MODULATION

"Standard" inter-trial phase coherence (also called cross-trial phase-locking) assumes that oscillation phase is relevant when the oscillation has a similar phase value across trials at each time–frequency point. Therefore, this approach mixes a number of potential causes of phase coherence, including stimulus-evoked responses, general orienting or attention responses, and task-specific dynamics. This approach precludes discovery of phase dynamics that are related to the task but are not consistent across trials. In contrast, the single-trial "weighted" phase modulation analysis performed here does not require phase values to be similar across trials; rather, this analysis is sensitive to modulations of phase values even if those phases are randomly distributed across trials. Indeed, an absence of pre-response cross-trial phase coherence would be expected if pre-response theta phase were modulated by reaction time (which differs from trial to trial).

Similarly, phase synchronization modulation has advantages over standard phase synchronization because, as described above, inter-site phase synchronization may result from a combination of specific task-related parameters and also more general cognitive/orienting processes. In contrast, the specific modulation of phase synchronization by reaction time provides a more focal interpretation of synchronization *vis-à-vis* conflict dynamics: phase angle differences do not need to be consistent over trials at each time–frequency point; rather, they need only to be consistently related to behavioral or experimental variables.

These two analytic approaches are complementary. Inter-trial phase coherence provides insights into the overall stimulus- or response-related phase consistencies, whereas phase modulation is process-specific. Note, however, that this specificity should be taken into consideration when interpreting results. For example, the phase modulation analyses in **Figure 6** do not indicate that delta-band phase is irrelevant for the task; rather, they show that only theta-band phase is modulated by reaction time, whereas delta-band phase coherence may support a more general cognitive function that is time-locked to the response but unrelated to variations in response time. Similarly, comparing the inter-electrode phase synchronization with the phase synchronization modulation (**Figure 6**) suggests that the MFC-lateral prefrontal theta-band synchronization reflects both general response initiation processes, and, particularly during conflict, reaction time-specific processes.

Based on these findings, it seems that (1) medial frontal cortex has both a general role in generating responses, as well as a specific role in conflict-modulated decision time; (2) lateral prefrontal

cortex is recruited by the medial frontal cortex during situations of conflict, and synchronized theta-band activity may be the substrate of their communication; (3) lateral prefrontal cortex is additionally involved in influencing reaction time according to modulations of conflict (stimulus luminance).

Results from the independent components analyses were generally less robust compared to those based on electrode time courses. It is possible that more sophisticated component clustering techniques would reveal the findings to a similar magnitude as with the electrode-based analyses. However, in this case, it seems that independent components analysis may not necessarily be an optimal approach for single-trial analyses.

Another significant advantage of the methods used here is that they are robust to potential outliers at the single-trial level. During standard trial averaging, trials are not typically inspected for outliers, and it is thus possible that a minority of trials with large oscillation power values bias the average activity levels. In contrast, robust regression minimizes the contribution of outliers, and the phase modulation analyses are based on permutation testing, therefore minimizing the danger of outliers biasing estimates of trial-averaged results.

POSSIBLE EXTENSIONS TO THE SINGLE-TRIAL MULTIPLE REGRESSION APPROACH

One could extend this framework to apply hierarchical regression models in which the variance from the single-subject trial-level data is used to inform group-level results. This might be particularly useful when comparing groups, e.g., if patients and control subjects have similar average effects but patients have more variable responses. This approach could also be applied to fMRI data, for example using as the dependent measure the observed data from ~6 s post-stimulus (when the hemodynamic response is expected to peak) or a beta parameter corresponding to the fit of the post-stimulus data to a canonical hemodynamic response. Although single-trial correlations have been performed with the hemodynamic response (e.g., Rissman et al., 2004; Weissman et al., 2006), multiple regression may prove more powerful for reasons highlighted in the Section "Introduction." Finally, the regressions could be turned around, such that behavior or experimental condition is predicted from brain activity, rather than predicting brain activity from behavior and experimental condition. In this case, logistic regression should be used if predicting binary outcomes (e.g., accuracy or condition; Dixon, 2008; Jaeger, 2008).

CONCLUSIONS

Cross-trial averaging has advantages, but in many cases should be only the first step of a data analysis protocol. Many experiments contain rich and untapped experimental, behavioral, and neural dynamics that remain to be explored and tested. In this study, for example, single-trial multiple regression (and corresponding phase modulation analyses) revealed complementary and interactive roles of the medial and lateral prefrontal cortex during response conflict. The methods used here, of course, are general and could be applied to any design or topic within human neuroscience.

ACKNOWLEDGMENTS

We thank Claudia Arena for assistance with data collection. Michael X Cohen is, and the present experiment was, funded by a VIDI grant from the Dutch Organization for Scientific Research (NWO).

REFERENCES

Aron, A. R., Robbins, T. W., and Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends Cogn. Sci.* 8, 170–177.

Boehler, C. N., Bunzeck, N., Krebs, R. M., Noesselt, T., Schoenfeld, M. A., Heinze, H. J., Munte, T. F., Woldorff, M. G., and Hopf, J. M. (2010). Substantia nigra activity level predicts trial-to-trial adjustments in cognitive control. *J. Cogn. Neurosci.* 23, 362–373.

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

Cavanagh, J. F., Cohen, M. X., and Allen, J. J. (2009). Prelude to and resolution of an error: EEG phase synchrony reveals cognitive control dynamics during action monitoring. *J. Neurosci.* 29, 98–105.

Cavanagh, J. F., Frank, M. J., Klein, T. J., and Allen, J. J. B. (2010). Frontal theta links prediction errors to behavioral adaptation in reinforcement learning. *Neuroimage* 49, 3198–3209.

Christie, G. J., and Tata, M. S. (2009). Right frontal cortex generates reward-related theta-band oscillatory activity. *Neuroimage* 48, 415–422.

Cohen, M. X. (2011a). Error-related medial frontal theta activity predicts cingulate-related structural connectivity. *Neuroimage* 2010, 30.

Cohen, M. X. (2011b). It's about time. *Front. Hum. Neurosci.* 5:2. doi: 10.3389/fnhum.2011.00002

Cohen, M. X., Ridderinkhof, K. R., Haupt, S., Elger, C. E., and Fell, J. (2008). Medial frontal cortex and response conflict: evidence from human intracranial EEG and medial frontal cortex lesion. *Brain Res.* 1238, 127–142.

Cohen, M. X., van Gaal, S., Ridderinkhof, K. R., and Lamme, V. A. (2009). Unconscious errors enhance prefrontal-occipital oscillatory synchrony. *Front. Hum. Neurosci.* 3:54. doi: 10.3389/neuro.09.054.2009

De Lucia, M., Michel, C. M., and Murray, M. M. (2010). Comparing ICA-based and single-trial topographic ERP analyses. *Brain Topogr.* 23, 119–127.

Debener, S., Ullsperger, M., Siegel, M., and Engel, A. K. (2007). Towards single-trial analysis in cognitive brain research. *Trends Cogn. Sci.* 11, 502–503.

Debener, S., Ullsperger, M., Siegel, M., Fiehler, K., von Cramon, D. Y., and Engel, A. K. (2005). Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. *J. Neurosci.* 25, 11730–11737.

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

Delorme, A., Westerfield, M., and Makeig, S. (2007). Medial prefrontal theta bursts precede rapid motor responses during visual selective attention. *J. Neurosci.* 27, 11949–11959.

Dixon, P. (2008). Models of accuracy in repeated-measures designs. *J. Mem. Lang.* 59, 447–456.

Egner, T. (2007). Congruency sequence effects and cognitive control. *Cogn. Affect. Behav. Neurosci.* 7, 380–390.

Eichele, H., Juvodden, H. T., Ullsperger, M., and Eichele, T. (2010). Mal-adaptation of event-related EEG responses preceding performance errors. *Front. Hum. Neurosci.* 4:65. doi: 10.3389/fnhum.2010.00065

Eichele, T., Calhoun, V. D., and Debener, S. (2009). Mining EEG-fMRI using independent component analysis. *Int. J. Psychophysiol.* 73, 53–61.

Forstmann, B. U., Jahfari, S., Scholte, H. S., Wolfensteller, U., van den Wildenberg, W. P., and Ridderinkhof, K. R. (2008). Function and structure of the right inferior frontal cortex predict individual differences in response inhibition: a model-based approach. *J. Neurosci.* 28, 9790–9796.

Gratton, G., Coles, M. G., and Donchin, E. (1992). Optimizing the use of information: strategic control of activation of responses. *J. Exp. Psychol. Gen.* 121, 480–506.

Hanslmayr, S., Pastotter, B., Bauml, K. H., Gruber, S., Wimber, M., and Klimesch, W. (2008). The electrophysiological dynamics of interference during the Stroop task. *J. Cogn. Neurosci.* 20, 215–225.

Jaeger, T. F. (2008). Categorical Data Analysis: away from ANOVAs (transformation or not) and towards Logit Mixed Models. *J. Mem. Lang.* 59, 434–446.

Lachaux, J. P., Rodriguez, E., Martinerie, J., and Varela, F. J. (1999). Measuring phase synchrony in brain signals. *Hum. Brain Mapp.* 8, 194–208.

Luu, P., and Tucker, D. M. (2001). Regulating action: alternating activation of midline frontal and motor cortical networks. *Clin. Neurophysiol.* 112, 1295–1306.

Marco-Pallares, J., Cucurell, D., Cunillera, T., Garcia, R., Andres-Pueyo, A., Münte, T. F., and Rodriguez-Fornells, A. (2008). Human oscillatory activity associated to reward processing in a gambling task. *Neuropsychologia* 46, 241–248.

Mars, R. B., Debener, S., Gladwin, T. E., Harrison, L. M., Haggard, P., Rothwell, J. C., and Bestmann, S. (2008). Trial-by-trial fluctuations in the event-related electroencephalogram reflect dynamic changes in the degree of surprise. *J. Neurosci.* 28, 12539–12545.

Mathalon, D. H., Whitfield, S. L., and Ford, J. M. (2003). Anatomy of an error: ERP and fMRI. *Biol. Psychol.* 64, 119–141.

Mazaheri, A., Nieuwenhuis, I. L., van Dijk, H., and Jensen, O. (2009). Prestimulus alpha and mu activity predicts failure to inhibit motor responses. *Hum. Brain Mapp.* 30, 1791–1800.

Miltner, W. H., Lemke, U., Weiss, T., Holroyd, C., Scheffers, M. K., and Coles, M. G. (2003). Implementation of error-processing in the human anterior cingulate cortex: a source analysis of the magnetic equivalent of the error-related negativity. *Biol. Psychol.* 64, 157–166.

O'Leary, D. P. (1990). Robust regression computation using iteratively reweighted least squares. *J. Matrix Anal. Appl.* 11, 466–480.

Onton, J., Delorme, A., and Makeig, S. (2005). Frontal midline EEG dynamics during working memory. *Neuroimage* 27, 341–356.

Philiastides, M. G., and Sajda, P. (2007). EEG-informed fMRI reveals spatiotemporal characteristics of perceptual decision making. *J. Neurosci.* 27, 13082–13091.

Ratcliff, R., Philiastides, M. G., and Sajda, P. (2009). Quality of evidence for perceptual decision making is indexed by trial-to-trial variability of the EEG. *Proc. Natl. Acad. Sci. U.S.A.* 106, 6539–6544.

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., and Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science* 306, 443–447.

Rissman, J., Gazzaley, A., and D'Esposito, M. (2004). Measuring functional connectivity during distinct stages of a cognitive task. *Neuroimage* 23, 752–763.

Rousselet, G. A., Husk, J. S., Pernet, C. R., Gaspar, C. M., Bennett, P. J., and Sekuler, A. B. (2009). Age-related delay in information accrual for faces: evidence from a parametric, single-trial EEG approach. *BMC Neurosci.* 10, 114. doi: 10.1186/1471-2202-10-114

Rousselet, G. A., Pernet, C. R., Bennett, P. J., and Sekuler, A. B. (2008). Parametric study of EEG sensitivity to phase noise during face processing. *BMC Neurosci.* 9, 98. doi: 10.1186/1471-2202-9-98

Scholte, H. S., Ghebreab, S., Waldorp, L., Smeulders, A. W., and Lamme, V. A. (2009). Brain responses strongly correlate with Weibull image statistics when processing natural images. *J. Vis.* 9, 1–15.

Schwind, J., and Dormann, W. U. (1986). Off-line removal of ocular artifacts from event-related potentials using a multiple linear regression model. *Int. J. Psychophysiol.* 4, 203–208.

Szmalec, A., Verbruggen, F., Vandierendonck, A., De Baene, W., Verguts, T., and Notebaert, W. (2008). Stimulus ambiguity elicits response conflict. *Neurosci. Lett.* 435, 158–162.

Trujillo, L. T., and Allen, J. J. (2007). Theta EEG dynamics of the error-related negativity. *Clin. Neurophysiol.* 118, 645–668.

van Veen, V., Cohen, J. D., Botvinick, M. M., Stenger, V. A., and Carter, C. S. (2001). Anterior cingulate cortex, conflict monitoring, and levels of processing. *Neuroimage* 14, 1302–1308.

Vocat, R., Pourtois, G., and Vuilleumier, P. (2008). Unavoidable errors: a spatiotemporal analysis of time-course and neural sources of evoked potentials associated with error processing in a speeded task. *Neuropsychologia* 46, 2545–2555.

Wang, C., Ulbert, I., Schomer, D. L., Marinkovic, K., and Halgren, E. (2005). Responses of human anterior cingulate cortex microdomains to error detection, conflict monitoring, stimulus-response mapping, familiarity, and orienting. *J. Neurosci.* 25, 604–613.

Weissman, D. H., Roberts, K. C., Visscher, K. M., and Woldorff, M. G. (2006). The neural bases of momentary lapses in attention. *Nat. Neurosci.* 9, 971–978.

Wessel, J. R., and Ullsperger, M. (2010). Selection of independent components representing event-related brain potentials: a data-driven approach for greater objectivity. *Neuroimage* 54, 2105–2115.

Yarkoni, T., Barch, D. M., Gray, J. R., Conturo, T. E., and Braver, T. S. (2009). BOLD correlates of trial-by-trial reaction time variability in gray and white matter: a multi-study fMRI analysis. *PLoS ONE* 4, e4257. doi: 10.1371/journal.pone.0004257

Yeung, N., Ralph, J., and Nieuwenhuis, S. (2007). Drink alcohol and dim the lights: the impact of cognitive deficits on medial frontal cortex function. *Cogn. Affect. Behav. Neurosci.* 7, 347–355.

Zanto, T. P., Rubens, M. T., Bollinger, J., and Gazzaley, A. (2010). Top-down modulation of visual feature processing: the role of the inferior frontal junction. *Neuroimage* 53, 736–745.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 07 January 2011; paper pending published: 04 February 2011; accepted: 14 February 2011; published online: 28 February 2011.

Citation: Cohen MX and Cavanagh JF (2011) Single-trial regression elucidates the role of prefrontal theta oscillations in response conflict. Front. Psychology ***2****:30. doi: 10.3389/fpsyg.2011.00030*

This article was submitted to Frontiers in Perception Science, a specialty of Frontiers in Psychology.

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 10 May 2011
doi: 10.3389/fpsyg.2011.00082

Trial-by-trial variations in subjective attentional state are reflected in ongoing prestimulus EEG alpha oscillations

James S. P. Macdonald[1]*, Santosh Mathan[2] and Nick Yeung[1]

[1] *Department of Experimental Psychology, University of Oxford, Oxford, UK*
[2] *Human Centered Systems Group, Honeywell Labs, Redmond, WA, USA*

Edited by:
Paul Sajda, Columbia University, USA

Reviewed by:
Tom Eichele, University of Bergen, Norway
Lucas C. Parra, City College of New York, USA

****Correspondence:***
James S. P. Macdonald, Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, UK.
e-mail: james.macdonald@psy.ox.ac.uk

Parieto-occipital electroencephalogram (EEG) alpha power and subjective reports of attentional state are both associated with visual attention and awareness, but little is currently known about the relationship between these two measures. Here, we bring together these two literatures to explore the relationship between alpha activity and participants' introspective judgments of attentional state as each varied from trial-to-trial during performance of a visual detection task. We collected participants' subjective ratings of perceptual decision confidence and attentional state on continuous scales on each trial of a rapid serial visual presentation detection task while recording EEG. We found that confidence and attentional state ratings were largely uncorrelated with each other, but both were strongly associated with task performance and post-stimulus decision-related EEG activity. Crucially, attentional state ratings were also negatively associated with prestimulus EEG alpha power. Attesting to the robustness of this association, we were able to classify attentional state ratings via prestimulus alpha power on a single-trial basis. Moreover, when we repeated these analyses after smoothing the time series of attentional state ratings and alpha power with increasingly large sliding windows, both the correlations and classification performance improved considerably, with the peaks occurring at a sliding window size of approximately 7 min worth of trials. Our results therefore suggest that slow fluctuations in attentional state in the order of minutes are reflected in spontaneous alpha power. Since these subjective attentional state ratings were associated with objective measures of both behavior and neural activity, we suggest that they provide a simple and effective estimate of task engagement that could prove useful in operational settings that require human operators to maintain a sustained focus of visual attention.

Keywords: attention, alpha, mind-wandering, vigilance, detection, confidence, P300

INTRODUCTION

Attention waxes and wanes during the undertaking of a task, as our minds wander and subsequently refocus and as our levels of vigilance vary (e.g., Robertson et al., 1997; Gilden, 2001; Wagenmakers et al., 2004; Monto et al., 2008), particularly during relatively straightforward tasks for which processing becomes automated after initial orientation (Smallwood and Schooler, 2006). We refer to such fluctuations in task engagement as "attentional state." Here, we assess two very different measures of attentional state: participants' own introspective judgments, and spontaneous prestimulus electroencephalogram (EEG) alpha activity. On a trial-by-trial basis, we investigate the relationship between these disparate measures of attentional state, the timescales over which they fluctuate, and their interplay with performance in a rapid serial visual presentation (RSVP) detection task.

Alpha is EEG oscillatory activity between approximately 8 and 12 cycles per second that can occur over the entire scalp but is typically highest in amplitude in parieto-occipital areas. It increases in amplitude when the eyes are closed and is attenuated by visual stimulation (Berger, 1929); it has long been considered to reflect general arousal in that low alpha is associated with a state of alertness and high alpha is associated with relaxation or drowsiness (Pollen and Trachtenberg, 1972; Ray and Cole, 1985).

Over the last decade a wealth of research into alpha activity has shown that it is intimately linked to attention and it is now considered by many to reflect cortical excitability, with low alpha indicating active neuronal processing and high alpha denoting inhibition or disengagement of brain areas uninvolved in task performance (e.g., Klimesch et al., 1998; Worden et al., 2000; Pfurtscheller, 2001; Sauseng et al., 2005b; Kelly et al., 2006; Romei et al., 2008a,b; Snyder and Foxe, 2010; for a review see Klimesch et al., 2007). It has frequently been reported that when participants are spatially cued to anticipate a visual stimulus appearing on one side of space, alpha decreases contralaterally and increases ipsilaterally (Worden et al., 2000; Yamagishi et al., 2003; Sauseng et al., 2005b; Kelly et al., 2006, 2009; Thut et al., 2006; Rihs et al., 2007; Wyart and Tallon-Baudry, 2009). A similar finding has been reported for the allocation of attentional focus in the upper and lower visual fields (Worden et al., 2000; Rihs et al., 2007). Furthermore, some studies have reported this effect in tandem with an association of alpha power and task performance: Faster response latencies (Thut et al., 2006; Kelly et al., 2009) and increased accuracy of detection or discrimination (Kelly et al., 2009; Wyart and Tallon-Baudry, 2009) have been shown to coincide with lower alpha power contralaterally and higher alpha power ipsilaterally.

Recently, the causal nature of alpha oscillations in selective attention has been demonstrated: Romei et al. (2010) have shown that alpha frequency (10 Hz) repetitive transcranial magnetic stimulation (r-TMS) enhances visual detection when applied to ipsilateral visual cortex and impairs visual detection when applied to contralateral visual cortex. Doesburg et al. (2009) have reported that the phase-locking of alpha activity between low-level visual cortex and parietal cortex during the interval between an informative spatial cue and target stimulus onset increases contralaterally and decreases ipsilaterally. In a similar vein, Capotosto et al. (2009) have demonstrated that r-TMS to the right intra-parietal sulcus and right frontal eye fields disrupts attentional modulation of alpha power in visual cortex and impairs identification accuracy and response latency to a target stimulus presented approximately 2 s later. Collectively, these results implicate the parietal cortex in the control of attention by alpha suppression and enhancement in visual cortex.

There is a good deal of evidence, therefore, that relative alpha power within occipital cortex reflects the spatial focus of visual attention. In addition, the overall level of alpha power in visual cortex has been shown to be indicative of disengagement with external visual input and a focus on other senses or internal thoughts. For example, focusing attention on auditory (Foxe et al., 1998; Fu et al., 2001) or somatosensory input (Linkenkaer-Hansen et al., 2004) results in increased alpha power in parieto-occipital cortex. Alpha power in parieto-occipital cortex is also greater during internal cognitive tasks such as mental arithmetic (Ray and Cole, 1985; Palva et al., 2005) and imagery (Ray and Cole, 1985; Hari et al., 1997; Cooper et al., 2006, 2003), and during short-term and working memory retention (Jensen et al., 2002; Busch and Herrmann, 2003; Sauseng et al., 2005a). In a recent study, alpha power was found to be greater while participants were focused on an internal counting task (Braboszcz and Delorme, 2011). Conversely, when attention is redirected to visual input, alpha power is attenuated. For example, alpha power decreases in response to a warning cue that the appearance of a task-related stimulus is imminent (Klimesch et al., 1998), and is reduced following errors relative to correct trials in the Stroop task (Carp and Compton, 2009) and in digit discrimination tasks (Mazaheri et al., 2009). Since many models of cognitive control propose that detection of an error is a sign that task approach needs to be improved (Holroyd and Coles, 2002; Ridderinkhof et al., 2004; Yeung et al., 2004), this finding implies that after a lapse in attention, participants refocus on the task at hand.

These observations of elevated levels of alpha activity in visual cortex in association with disengagement from visual input have recently led researchers to investigate whether visual awareness is negatively correlated with spontaneous alpha power. However, whereas there have been many demonstrations of alpha power changes due to directed attention in a variety of paradigms, as discussed above, effects of alpha power on visual awareness are seemingly more difficult to obtain. Some authors have reported such an association in simple threshold contrast detection tasks with precisely controlled stimulus conditions, with greater prestimulus alpha preceding trials in which a target was not detected than trials in which a target was detected (Ergenoglu et al., 2004; van Dijk et al., 2008; Busch et al., 2009); during a visual stimulus duration monitoring task (O'Connell et al., 2009), alpha power was found to steadily increase starting up to 20 s before a task error occurred. Others, however, with similar paradigms have found no such association (e.g., Thut et al., 2006). Therefore it seems that whereas the link between alpha power in parieto-occipital cortex and directed attention is strong, subsequent effects on visual awareness are less consistently observed.

To summarize, prestimulus alpha power is a good candidate for use as an index of attentional state on a trial-by-trial basis since it has repeatedly been demonstrated to reflect the spatial locus of visual attention and the extent to which attention is focused on visual input. In this study, we sought to investigate the relationship between alpha power and participants' subjective ratings of their attentional state.

There is a growing literature demonstrating that participants' introspective judgments of attentional state are reliable and meaningful. Typically, such judgments are recorded via direct questioning at random intervals. The type of measure used has varied from a simple binary response in which participants indicate whether they were focused on the task or not immediately preceding the question (e.g., Smallwood et al., 2004, 2008; Mason et al., 2007; Christoff et al., 2009; Forster and Lavie, 2009; Kam et al., 2011), to asking participants to comment on what they were thinking about just prior to the question (for a review see Smallwood and Schooler, 2006). It has been proposed that during periods in which participants report that they are not focused on the task, attention has switched to thoughts unrelated to the task, i.e., mind-wandering (Smallwood and Schooler, 2006). This interpretation is supported by studies that asked participants to report what they were thinking about rather than simply report whether they were focused on the task or not (Teasdale et al., 1993, 1995).

During periods of reportedly low attentional state, sensory awareness of the external world may be reduced (Smallwood and Schooler, 2006). For example, in a task in which a response must be withheld upon presentation of a target digit, errors are more likely preceding reports of being unfocused on the task, assessed via probes occurring at random intervals (Smallwood et al., 2004, 2008; McVay and Kane, 2009). Furthermore, in EEG studies, early event-related potential (ERP) components related to perception (the P1 for visual stimuli and the N1 for auditory stimuli) have been shown to be attenuated in trials preceding reports of being unfocused on the task (Kam et al., 2011), as have later decision-related components, such as the P300, in response to targets (Smallwood et al., 2008). Episodes of low attentional state have also been associated with increases in neural activity in the default mode network as measured with fMRI (Mason et al., 2007; Christoff et al., 2009), and such neural changes in this network have in turn been associated with task errors (Eichele et al., 2008; Christoff et al., 2009).

One aspect of attentional state that has not yet been investigated is its temporal properties, i.e., how frequently do fluctuations in attentional state occur? Typically, attentional state probes are interspersed at random intervals of 30–90 s and all trials occurring during the 15 s prior to each mind probe are assigned with the subsequent response (e.g., Smallwood et al., 2008), as a compromise between the temporal resolution of reported attentional state and trial economy. Given that studies of this kind have found an association between subjective reports of attentional state and task performance (Smallwood et al., 2004, 2008; McVay and Kane, 2009), we might assume that fluctuations of attentional state occur with

cycle durations in the order of at least tens of seconds. In support of this notion, a recent study (Monto et al., 2008) has demonstrated that somatosensory detection performance is correlated with the phase of infraslow fluctuations (0.1–0.01 Hz, i.e., 10–100 s cycles) of EEG. Furthermore, power in all frequency bands was similarly correlated, suggesting that the infraslow fluctuations reflect the excitability dynamics of cortical networks. Both of these findings could be interpreted as reflecting slow variations in attentional state.

Here, we attempt to bring together the research into EEG alpha activity and attention, and studies of subjective attentional state, by exploring the robustness of the relationship between spontaneous prestimulus EEG alpha activity and participants' introspective ratings of attentional state on a continuous scale, with a focus on trial-by-trial variability. If these measures both relate to attentional state, they should be negatively correlated. However, it is an open question as to whether they would be associated in this way. On the one hand they operate on disparate levels: Parieto-occipital alpha power has been suggested to reflect the excitability of visual cortex (e.g., Klimesch et al., 2007), implying that it reflects relatively low-level physiological factors, whereas reporting on one's own mind-state is clearly a high-level process. On the other hand, both have been shown to be associated with hallmark attentional effects, as reviewed above. We also anticipated that high attentional state ratings and low prestimulus alpha power would be associated with better detection performance – as the results of previous studies' that we have reviewed above would suggest. To test these predictions, and to explore the timescales over which these measures co-varied during the course of the experiment, we correlated the time series of attentional state ratings, prestimulus alpha power, and task performance after smoothing them with increasingly large sliding windows. Finally, we considered that attentional state ratings might be positively associated with the amplitude of ERPs, for example, early perceptual components (e.g., the N1 and P1) in response to stimulus onset warning cues, steady-state visually evoked potentials (SSVEPs) generated by the RSVP stream, and post-stimulus P300 amplitude in response to detected targets.

We also asked participants to rate the confidence of their perceptual decision. We predicted that, like attentional state ratings, this measure would correlate positively with detection performance and perhaps P300 amplitude, but we did not expect it to be associated with prestimulus alpha power. We also anticipated that confidence ratings could potentially be somewhat positively correlated with attentional state ratings, because a high attentional state could improve perceptual representation of the target leading to higher decision confidence, and conversely, that decision confidence might retroactively bias judgments of attentional state.

This study is part of a wider project looking at the use of EEG to improve image triage efficiency in an operational context (e.g., Gerson et al., 2006; Mathan et al., 2006, 2008; Mathan, 2008; Parra et al., 2008; Poolman et al., 2008). As such, the task we employ here is intended to be an experimental analog of the work performed by intelligence analysts searching satellite imagery. Previous work on this project has shown that efficiency is improved if images are subdivided into smaller images and presented in an RSVP stream (Gerson et al., 2006; Mathan et al., 2006). Hence, our task was a simple detection task with complex stimuli: An RSVP stream of noise images lasting 1 s during which a target geometric pattern was sometimes presented, embedded in the noise of one of the images (see **Figure 1** for a schematic illustration of the trial procedure). The stimuli we used were synthetic to allow precise control over stimulus visibility, but were designed to mimic satellite imagery used by other groups in the project. At the end of each trial participants reported whether they had seen the target or not, and rated both their confidence of this decision and their attentional state with respect to the task during that trial. To minimize the time taken to collect responses, and to provide participants with an intuitive response method, all three of these judgments were reported via a single click of a mouse within a large square (see **Figure 1** for an illustration). The vertical axis of the square reflected their attentional state, and the horizontal axis reflected both their target present or absent response and their confidence of this decision: Clicks within the right half of the square indicated

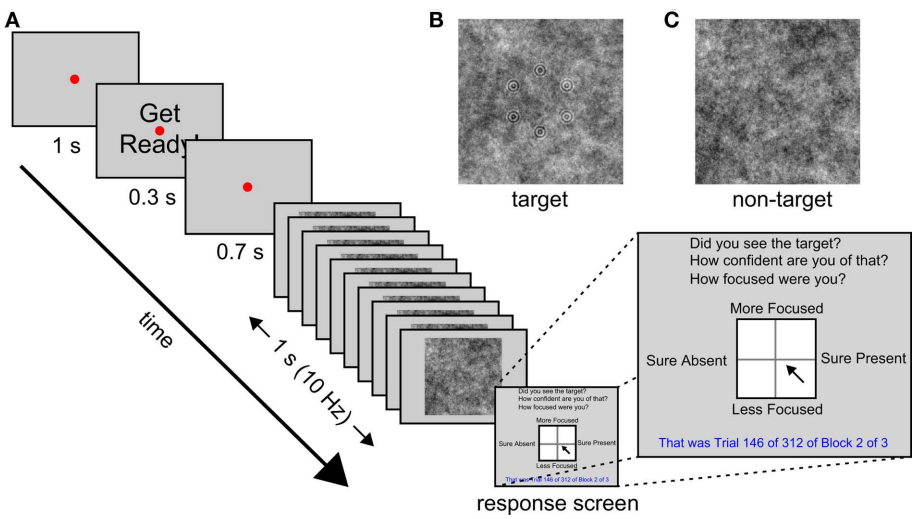


FIGURE 1 | Trial procedure and stimuli. (A) A schematic illustration of the sequence of events in each trial. **(B)** Example target image. **(C)** Example non-target image.

target present and clicks within the left half indicated target absent; distance from the vertical midline to either edge of the square indicated confidence. Participants' EEG in 32 channels was recorded throughout the experiment.

MATERIALS AND METHODS

PARTICIPANTS

Twelve participants were recruited at the University of Oxford and were paid for their participation. One participant was excluded and replaced due to consistent reporting of target present responses as high attentional state and target absent responses as low attentional state with very little other variability. The age range of those included was 18–29 years ($M = 22.33$ years, SD = 4.36 years), and there were five males. All of the participants had normal or corrected-to-normal vision. The research was conducted in accordance with the American Psychological Association's standards for ethical treatment of participants and with the approval of the University of Oxford's institutional review board.

STIMULI AND PROCEDURE

The experiments were created and run with the Psychophysics Toolbox version 3 (Brainard, 1997) in Matlab 2009b (The Mathworks, Inc., 2009) on a Windows PC attached to a 20″ monitor at a resolution of 1024 × 768 and a refresh rate of 60 Hz. The participants' task was to monitor an RSVP stream of images for the presence of a target image. The sequence of events on each trial is illustrated schematically in **Figure 1**.

In each trial a set of 10 images was presented serially over the course of 1 s (i.e., at 10 Hz). Each image was centered at fixation, subtended 18.5° by 18.5° of visual angle at a viewing distance of 57 cm, and was presented for 50 ms, followed by a blank gray screen for another 50 ms. The images were gray-scale white noise patterns that were randomly selected from a pre-generated set of 60 used for all participants. In target images, a set of six concentric circle patterns (each subtending 0.4° radius and consisting of two concentric circles) arranged in a randomly oriented hexagon of 3.3° radius, was embedded in the background noise. Targets were presented in 50% of trials, and their position in the RSVP stream was counterbalanced, although they were never presented in the first or last two positions. The contrast of the target pattern was determined for each participant during a brief pre-experimental session consisting of three blocks of 24 trials, in which a staircase procedure (QUEST from the Psychophysics Toolbox 3, Brainard, 1997) was used to titrate detection rate at 75%. This pre-experimental session also served as a practice session for the participants. Feedback was provided at the end of each trial during the first practice block only. The experimental session consisted of three blocks of 312 trials each.

The RSVP image stream was preceded by a red fixation point of 0.1° radius at the center of the gray screen for a period of 2 s that was interrupted after 1 s by the presentation of the words "Get Ready!" in black, centered at fixation, for 0.3 s. The letters of the words subtended 0.6° by 0.7°. After the offset of these words the screen remained blank, except for the red fixation point, for 700 ms before the RSVP stream began. It was followed by a response screen (see **Figure 1**) consisting of a white square (8.0° by 8.0°) with a black border subdivided into four quadrants by a faint gray line, with the response questions, "Did you see the target?", "How confident are you of that?", and "How focused were you?" in black at the top of the screen, the number of the current trial and block, as well as the number of trials per block and total number of blocks, in blue at the bottom of the screen. The words "Sure Present" appeared on the right hand side of the square, "Sure Absent" on the left, "More Focused" above, and "Less Focused" below, in black. Participants were instructed to click once with a mouse within the square to indicate their response to all three task questions. The vertical axis of the square reflected their attentional state, whereas the horizontal axis reflected both their target present or absent response and their confidence of this decision: Clicks within the right half of the square indicated target present and clicks within the left half indicated target absent; distance from the vertical midline to either edge of the square indicated confidence. The square was 201 by 201 pixels so attentional state ratings were measured on a 201-point scale and confidence was measured on a 101-point scale (since confidence increased from 0 up to 100 for target present responses and decreased from 0 to −100 for target absent responses). Participants were asked to rate their attentional state with respect to the current trial only in terms of the extent to which they were focused on performing the detection task. They were asked to incorporate their levels of vigilance and distraction (whether from internal sources, i.e., mind-wandering, or external sources, e.g., sounds outside the testing room) into this one response.

EEG RECORDING AND PROCESSING

A Neuroscan Synamps2 system (10 GΩ input impedance; 29.8 nV resolution) was used to record EEG data from 32 Ag/AgCl electrodes mounted in an elastic cap at locations FP1, FPZ, FP2, F7, F3, FZ, F4, F8, FT7, FC3, FCZ, FC4, FT8, T7, C3, CZ, C4, T8, TP7, CP3, CPZ, CP4, TP8, P7, P3, PZ, P4, P8, POZ, O1, OZ, and O2. An additional six external electrodes were attached to the left and right mastoids, of which the left acted as a reference, the outer canthi of the left and right eyes, and above and below the right eye to measure electro-oculograms (EOGs). All electrode impedances were kept below 50 kΩ. EEG data were recorded at a sampling rate of 1000 Hz and were high-pass filtered online above 0.1 Hz. Data were downsampled off-line at 100 Hz, then low-pass filtered at 48 Hz, and subsequently epoched from 1.5 s before RSVP onset to 1 s after offset. For the alpha power analyses, the continuous data were additionally high-pass filtered at 0.5 Hz to reduce noise before the epochs were extracted. EEG epochs were baseline-corrected by subtracting the average of the data points between 1.1 and 1 s before RSVP onset. Eye blink correction was conducted using an independent components analysis approach via the EEGLab toolbox for Matlab (Delorme and Makeig, 2004). All EEG processing was conducted with custom-written scripts using native Matlab commands and commands from the EEGLab toolbox. Filtering was conducted with the "eegfilt" command from the EEGLab toolbox, which utilizes a two-way least-squares FIR filter. The order of the filter was equal to the sampling rate (100 Hz) divided by the lower edge of the band, rounded down and multiplied by three, with a minimum value of 15.

ANALYSIS

Our analysis was intended to uncover neural correlates of detection performance and continuous ratings of subjective attentional state and decision confidence. We therefore compared ERPs and

prestimulus alpha power as a function of each type of rating, and correlated increasingly smoothed time series of attentional state ratings, alpha power, and task performance to investigate the timescales over which they co-varied. Finally, we used single-trial classification to appraise the robustness of the association between attentional state ratings and prestimulus alpha power. Details of these analyses are given in the following sections.

To quantify the P300 for each trial, we re-epoched the EEG data time-locked to target image onset, and baseline-corrected by subtracting the average of data points between 0.1 s before and after target image onset. We then low-pass filtered the EEG epochs below 8 Hz to remove the strong 10 Hz SSVEP signal generated by the RSVP stream, averaged them across central, parietal, and occipital midline electrodes CZ, CPZ, PZ, POZ, and OZ, and took the maximum voltage between 350 and 450 ms post-target image onset as the amplitude of the P300 for each trial.

To quantify prestimulus alpha power for each trial, we first band-pass filtered the EEG epochs from parieto-occipital channels P7, P3, PZ, P4, P8, POZ, O1, OZ, and O2, using a 4-Hz band centered on each participant's modal alpha frequency (typically 10 Hz, Klimesch et al., 2007). We subsequently computed the envelope of the amplitude-modulated signal via the Hilbert transform ("hilbert" function in Matlab), which discards phase information and reveals oscillatory power fluctuations over time, and averaged this signal within the 1-s time period between the onset of the words "Get Ready!" and the onset of the RSVP stream. Hence, data from the first 0.5 s and last 2 s of the epochs were discarded, avoiding contamination from edge effects after filtering. We then divided by the average Fourier spectrum power for all frequencies except those in the alpha band, from the same time interval, to normalize alpha power to that of the rest of the EEG frequency spectrum. We performed this step because we were concerned that a simple measure of alpha power might be unduly affected by changes in broadband EEG power (i.e., recording noise) during the session, which might mask the activity of interest. By dividing alpha power by the average broadband power from the same time interval, we eliminate this source of extraneous variance in our alpha quantifications. Without this step the results were slightly less clear in some participants, but the overall pattern was not materially affected. Therefore while the normalization procedure does have an effect, it only served to reduce noise.

The aim of the timescales analyses was to investigate the pairwise associations between prestimulus alpha power, attentional state ratings and task performance (trial accuracy), at different timescales within the experimental session. Our novel approach was to hold the trial-by-trial time series of one variable constant and increasingly smooth the corresponding time series of the other variable, calculating Pearson's Product Moment Correlation Coefficient after each increase in the sliding window size. Prior to smoothing, we high-pass filtered the time series of each variable at two cycles per experimental session and subsequently discarded the first and last 72 trials to avoid contamination from edge effects. The purpose of this pre-processing step was to remove the gradual downward drift that dominated the time series of attentional state ratings of 4 of the 12 participants, while having minimal effects for the other eight participants. Hence, the analysis focused on fluctuations within the experimental session. Our smoothing method was to average over trials within a sliding window. At the beginnings and ends of the time series, the number of available trials was used instead of the specified sliding window size. Since smoothing also resulted in edge effects, we discarded an additional 36 trials at the start and end of each time series, leaving 720 trials. We then repeated the smooth and correlate procedure, leaving the previously smoothed time series unsmoothed and smoothing the other time series.

Due to the novelty of our approach, we repeated the analysis on simulated time series data to verify that it could reveal correlated fluctuations in pairs of time series and would not generate spurious correlations. Thus, the simulated time series could be uncorrelated or could be correlated over fast timescales, slow timescales, or both fast and slow timescales. Of interest was the effectiveness with which the smoothing analysis, applied exactly as it was to our empirical EEG data, would identify the presence or absence of these correlations in the time series data. Specifically, we generated four sets of simulated time series: The first set contained no oscillations; the second consisted of sine waves of 1–9 cycles per experiment (i.e., slow oscillations); the third, sine waves of 46–90 cycles per experiment (i.e., fast oscillations); the fourth, sine waves of 1–9 and 46–90 cycles per experiment (i.e., both slow and fast oscillations). We subsequently inverted the values of each signal to form a complementary signal that was perfectly negatively correlated with the first. We then added independently generated white noise to each simulated time series, and finally performed our smoothed time series correlation analysis on 12 time series pairs (to match the number of participants).

We subsequently used a single-trial classification analysis to appraise the robustness of the association between attentional state ratings and prestimulus alpha power. Specifically, we classified upper vs. lower attentional state rating quartiles on the basis of prestimulus alpha power. We used a logistic regression classifier (Parra et al., 2002) that identifies the spatial distribution of scalp EEG activity in a given time window that maximally distinguishes two conditions to deliver a scalar estimate of component amplitude on each trial. The derived estimates are robust (i.e., have high signal-to-noise) because the discriminating components act as a spatial filter that estimates component amplitude as a spatially weighted average across electrodes for each trial, in much the same way that conventional ERP analysis averages across trials to reduce noise (Parra et al., 2002).

For the classification analyses, alpha amplitude was quantified as Fourier spectral power in a 4-Hz band centered on each participant's modal alpha frequency (typically 10 Hz). We used the 1 s of EEG data between the onset of the words "Get Ready!" and the onset of the of the RSVP stream to compute a Fourier spectrum for each trial. Since we had downsampled the data to 100 Hz, an array of five values – one for each integer frequency within the 4-Hz band – was obtained per electrode, per trial. The classifier identified an optimal weighting of electrodes for each of these five frequencies as a predictor of upper vs. lower attentional state quartile. We averaged the classifier output for the five frequencies to obtain a scalar value for each trial from 0 to 1, which can be conceptualized as the estimated probability that attentional state rating was in the upper quartile on that trial. By comparing these values with the objective truth label of the trial (i.e., 0 = lower attentional state

rating quartile and 1 = upper), we computed the Az score (the area under the receiver operating characteristic curve, Stanislaw and Todorov, 1999) for each participant.

We employed a 2-fold cross validation approach, such that we trained the classifier on half of the data and tested it on the other half, then repeated this procedure after switching training and validation data. We employed a random-fold assignment procedure, in which data samples are randomly assigned to training and validation sets, rather than a sequential-fold assignment procedure, in which data from each class are split into training and validation sets based on their temporal order. Whereas sequential folds provide a better indication of cross-session generalization, the advantage of random folds is that non-stationary aspects of the EEG signal are factored out and the invariant correlates of the conditions being classified become more prominent. Another advantage of a random-fold assignment procedure is that it introduces a source of variability into the computation, such that classifications can be repeated a number of times and Az score confidence intervals can be computed. We repeated all classifications 1000 times and give the mean and 95% confidence interval for individual participants' Az scores.

RESULTS

We first present analyses of the continuous trial-by-trial ratings of attentional state and perceptual decision confidence, and their relation to task performance. We then relate these measures to key EEG indices of attention and task performance: ERPs and ongoing prestimulus alpha power as a function of quantile-split attentional state and confidence ratings. Having observed an overall relationship between these measures, our next analyses focus on the timescales of variation of attentional state ratings, prestimulus EEG alpha power, and task performance. Finally, to assess the robustness of EEG alpha power as a measure of attentional state, we investigate whether this measure can be used to predict participants' subjective attentional state ratings as they vary from trial-to-trial using multivariate classification.

TASK PERFORMANCE AND ATTENTIONAL STATE AND CONFIDENCE RATING DISTRIBUTIONS

Target contrast corresponding to a detection rate of 75% was determined for each participant with a staircase procedure during a pre-experimental session. The mean contrast value was 0.15 (SD = 0.02). The mean detection rate across participants was 71% (SD = 8%) and the mean false alarm rate (target present responses when no target was presented) was 9% (SD = 5%). Mean d′ was 1.98 (SD = 0.49).

Figure 2A shows the locations within the response square of all clicks by all participants for target present trials (in black) and target absent trials (in red), and the grand average distributions of attentional state and confidence ratings for target present and absent trials separately (also in black and red). Since clicks within the right half of the square denoted target present responses and clicks within the left half denoted target absent responses, black dots on the right indicate trials in which a target was presented and detected (hits), whereas those on the left indicate trials in which a target was presented but not detected (misses). Correspondingly,

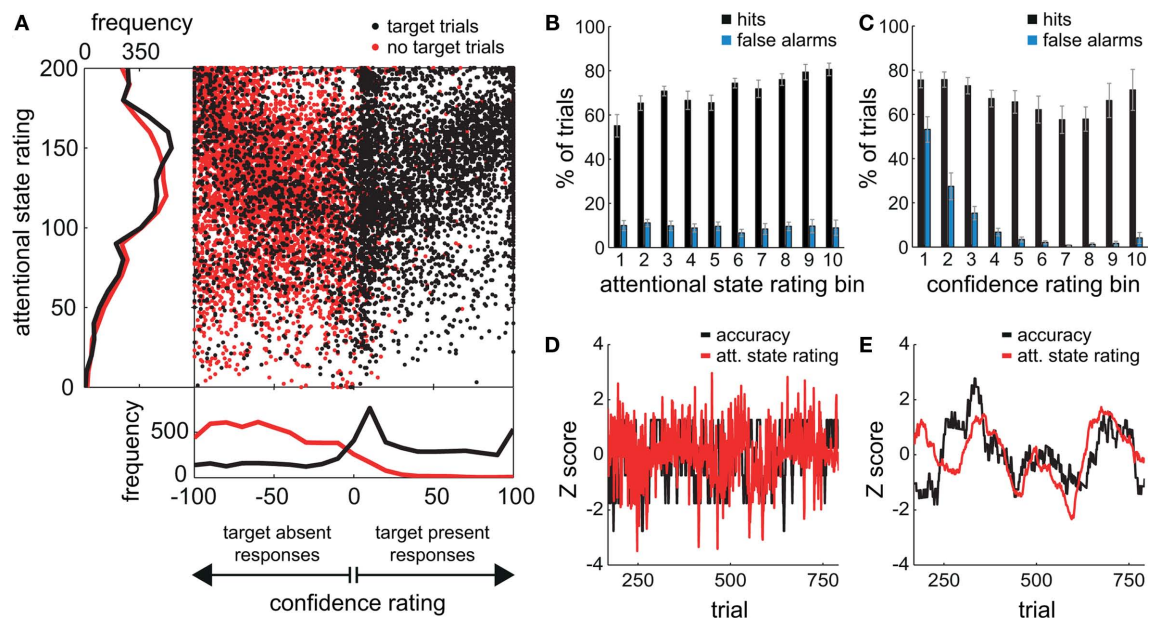


FIGURE 2 | Behavioral results from the visual detection task. (A) The locations of all responses (mouse clicks) from all participants. Responses in target trials are shown as black dots; responses in non-target trials are shown as red dots. The frequency distributions of attentional state ratings for target trials (black line) and non-target trials (red line) are shown on the left hand side, and the frequency distributions of decision confidence ratings are shown below. **(B,C)** Grand average percentages of hits and false alarms as a function of attentional state rating decile and confidence rating decile, respectively. Error bars show the SE of the mean across participants. **(D)** The normalized time series of attentional state ratings (red line) and detection task accuracy (black line) for one participant. **(E)** The normalized time series of attentional state ratings (red line) and detection task accuracy (black line) for one participant after smoothing each time series with a sliding window size of 101 trials.

red dots on the right indicate trials in which no target was presented but one was reported (false alarms), and those on the left indicate trials in which it was correctly reported that no target had been presented (correct rejections). An association between attentional state and confidence rating can be observed for both target present and absent responses: A dense diagonal cloud of dots is evident from the center to the top right corner for target present responses (black dots) and from the center to the top left corner for target absent responses (red dots), indicating that participants were more confident about their judgments of both target presence and target absence on trials in which they felt they had paid greater attention to the task.

The grand average distribution of confidence ratings (lower panel of **Figure 2A**) for target present trials (black line) shows that confidence ratings on hit trials were bimodally distributed with a peak at the maximum confidence rating (100), and another close to the minimum confidence rating (0). Confidence ratings on miss trials were more distributed, with no peak at the maximum (−100), but a small one at the minimum confidence rating (0). Hence, when confidence was low, participants tended to opt for target present rather than absent. In addition, these low confidence target present responses were very frequently correct since there is no corresponding peak in the distribution of confidence ratings for target absent trials (red line). Instead, this distribution is biased toward higher confidence ratings, so there seems to have been a tendency to rate correct rejections as high confidence. False alarms on the other hand were very rarely rated as high confidence and tend to cluster close to 0, suggesting that they were guesses.

The grand average distributions of attentional state ratings (**Figure 2A**, left panel) are negatively skewed for both target present and absent trials (black and red lines, respectively), indicating that participants tended to rate their attentional state as high rather than low, with modes of 150 (out of 200) for target present trials, and 118 (out of 200) for target absent trials. The mean attentional state rating for target present trials was 130.71 (SD = 20.46) and for target absent trials was 126.36 (SD = 19.95), a small but significant reduction, $F(1, 11) = 17.01$, MSE = 113.75, $p = 0.002$, $\eta_p^2 = 0.61$. This result reveals a slight but consistent bias in attentional state ratings, since whether a target was presented in a trial or not was randomly determined and could not be related to fluctuations in attentional state. We interpret this effect as a small retroactive bias of attentional state ratings in hit trials: Noticing a target may lead to increased alertness and hence a higher attentional state rating, or might lead participants to assume their attention must have been high given that they detected a target.

TASK PERFORMANCE AND ATTENTIONAL STATE RATINGS

To investigate the relationship between attentional state ratings and task performance, we binned trials by attentional state rating and subsequently performed one-way ANOVAs on hit rate, false alarm rate and d′ (Stanislaw and Todorov, 1999) by bin. Higher attentional state ratings were associated with better performance in terms of hit rate, $F(9, 99) = 8.65$, MSE = 709.47, $p < 0.001$, $\eta_p^2 = 0.44$, and d′, $F(9, 99) = 6.84$, MSE = 1.19, $p < 0.001$, $\eta_p^2 = 0.38$, but not false alarm rate, $F < 1$. Grand average hit rate and false alarm rate as a function of attentional state rating bin are shown in **Figure 2B**. A clear linear trend is evident with hit rate increasing as attentional state rating increases, $F(1, 99) = 63.43$, MSE = 5203.05, $p < 0.001$, $\eta_p^2 = 0.39$. There was also a significant linear trend for d′, $F(1, 99) = 53.42$, MSE = 9.29, $p < 0.001$, $\eta_p^2 = 0.35$.

As noted above, we observed a retroactive bias to rate attentional state as higher in target present trials. To rule out the possibility that this bias produced an artifactual association between attentional state ratings and task performance, we repeated the analysis using the rating from the following trial, thus eliminating any bias arising from the presence or absence of a target on the current trial. The association persisted in hit rates, $F(9, 99) = 2.21$, MSE = 152.65, $p = 0.027$, $\eta_p^2 = 0.17$, as did the linear trend, $F(1, 99) = 9.53$, MSE = 657.92, $p = 0.003$, $\eta_p^2 = 0.09$. Additionally, we repeated the analysis after excluding any participant for whom the distributions of attentional state ratings for target present and absent trials did not overlap ($N = 7$). Again, the association remained for hit rates, $F(9, 36) = 4.74$, MSE = 471.80, $p < 0.001$, $\eta_p^2 = 0.54$, as did the linear trend, $F(1, 36) = 31.52$, MSE = 3138.03, $p < 0.001$, $\eta_p^2 = 0.47$. Effect sizes for both the ANOVA and the test of trend increased rather than decreased after removing participants who exhibited a retroactive bias, so if anything the association between attentional state ratings and task performance was weakened by these participants rather than enhanced.

These analyses demonstrate that subjective attentional state ratings were meaningful: Detection performance was superior during periods of higher reported attentional state.

TASK PERFORMANCE AND CONFIDENCE RATINGS

To investigate the association between confidence ratings and task performance, we performed a corresponding analysis to that above: We binned trials on the basis of confidence rating decile, then performed one-way ANOVAs on hit rate, false alarm rate and d′ by bin. Note that our confidence scale increases from 0 up to 100 for target present responses and decreases from 0 to −100 for target absent responses, hence we took the absolute value of confidence ratings for this analysis. Grand average hit rate and false alarm rate as a function of confidence rating bin are shown in **Figure 2C**. Higher confidence ratings were associated with better performance in terms of false alarm rate, $F(9, 99) = 34.99$, MSE = 3409.16, $p < 0.001$, $\eta_p^2 = 0.76$, and d′, $F(9, 99) = 14.76$, MSE = 4.69, $p < 0.001$, $\eta_p^2 = 0.57$, but not hit rate, $F(9, 99) = 1.76$, MSE = 532.91, $p = 0.086$. A clear linear trend is evident with false alarm rate decreasing as confidence rating increases and a test of linear trend revealed this to be significant, $F(1, 99) = 189.19$, MSE = 18,435.18, $p < 0.001$, $\eta_p^2 = 0.66$. There was also a significant linear trend for d′, $F(1, 99) = 104.70$, MSE = 33.24, $p < 0.001$, $\eta_p^2 = 0.51$. Hence, higher confidence ratings were associated with better detection performance, as expected; however, it is surprising that the effect was not found for hit rate. This finding seems to be related to the bimodality seen in the distribution of hit trial confidence ratings, in that a large proportion of hits were rated as very low confidence. It could reflect a strong tendency to guess target present rather than absent, or alternatively indicates implicit target detection.

TASK PERFORMANCE AND ATTENTIONAL STATE RATINGS TIME SERIES

One of the main aims of this study was to look at fluctuations of attentional state over the course of the experiment. As an example, **Figure 2D** presents the normalized time series of attentional

state ratings and detection task accuracy for one participant, and **Figure 2E** presents the same time series after smoothing by averaging with a sliding window size of 101 trials. There is a positive correlation between attentional state rating and detection accuracy for this participant. We will return to this issue later with a timescale analysis for all participants on attentional state rating, prestimulus alpha power, and detection performance.

EVENT-RELATED POTENTIALS AND ATTENTIONAL STATE AND CONFIDENCE RATINGS

Figure 3A presents grand averaged ERPs at central and posterior midline scalp electrodes as a function of attentional state rating quartile. **Figure 3B** presents grand averaged ERPs as a function of confidence rating quartile for hit trials only, with miss trials represented by an additional line. The waveforms are time-locked to the onset of the words "Get Ready!" that signified that the RSVP stream would begin in 1 s.

A clear P1–N1 complex followed by an N400 is visible in response to the presentation of the words "Get Ready!", and a SSVEP is generated by the RSVP stream (onset at 1 s). We might have expected these responses to be of greater amplitude in higher rated attentional state trials; however, this was not the case – ERP and SSVEP amplitudes varied little, if at all, as a function of attentional state rating. The P300 in response to targets is evident in the latter half of the SSVEP for the two upper quartiles of attentional state and confidence ratings. For a clearer view, **Figure 3** presents grand averaged ERPs time-locked to target onset, for hit trials only as a function of attentional state rating quartile (**Figure 3C**), and confidence rating quartile (**Figure 3D**), with miss trials presented in a separate line. P300 amplitude in hit trials was significantly larger in trials with higher attentional state ratings, $F(3, 33) = 7.82$, MSE = 24.02, $p < 0.001$, $\eta_p^2 = 0.42$, and in trials with higher confidence ratings, $F(3, 33) = 25.95$, MSE = 43.64, $p < 0.001$, $\eta_p^2 = 0.70$. Both effects followed a linear trend: For attentional state ratings, $F(1, 33) = 20.57$, MSE = 63.20, $p < 0.001$, $\eta_p^2 = 0.38$, and for confidence ratings, $F(1, 33) = 75.83$, MSE = 127.53, $p < 0.001$, $\eta_p^2 = 0.70$. These effects were not driven by differential hit rates because only hit trials were included in the analyses.

EEG ALPHA POWER AND ATTENTIONAL STATE AND CONFIDENCE RATINGS

To assess changes in alpha power as a function of attentional state and confidence rating, we band-pass filtered EEG epochs at alpha frequencies (see Materials and Methods for details), and subsequently computed the envelope of the amplitude-modulated signal, thereby discarding phase information and revealing fluctuations in oscillatory power over time. **Figure 4** shows the grand average amplitude-modulated signal as a function of attentional state rating quartile (**Figure 4A**) and confidence rating quartile (**Figure 4B**), time-locked to the onset of the words "Get Ready!". Alpha power can be seen to steadily increase after the initial response to the words, and peaks at the start of the SSVEP in response to the RSVP stream. Subsequently, alpha power declines sharply up to 0.4 s after RSVP offset.

Inspection of **Figure 4A** reveals that alpha power was lower throughout the epoch in trials rated as high attentional state, as predicted. There was no association of alpha power and confidence rating, except during the second half of the RSVP stream and beyond, perhaps as a result of greater P300 amplitudes in trials with decisions rated as higher confidence. The relationships between

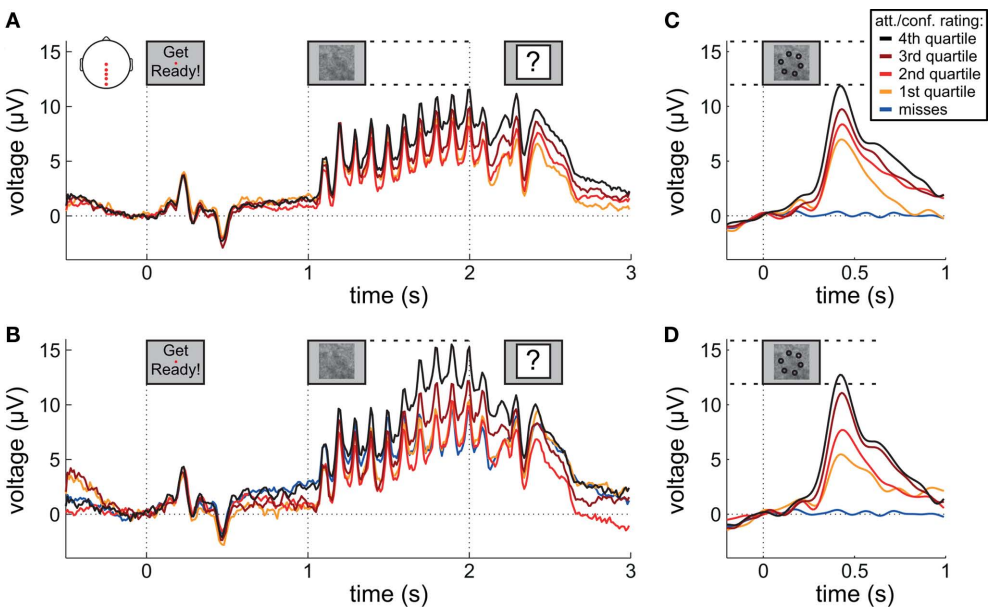


FIGURE 3 | Event-related potentials during preparation and stimulus presentation. (A,B) Grand average ERPs time-locked to the onset of the words "Get Ready!" by attentional state rating quartile and confidence rating quartile, respectively. **(C,D)** Grand average ERPs time-locked to target onset by attentional state rating quartile and confidence rating quartile, respectively. In **(B–D)**, miss trials are shown separately as a blue line. For **(C,D)** EEG epochs were low-pass filtered at 8 Hz. Scalp plot in **(A)** shows the locations of the electrodes included in the analyses in **(A–D)**.

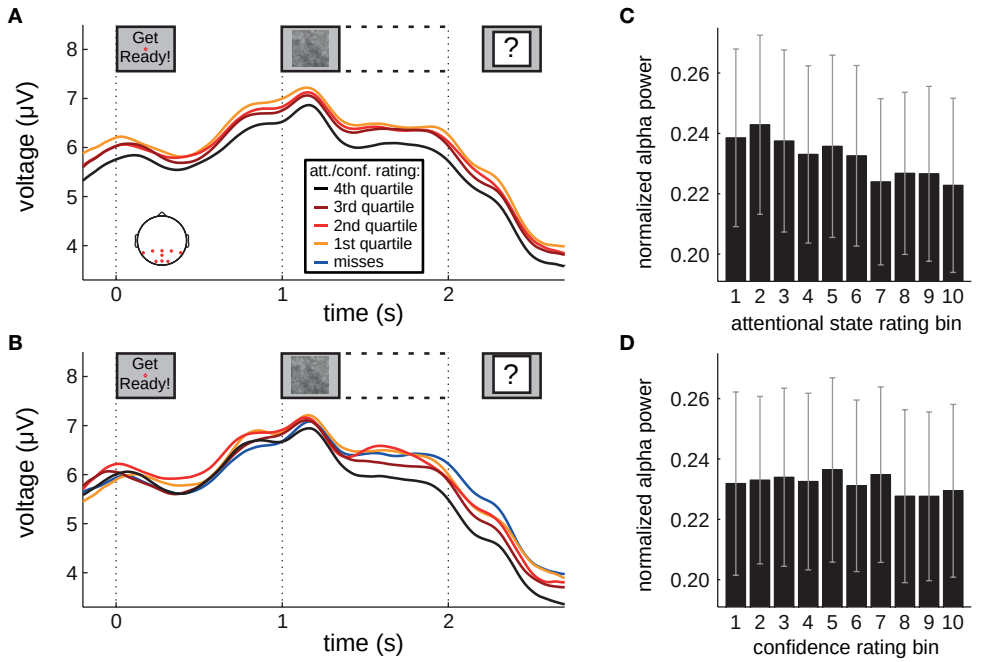


FIGURE 4 | Alpha power during preparation and stimulus presentation. **(A,B)** Grand average amplitude-modulated EEG signal after band-pass filtering at alpha frequencies, time-locked to the onset of the words "Get Ready!", by attentional state rating quartile and confidence rating quartile, respectively. In **(B)**, miss trials are shown separately as a blue line. **(C,D)** Grand average normalized prestimulus alpha power as a function of attentional state rating decile and confidence rating decile, respectively. Error bars show the SE of the mean across participants. Scalp plot in **(A)** shows the locations of the electrodes included in the analyses in **(A–D)**.

prestimulus alpha power and attentional state and confidence ratings can be clearly seen in **Figures 4C,D**, respectively, in which grand average normalized prestimulus alpha power is plotted as a function of attentional state and confidence rating decile. Higher prestimulus alpha power was associated with lower attentional state ratings, $F(9, 99) = 4.27$, MSE = 0.000549, $p < 0.001$, $\eta_p^2 = 0.28$, with a significant linear trend, $F(1, 99) = 32.57$, MSE = 0.00419, $p < 0.001$, $\eta_p^2 = 0.25$, but was not reliably associated with confidence rating, $F(9, 99) = 1.25$, MSE = 0.000103, $p = 0.273$.

This basic analysis illustrates clearly the relationship between attentional state ratings and prestimulus alpha power that we anticipated: Trials in which attentional state was rated as high featured lower prestimulus alpha power. Combined with the finding that attentional state rating was associated with detection performance, this result suggests that both measures reflect attentional state.

PRESTIMULUS EEG ALPHA POWER, ATTENTIONAL STATE RATING, AND TASK PERFORMANCE TIMESCALES

The above analyses reveal a strong relationship between attentional state rating and prestimulus alpha power. To assess whether the strength of this association varies across different timescales, we performed a series of correlations between individual participants' time series of attentional state ratings and normalized prestimulus alpha power, after smoothing the time series of attentional state ratings with increasingly large sliding windows. We then repeated this procedure, leaving the time series of attentional state ratings unsmoothed and instead smoothing the time series of prestimulus alpha power. This analysis is illustrated in a flow chart in **Figure 5A**.

For comparison, we performed a corresponding analysis of the relationship between prestimulus alpha power and task performance (detection task accuracy). The grand average power spectra of the time series of attentional state ratings, prestimulus alpha power, and task accuracy are shown in **Figures 5B–D**, respectively, for reference.

Figures 5E,F present the resulting grand average correlations between alpha power and attentional state ratings (black lines), and between alpha power and detection task accuracy (blue lines). The time series of attentional state ratings and alpha power were negatively correlated, replicating the basic relationship described above. Crucially, this association increased in strength when we smoothed either the time series of attentional state ratings (**Figure 5E**) or alpha power (**Figure 5F**), with the peak correlation occurring at a sliding window size of 75 trials (approximately 7 min) in both analyses. The correlations between alpha power and task performance were much weaker.

We tested the correlations for significance with a series of two-tailed paired-samples *t*-tests between individual participants' *r* values and zero, one for each sliding window size, controlling for multiple comparisons via the Benjamini and Hochberg (1995) procedure, which optimally balances statistical power and the control of false discovery rate for ERP analyses (Lage-Castellanos et al., 2010). Significant correlations are shown in **Figure 5** in red. Whereas none of the correlations between task performance and alpha power were significant, the correlations between attentional state rating and alpha power were significantly different from zero between sliding window sizes of 3 and 133 trials, inclusive, in both analyses.

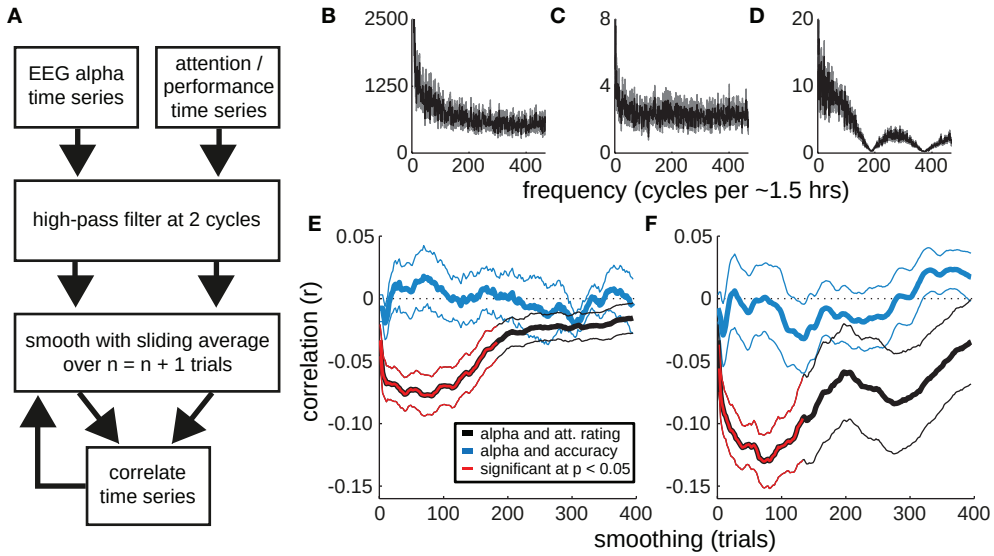


FIGURE 5 | Correlations between attentional state ratings, prestimulus EEG alpha power, and task performance (detection accuracy). (A) Flow chart illustrating the smoothing analysis. **(B–D)** Grand average Fourier power spectra of the trial-by-trial time series of attentional state ratings, prestimulus EEG alpha power, and detection task accuracy, respectively. The SE of the mean is shown in gray. **(E,F)** Grand average correlations (*r*) as a function of the degree of smoothing (sliding window size). The correlations between the time series of attentional state ratings and prestimulus EEG alpha power are shown in black. The correlations between the time series of detection task accuracy and prestimulus EEG alpha power are shown in blue. Significant correlations are highlighted in red. Thin lines show the SE of the mean. In **(E)**, smoothing is applied to the time series of attentional state ratings and detection task accuracy; in **(F)**, smoothing is applied to the time series of prestimulus EEG alpha power.

At first glance, the results of this analysis indicate that attentional state ratings and alpha power co-vary maximally over periods of minutes. However, increasing smoothing window size may be confounded with increasing signal-to-noise ratio (because trial-varying noise should increasingly average to zero with increasing smoothing). To investigate the extent of this issue, and to illustrate further the properties of our novel smoothing analysis, we repeated the analysis on four sets of simulated data. We simulated time series of attentional state ratings and alpha power that were negatively correlated at short timescales, long timescales, both short and long timescales, or at neither timescale. The results, shown in **Figure 6**, indicate that the analysis was effective in identifying whichever correlations were present in the simulated time series, with peaks apparent at corresponding smoothing window sizes, and did not create artifactual correlations where none were present. These simple simulations demonstrate that our smoothed time series correlation analysis was effective in revealing the timescales of covariation between attentional state ratings and prestimulus alpha power. This point is further supported by the fact that there were no significant correlations in the alpha power and task performance analysis. It should be noted, however, that it is still theoretically possible that the associations at short timescales were stronger than our analysis suggests, if there had been increased noise at these timescales.

The peak negative correlation between the time series of attentional state ratings and alpha power occurred at a sliding window size of 75 trials (approximately 7 min). **Figure 7** plots the time series of attentional state ratings and alpha power both smoothed with a sliding window size of 75 trials for each participant individually, together with the correlation coefficient (*r*) between these time series. Strong negative correlations are evident for all participants except one, with fluctuations in attentional state ratings (red lines) varying in the opposite direction to fluctuations in alpha power (black lines).

Overall, this analysis of the timescales of covariation between attentional state ratings and prestimulus alpha power reveals that the association between the two measures was even stronger at longer timescales: The negative correlation between the two measures peaked when we smoothed either of the time series with a sliding window size of approximately 7 min worth of trials.

SINGLE-TRIAL CLASSIFICATION OF ATTENTIONAL STATE RATING BY PRESTIMULUS EEG ALPHA POWER

To explore the reliability of the association between attentional state ratings and prestimulus EEG alpha power on single trials, we conducted single-trial logistic regression classification of upper vs. lower subjective attentional state rating quartile on the basis of prestimulus EEG alpha power for each participant individually. Fourier spectral power values within a 4-Hz band centered on each participant's modal alpha frequency (typically 10 Hz), across 32 electrodes, were used as the input features with which the classifier estimated attentional state rating quartile.

To explore classification performance at longer timescales, we smoothed the time series of attentional state ratings with increasingly large sliding windows, and repeated our classification procedure after each increase in sliding window size. **Figure 8A** presents grand average classification performance (Az score) as a function of the sliding window size. The performance of the classifier improved dramatically as we increased the extent of smoothing. The point of inflection on the curve occurs at a sliding window size of approximately 80 trials, indicating where further smoothing ceases to improve classification performance.

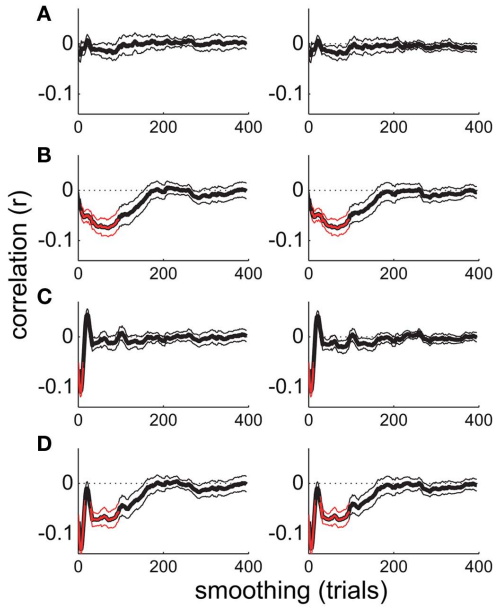


FIGURE 6 | Smoothed time series correlation analysis performed on simulated time series. (A) Uncorrelated time series. **(B–D)** Time series negatively correlated at long timescales, short timescales, and both long and short timescales, respectively. Thin lines show the SE of the mean across the 12 pairs of simulated time series.

Figure 8B presents the average Az score and 95% confidence interval over 1000 repetitions with unsmoothed attentional state ratings (in black) and with attentional state ratings smoothed with a sliding window size of 75 trials (in red), for each participant, ordered from the left by descending classification performance with smoothed data. As a bootstrap test of significance, we repeated the classification procedure a further 1000 times for each participant, shuffling attentional state rating trial labels each time. The resulting average Az score and 95% confidence intervals are shown in **Figure 8B** (in gray). Eight of the 12 participants' average Az scores for unsmoothed attentional state ratings were greater than their bootstrap significance values (the upper end of the 95% confidence interval for Az scores from classification with shuffled attentional state rating trial labels), meaning that classification performance without smoothing was significantly better than chance for those participants. For the smoothed attentional state ratings, all 12 participants' average Az scores were greater than their bootstrap significance values, and nine of them had average Az scores greater than the upper end of the 95% confidence interval for Az scores from classification with unsmoothed attentional state ratings. The three participants whose classification performance did not improve with smoothing, however, may have already been at ceiling – their average Az scores with unsmoothed attentional state ratings were already above 0.85. Smoothing attentional state ratings therefore improved classification performance significantly for all participants with average Az scores below 0.85.

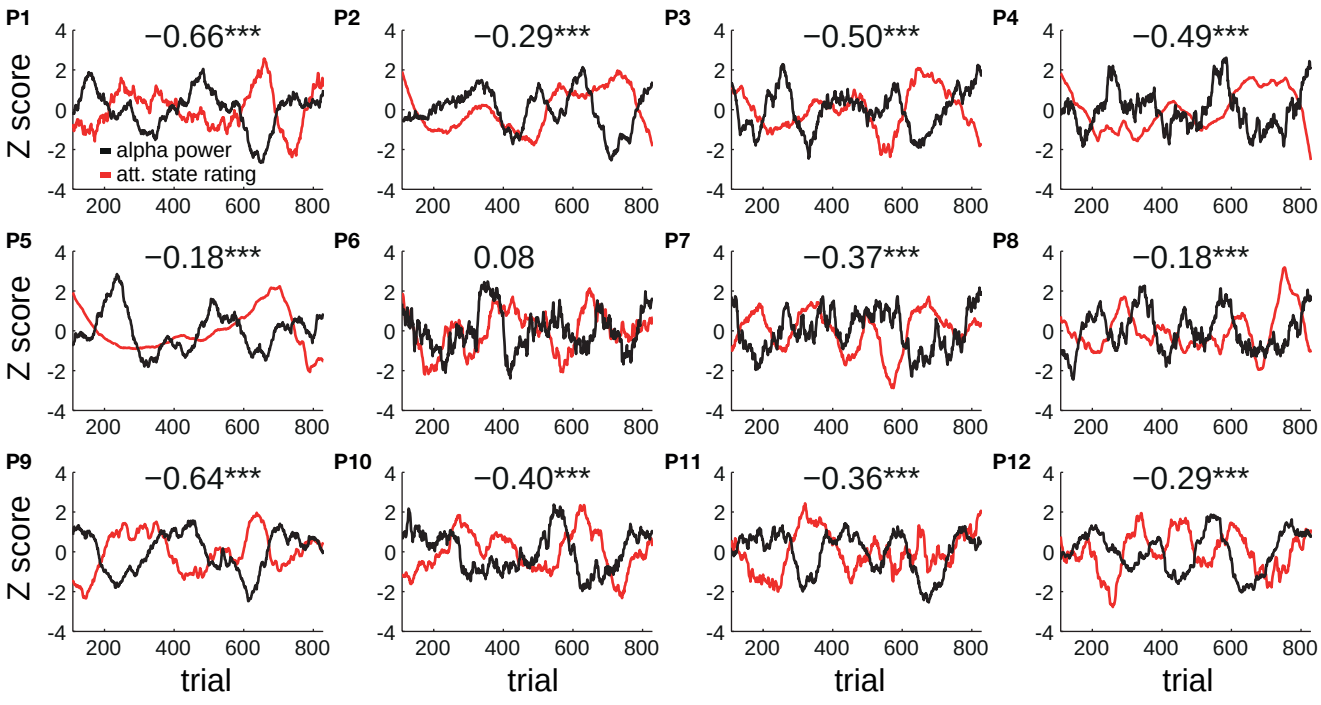


FIGURE 7 | Individual participants' smoothed time series of attentional state ratings and prestimulus EEG alpha power. Normalized time series smoothed with a sliding window size of 75 trials are shown for each participant individually (P1–P12). The correlation coefficient (*r*) between the two time series for each participant is shown at the top of each plot. Significant correlations are indicated by a star. Triple starred correlations are significant at $p < 0.001$.

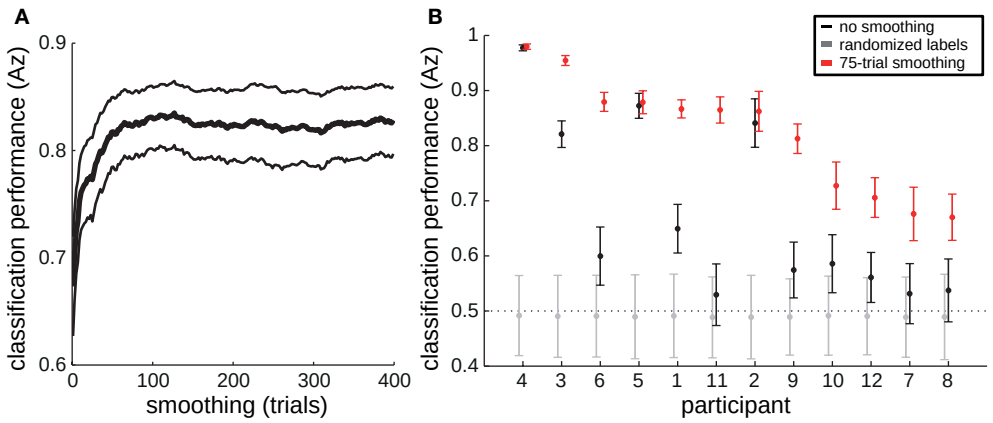


FIGURE 8 | Single-trial classification of attentional state rating on the basis of prestimulus EEG alpha power. (A) Grand average (thick line) and SE (thin lines) Az score as a function of the degree of smoothing (sliding window size) to the time series of attentional state ratings. **(B)** Average Az scores and 95% confidence intervals over 1000 repetitions for each participant, for attentional state ratings unsmoothed (in black), smoothed with a sliding window size of 75 trials (in red), and with shuffled attentional state rating trial labels (in gray).

In summary, trial-by-trial variations in prestimulus EEG alpha power reflected subjective attentional state for most participants, and smoothing attentional state ratings with a sliding window size of approximately 7 min improved classification performance considerably.

DISCUSSION

In this study, participants' subjective ratings of attentional state and perceptual decision confidence were collected on continuous scales at the end of each trial of an RSVP target detection task while recording EEG. We found that attentional state and confidence ratings were both strongly positively associated with task performance and P300 amplitude, and that attentional state ratings were also associated negatively with spontaneous prestimulus EEG alpha activity. Furthermore, we could reliably classify single trials as upper or lower quartile reported attentional state on the basis of prestimulus EEG alpha power for most participants. We also investigated at what timescale these associations were strongest, and found that smoothing with a sliding window size of around 7 min worth of trials yielded the maximum correlation values. In addition, classification performance improved when we smoothed attentional state ratings to the same extent beforehand. Our results suggest that participants' subjective ratings of their attentional state are informative in behavioral terms and reflect the excitability of visual cortex.

ATTENTIONAL STATE RATINGS PREDICTED PERFORMANCE

Higher attentional state ratings were associated with better task performance. This association persisted – in fact, became stronger – when participants with a bias for reporting attentional state as higher on target present trials were excluded from the analysis, and persisted when we repeated the analysis using the attentional state rating from the next trial. This finding supports the proposal that perceptual awareness of external events is attenuated during periods of low attentional state (Smallwood and Schooler, 2006).

Previous studies have demonstrated that perceptual task errors are more likely preceding reports of being unfocused on the task, in paradigms in which subjective attentional state was assessed on a discrete scale via probes occurring at random intervals (e.g., Smallwood et al., 2004, 2008; McVay and Kane, 2009). We queried participants' attentional state on a continuous scale and after every trial, and yet ratings were strongly associated with performance, hence our study demonstrates that attentional state can be indexed efficiently on a continuous scale and in individual trials. These aspects of our design are advantageous because they capture more of the variability in attentional state, both of the construct itself and of its temporal characteristics.

One disadvantage of our design, however, is that in the interests of speed, we did not ask participants to report what they were thinking about when they rated their attentional state, as some previous studies have done (Teasdale et al., 1993, 1995). We therefore do not know whether reports of low attentional state were accompanied by task-unrelated thoughts (i.e., mind-wandering) or would be better characterized as periods of low vigilance. This issue could be addressed in a future study.

Studies of subjective attentional state have typically employed tasks that are optimized to find an association between task errors and reports of low attentional state, such as response inhibition tasks, in which a response that is required for most trials must be withheld on infrequently occurring target trials (Smallwood et al., 2004, 2008; McVay and Kane, 2009). Since our task required a response to both targets and non-targets, and targets were presented frequently (in 50% of trials), we extend previous demonstrations that subjective judgments of attentional state are associated with task performance (Smallwood et al., 2004, 2008; McVay and Kane, 2009) to a standard type of perceptual decision task. Hence, our results lend wider support to the proposition that sensory awareness of the external world is reduced during periods of low attentional state (Smallwood and Schooler, 2006).

ATTENTIONAL STATE RATINGS WERE NOT ASSOCIATED WITH ERP OR SSVEP AMPLITUDE

From the proposal that sensory perception is attenuated during periods of low attentional state (Smallwood and Schooler, 2006), we might have expected to find that EEG responses related to the perception of visual events in our task – such as the ERP in response

to the words "Get Ready!", or the SSVEP generated by the RSVP stream – would be reduced in trials in which attentional state had been rated as low. Previous studies have produced discrepant results in this regard. Kam et al. (2011) reported a smaller P1 in response to visual stimuli in trials preceding reports of low task focus, whereas Smallwood et al. (2008) found no such effects in a similar paradigm, and O'Connell et al. (2009) also reported no such effects preceding stimulus duration judgment errors in a study showing steady increases in parieto-occipital alpha power in the 20 s leading up to errors. Kam et al. (2011) suggest that these two studies, both of which featured visual task stimuli presented foveally, failed to find variations in visually evoked EEG responses because the effect is limited to peripherally presented stimuli and does not occur, or is far weaker, for foveally presented stimuli, as has been suggested to be generally true of attentional effects on foveated stimuli (e.g., Handy and Khoe, 2005). Our RSVP stream images, although larger in retinal representation than the fovea, were centered at fixation, and target patterns were not much larger than foveal area, which might explain our lack of such an effect.

ATTENTIONAL STATE RATINGS AND CONFIDENCE RATINGS PREDICTED P300 AMPLITUDE

P300 amplitude in hit trials was strongly associated with both perceptual decision confidence rating and attentional state rating. Binary subjective judgments of attentional state recorded at random intervals have previously been shown to be associated with P300 amplitude (Smallwood et al., 2008), as has the allocation of spatial attention (e.g., Duncan-Johnson and Donchin, 1977; Johnson, 1993), and stimulus intensity (Roth et al., 1984; Covington and Polich, 1996). The effects of attentional state and spatial attention might therefore operate via a common mechanism: The enhancement of perceptual representations of task stimuli.

P300 amplitude was also predicted by confidence rating. If we assume that confidence ratings are primarily determined by target visibility (in our task images the background noise generated variance in target visibility), then this result could reflect the influence of stimulus intensity on P300 amplitude as has been reported previously (Roth et al., 1984; Covington and Polich, 1996). On the other hand, since neural representations are themselves noisy, it could alternatively suggest that P300 amplitude reflects the perceptual evidence available for a decision, with variability in attention and confidence reflecting the variability in perceptual representation of the target.

In combination with the fact that we found no differences in perceptual EEG responses as a function of attentional state rating, these strong associations of perceptual confidence and attentional state ratings with P300 amplitude suggest that the effects of attentional state occur at the decision level rather than the perceptual level, and hence further along the neural processing stream.

PRESTIMULUS EEG ALPHA POWER PREDICTED ATTENTIONAL STATE RATING

We propose that participants' subjective attentional state ratings reflect both vigilance and the balance of attentional focus between external task-relevant information (in our case visual information) and internal task-unrelated thoughts (i.e., mind-wandering). This assertion is supported firstly in behavioral terms by the finding that attentional state ratings were positively associated with task performance, and secondly in neurophysiological terms by the finding that attentional state ratings were negatively associated on a trial-by-trial basis with spontaneous prestimulus parieto-occipital EEG alpha power. It has been widely reported that alpha power in these regions of cortex tracks the spatial focus of visual attention (e.g., Worden et al., 2000; Sauseng et al., 2005b), and there is some evidence that overall parieto-occipital alpha power is indicative of disengagement with external visual input, since it increases when attention is focused on other modalities (Foxe et al., 1998; Fu et al., 2001; Linkenkaer-Hansen et al., 2004), or during internal cognitive tasks (e.g., Ray and Cole, 1985; Cooper et al., 2003, 2006; Sauseng et al., 2005a), and decreases when attention is redirected toward visual input (Klimesch et al., 1998) or is refocused following visual task errors (Carp and Compton, 2009; Mazaheri et al., 2009). We have now demonstrated that parieto-occipital alpha power reflects attentional state. Importantly, our results imply that alpha power changes related to attentional state are accessible to consciousness.

We found that the association between attentional state ratings and prestimulus EEG alpha power increased in strength when we smoothed either or both of them in time, with the peak negative correlation occurring at a sliding window size of approximately 7 min worth of trials. In addition, the performance of our classifier reached an asymptote when attentional state ratings were smoothed to the same extent. Although participants did report fluctuations in attentional state at shorter timescales than this – indeed, with significant variability from trial to trial – these faster fluctuations were not as strongly associated with prestimulus alpha power in our analysis. This finding does not definitively indicate that fluctuations at shorter timescales were more weakly associated, however, because these faster fluctuations could have been subject to increased noise, thereby curtailing the ability of our analysis to reveal them. Nevertheless, smoothing did more than merely boost the signal-to-noise characteristics of the data: The correlation between attentional state ratings and alpha power did not continue to improve with larger sliding windows, in fact, it returned to baseline well before the maximum sliding window size was reached. Furthermore, we did not see any significant correlations in the timescale analysis of alpha power and task performance.

Previous EEG alpha and attention studies, as well as investigations of subjective attentional state, have not systematically analyzed the timescales of variability, but previous research has shown that behavioral performance (e.g., Verplanck et al., 1952; Gilden et al., 1995, see Gilden, 2001, for a review) and neural activity (e.g., Linkenkaer-Hansen et al., 2001; Leopold et al., 2003; Nikulin and Brismar, 2005) both exhibit 1/f frequency spectra. This type of distribution implies that the largest fluctuations in behavioral and neural responses over extended periods are at the longer timescales, i.e., minutes rather than seconds. Hence, our finding that attentional state ratings and prestimulus alpha power co-vary strongly when smoothed with a sliding window size of several minutes' worth of trials is in agreement with this literature. In addition, many studies investigating periodicities in reaction times and error rates in detection tasks have reported cycle times in the range of 2–10 min (e.g., Wertheimer, 1953; Elliott, 1960; Makeig and Inlow, 1993; Conte et al., 1995; Smith et al., 2003; Arruda et al., 2009; Aue et al., 2009). In particular, Monto et al.

(2008) have demonstrated that somatosensory detection performance and power in all frequency bands are correlated with the phase of infraslow fluctuations (0.1–0.01 Hz, i.e., 10–100 s cycles) of EEG, a finding that could be interpreted as reflecting slow variations in attentional state. Collectively, these findings suggest that changes in attentional state over time are not solely characterized by simple monotonic declines, since fluctuations occur over periods of minutes. Our study provides a link between this literature and research into subjective attentional state, spontaneous EEG alpha, and attention. Our findings are also informative with respect to online monitoring of user attention in contexts in which the users' attentional state could be a critical determinant of task performance. In particular, our results suggest that measures of attentional state might be optimally assessed with a frequency in the region of minutes, and that periods of low attentional state can be relatively long lasting.

PRESTIMULUS EEG ALPHA POWER WAS NOT PREDICTIVE OF DETECTION PERFORMANCE

Previous studies featuring simple visual detection tasks have demonstrated that prestimulus alpha power is greater preceding failures to detect the target than successes (Ergenoglu et al., 2004; van Dijk et al., 2008; Busch et al., 2009). We found no such association in our data, which is surprising given the pairwise associations we found between prestimulus alpha power and attentional state ratings, and between attentional state ratings and detection accuracy. However, previous studies featured paradigms that were optimized for detecting associations between alpha power and visual awareness by precisely controlling stimulus conditions, such that target stimuli were identical from trial-to-trial. In contrast, our design was motivated by the desire to enable the extrapolation of our results to real-world tasks, for example, the detection of landmarks in satellite photographs (e.g., Poolman et al., 2008). In particular, our target stimuli were intended to vary in visibility considerably due to being embedded in white noise. Hence, an effect of prestimulus EEG alpha power on detection accuracy may have simply been lost in the inherent noise of EEG recordings coupled with the noise in the visibility of our target stimuli. Indeed, several other previous studies featuring much more controlled stimuli, akin to those aforementioned, have also failed to find an association between prestimulus alpha power and detection accuracy (e.g., Thut et al., 2006).

One other feature of our paradigm is potentially relevant in this context: Targets were never presented in the first two images of the RSVP stream, so there was always a delay of at least 200 ms between the start of the RSVP stream and target presentation. It is possible that this delay gave participants time to re-engage with the task following prestimulus periods of high alpha power, thereby attenuating any effect of alpha power on task performance. However, because attentional state ratings were associated with detection accuracy as well as prestimulus alpha power, this explanation implies that detection accuracy and prestimulus alpha power are associated with distinct portions of trial-to-trial variance in attentional state ratings. This idea could be addressed in a future study using a visual detection paradigm optimized to find an association between prestimulus alpha power and detection accuracy (e.g., Busch et al., 2009), in which attentional state ratings are collected on every trial, as we did. Such a study might find that attentional state ratings correlate with detection performance at shorter timescales than they do with alpha power, supporting the notion that attentional state is multi-faceted, featuring dissociable relationships with behavioral and neural measures.

CONCLUSION

We have demonstrated a strong association between subjective ratings of attentional state and prestimulus EEG alpha power. This association implies that the level of alpha activity in parieto-occipital cortex not only reflects cortical excitability (e.g., Klimesch et al., 2007), but is also indicative of attentional state with respect to a visual task. Furthermore, it suggests that fluctuations of alpha activity are accessible to consciousness. We propose that participants' subjective attentional state ratings reflect both vigilance and the balance of attentional focus between external task-relevant information and internal task-unrelated thoughts (i.e., mind-wandering). Because these ratings were associated with objective measures of both behavior and neural activity, we suggest that they are a simple and effective means of estimating task engagement. As such, they could provide valuable information in operational settings in which monitoring users' attentional state might permit the optimization of task parameters such as stimulus presentation rate or the weight assigned to their decisions as their attentional state waxes and wanes.

ACKNOWLEDGMENTS

We thank Misha Pavel at Oregon Health and Science University for the images used in the experiment, Patricia Ververs at Honeywell Laboratories for useful discussions and feedback, and Sam Rogers at the University of Oxford for help with data collection. This work was supported by the Defense Advanced Research Projects Agency (DARPA) under contract N10PC20048. The views, opinions, and/or findings contained in this article/presentation are those of the authors and should not be interpreted as representing the official views or policies, either expressed or implied, of DARPA or the United States Department of Defense.

REFERENCES

Arruda, J. E., Zhang, H., Amoss, R. T., Coburn, K. L., and Aue, W. R. (2009). Rhythmic oscillations in quantitative EEG measured during a continuous performance task. *Appl. Psychophysiol. Biofeedback* 34, 7–16.

Aue, W. R., Arruda, J. E., Kass, S. J., and Stanny, C. J. (2009). Cyclic variations in sustained human performance. *Brain Cogn.* 71, 336–344.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Series B Stat. Methodol.* 57, 289–300.

Berger, H. (1929). Uber das elektrenkephalogramm des menschen. *Arch. Psychiatr. Nervenkr.* 87, 527–570.

Braboszcz, C., and Delorme, A. (2011). Lost in thoughts: neural markers of low alertness during mind wandering. *Neuroimage* 54, 3040–3047.

Brainard, D. H. (1997). The psychophysics toolbox. *Spat. Vis.* 10, 433–436.

Busch, N. A., Dubois, J., and VanRullen, R. (2009). The phase of ongoing EEG oscillations predicts visual perception. *J. Neurosci.* 29, 7869–7876.

Busch, N. A., and Herrmann, C. S. (2003). Object-load and feature-load modulate EEG in a short-term memory task. *Neuroreport* 14, 1721–1724.

Capotosto, P., Babiloni, C., Romani, G. L., and Corbetta, M. (2009). Frontoparietal cortex controls spatial attention through modulation of anticipatory alpha rhythms. *J. Neurosci.* 29, 5863–5872.

Carp, J., and Compton, R. J. (2009). Alpha power is influenced by performance errors. *Psychophysiology* 46, 336–343.

Christoff, K., Gordon, A. M., Smallwood, J., Smith, R., and Schooler, J. W. (2009). Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. *Proc. Natl. Acad. Sci. U.S.A.* 106, 8719–8724.

Conte, S., Ferlazzo, F., and Renzi, P. (1995). Ultradian rhythms of reaction times in performance in vigilance tasks. *Biol. Psychol.* 39, 159–172.

Cooper, N. R., Burgess, A. P., Croft, R. J., and Gruzelier, J. H. (2006). Investigating evoked and induced electroencephalogram activity in task-related alpha power increases during an internally directed attention task. *Neuroreport* 17, 205–208.

Cooper, N. R., Croft, R. J., Dominey, S. J. J., Burgess, A. P., and Gruzelier, J. H. (2003). Paradox lost? Exploring the role of alpha oscillations during externally vs. internally directed attention and the implications for idling and inhibition hypotheses. *Int. J. Psychophysiol.* 47, 65–74.

Covington, J. W., and Polich, J. (1996). P300, stimulus intensity, and modality. *Electroencephalogr. Clin. Neurophysiol.* 100, 579–584.

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

Doesburg, S. M., Green, J. J., McDonald, J. J., and Ward, L. M. (2009). From local inhibition to long-range integration: a functional dissociation of alpha-band synchronization across cortical scales in visuospatial attention. *Brain Res.* 1303, 97–110.

Duncan-Johnson, C. C., and Donchin, E. (1977). On quantifying surprise: the variation of event-related potentials with subjective probability. *Psychophysiology* 14, 456–467.

Eichele, T., Debener, S., Calhoun, V. D., Specht, K., Engel, A. K., Hugdahl, K., von Cramon, D. Y., and Ullsperger, M. (2008). Prediction of human errors by maladaptive changes in event-related brain networks. *Proc. Natl. Acad. Sci. U.S.A.* 105, 6173–6178.

Elliott, E. (1960). Perception and alertness. *Ergonomics* 3, 357.

Ergenoglu, T., Demiralp, T., Bayraktaroglu, Z., Ergen, M., Beydagi, H., and Uresin, Y. (2004). Alpha rhythm of the EEG modulates visual detection performance in humans. *Brain Res. Cogn. Brain Res.* 20, 376–383.

Forster, S., and Lavie, N. (2009). Harnessing the wandering mind: the role of perceptual load. *Cognition* 111, 345–355.

Foxe, J. J., Simpson, G. V., and Ahlfors, S. P. (1998). Parieto-occipital approximately 10 Hz activity reflects anticipatory state of visual attention mechanisms. *Neuroreport* 9, 3929–3933.

Fu, K. M., Foxe, J. J., Murray, M. M., Higgins, B. A., Javitt, D. C., and Schroeder, C. E. (2001). Attention-dependent suppression of distracter visual input can be cross-modally cued as indexed by anticipatory parieto-occipital alpha-band oscillations. *Brain Res. Cogn. Brain Res.* 12, 145–152.

Gerson, A. D., Parra, L. C., and Sajda, P. (2006). Cortically coupled computer vision for rapid image search. *IEEE Trans. Neural Syst. Rehabil. Eng.* 14, 174–179.

Gilden, D. L. (2001). Cognitive emissions of 1/f noise. *Psychol. Rev.* 108, 33–56.

Gilden, D. L., Thornton, T., and Mallon, M. W. (1995). 1/f noise in human cognition. *Science* 267, 1837–1839.

Handy, T. C., and Khoe, W. (2005). Attention and sensory gain control: a peripheral visual process? *J. Cogn. Neurosci.* 17, 1936–1949.

Hari, R., Salmelin, R., Mäkelä, J. P., Salenius, S., and Helle, M. (1997). Magnetoencephalographic cortical rhythms. *Int. J. Psychophysiol.* 26, 51–62.

Holroyd, C. B., and Coles, M. G. H. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol. Rev.* 109, 679–709.

Jensen, O., Gelfand, J., Kounios, J., and Lisman, J. E. (2002). Oscillations in the alpha band (9-12 Hz) increase with memory load during retention in a short-term memory task. *Cereb. Cortex* 12, 877–882.

Johnson, R. (1993). On the neural generators of the P300 component of the event-related potential. *Psychophysiology* 30, 90–97.

Kam, J. W. Y., Dao, E., Farley, J., Fitzpatrick, K., Smallwood, J., Schooler, J. W., and Handy, T. C. (2011). Slow fluctuations in attentional control of sensory cortex. *J. Cogn. Neurosci.* 23, 460–470.

Kelly, S. P., Gomez-Ramirez, M., and Foxe, J. J. (2009). The strength of anticipatory spatial biasing predicts target discrimination at attended locations: a high-density EEG study. *Eur. J. Neurosci.* 30, 2224–2234.

Kelly, S. P., Lalor, E. C., Reilly, R. B., and Foxe, J. J. (2006). Increases in alpha oscillatory power reflect an active retinotopic mechanism for distracter suppression during sustained visuospatial attention. *J. Neurophysiol.* 95, 3844–3851.

Klimesch, W., Doppelmayr, M., Russegger, H., Pachinger, T., and Schwaiger, J. (1998). Induced alpha band power changes in the human EEG and attention. *Neurosci. Lett.* 244, 73–76.

Klimesch, W., Sauseng, P., and Hanslmayr, S. (2007). EEG alpha oscillations: the inhibition-timing hypothesis. *Brain Res. Rev.* 53, 63–88.

Lage-Castellanos, A., Martínez-Montes, E., Hernández-Cabrera, J. A., and Galán, L. (2010). False discovery rate and permutation test: an evaluation in ERP data analysis. *Stat. Med.* 29, 63–74.

Leopold, D. A., Murayama, Y., and Logothetis, N. K. (2003). Very slow activity fluctuations in monkey visual cortex: implications for functional brain imaging. *Cereb. Cortex* 13, 422–433.

Linkenkaer-Hansen, K., Nikouline, V. V., Palva, J. M., and Ilmoniemi, R. J. (2001). Long-range temporal correlations and scaling behavior in human brain oscillations. *J. Neurosci.* 21, 1370–1377.

Linkenkaer-Hansen, K., Nikulin, V. V., Palva, S., Ilmoniemi, R. J., and Palva, J. M. (2004). Prestimulus oscillations enhance psychophysical performance in humans. *J. Neurosci.* 24, 10186–10190.

Makeig, S., and Inlow, M. (1993). Lapses in alertness: coherence of fluctuations in performance and EEG spectrum. *Electroencephalogr. Clin. Neurophysiol.* 86, 23–35.

Mason, M. F., Norton, M. I., Van Horn, J. D., Wegner, D. M., Grafton, S. T., and Macrae, C. N. (2007). Wandering minds: the default network and stimulus-independent thought. *Science* 315, 393–395.

Mathan, S. (2008). Image search at the speed of thought. *Interactions* 15, 76–77.

Mathan, S., Erdogmus, D., Huang, Y., Pavel, M., Ververs, P., Carciofini, J., Dorneich, M., and Whitlow, S. (2008). "Rapid image analysis using neural signals," in *CHI '08 Extended Abstracts on Human Factors in Computing Systems, CHI EA '08* (New York, NY: ACM), 3309–3314.

Mathan, S., Whitlow, S., Erdogmus, D., Pavel, M., Ververs, P., and Dorneich, M. (2006). "Neurophysiologically driven image triage: a pilot study," in *CHI '06 Extended Abstracts on Human Factors in Computing Systems, CHI EA '06* (New York, NY: ACM), 1085–1090.

Mazaheri, A., Nieuwenhuis, I. L. C., van Dijk, H., and Jensen, O. (2009). Prestimulus alpha and mu activity predicts failure to inhibit motor responses. *Hum. Brain Mapp.* 30, 1791–1800.

McVay, J. C., and Kane, M. J. (2009). Conducting the train of thought: working memory capacity, goal neglect, and mind wandering in an executive-control task. *J. Exp. Psychol. Learn. Mem. Cogn.* 35, 196–204.

Monto, S., Palva, S., Voipio, J., and Palva, J. M. (2008). Very slow EEG fluctuations predict the dynamics of stimulus detection and oscillation amplitudes in humans. *J. Neurosci.* 28, 8268–8272.

Nikulin, V. V., and Brismar, T. (2005). Long-range temporal correlations in electroencephalographic oscillations: relation to topography, frequency band, age and gender. *Neuroscience* 130, 549–558.

O'Connell, R. G., Dockree, P. M., Robertson, I. H., Bellgrove, M. A., Foxe, J. J., and Kelly, S. P. (2009). Uncovering the neural signature of lapsing attention: electrophysiological signals predict errors up to 20 s before they occur. *J. Neurosci.* 29, 8604–8611.

Palva, J. M., Palva, S., and Kaila, K. (2005). Phase synchrony among neuronal oscillations in the human cortex. *J. Neurosci.* 25, 3962–3972.

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

Parra, L. C., Christoforou, C., Gerson, A. D., Dyrholm, M., Luo, A., Wagner, M., Philiastides, M. G., and Sajda, P. (2008). Spatiotemporal linear decoding of brain state: application to performance augmentation in high-throughput tasks. *IEEE Signal Process. Mag.* 25, 95–115.

Pfurtscheller, G. (2001). Functional brain imaging based on ERD/ERS. *Vision Res.* 41, 1257–1260.

Pollen, D. A., and Trachtenberg, M. C. (1972). Some problems of occipital alpha block in man. *Brain Res.* 41, 303–314.

Poolman, P., Frank, R. M., Luu, P., Pederson, S. M., and Tucker, D. M. (2008). A single-trial analytic framework for EEG analysis and its application to target detection and classification. *Neuroimage* 42, 787–798.

Ray, W. J., and Cole, H. W. (1985). EEG alpha activity reflects attentional demands, and beta activity reflects emotional and cognitive processes. *Science* 228, 750–752.

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., and Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science* 306, 443–447.

Rihs, T. A., Michel, C. M., and Thut, G. (2007). Mechanisms of selective inhibition in visual spatial attention are indexed by alpha-band EEG synchronization. *Eur. J. Neurosci.* 25, 603–610.

Robertson, I. H., Manly, T., Beschin, N., Daini, R., Haeske-Dewick, H., Hömberg, V., Jehkonen, M., Pizzamiglio, G., Shiel, A., and Weber, E.

(1997). Auditory sustained attention is a marker of unilateral spatial neglect. *Neuropsychologia* 35, 1527–1532.

Romei, V., Brodbeck, V., Michel, C. M., Amedi, A., Pascual-Leone, A., and Thut, G. (2008a). Spontaneous fluctuations in posterior alpha-band EEG activity reflect variability in excitability of human visual areas. *Cereb. Cortex* 18, 2010–2018.

Romei, V., Rihs, T. A., Brodbeck, V., and Thut, G. (2008b). Resting electroencephalogram alpha-power over posterior sites indexes baseline visual cortex excitability. *Neuroreport* 19, 203–208.

Romei, V., Gross, J., and Thut, G. (2010). On the role of prestimulus alpha rhythms over occipito-parietal areas in visual input regulation: correlation or causation? *J. Neurosci.* 30, 8692–8697.

Roth, W. T., Dorato, K. H., and Kopell, B. S. (1984). Intensity and task effects on evoked physiological responses to noise bursts. *Psychophysiology* 21, 466–481.

Sauseng, P., Klimesch, W., Doppelmayr, M., Pecherstorfer, T., Freunberger, R., and Hanslmayr, S. (2005a). EEG alpha synchronization and functional coupling during top-down processing in a working memory task. *Hum. Brain Mapp.* 26, 148–155.

Sauseng, P., Klimesch, W., Stadler, W., Schabus, M., Doppelmayr, M., Hanslmayr, S., Gruber, W. R., and Birbaumer, N. (2005b). A shift of visual spatial attention is selectively associated with human EEG alpha activity. *Eur. J. Neurosci.* 22, 2917–2926.

Smallwood, J., Beach, E., Schooler, J. W., and Handy, T. C. (2008). Going AWOL in the brain: mind wandering reduces cortical analysis of external events. *J. Cogn. Neurosci.* 20, 458–469.

Smallwood, J., Davies, J. B., Heim, D., Finnigan, F., Sudberry, M., O'Connor, R., and Obonsawin, M. (2004). Subjective experience and the attentional lapse: task engagement and disengagement during sustained attention. *Conscious. Cogn.* 13, 657–690.

Smallwood, J., and Schooler, J. W. (2006). The restless mind. *Psychol. Bull.* 132, 946–958.

Smith, K. J., Valentino, D. A., and Arruda, J. E. (2003). Rhythmic oscillations in the performance of a sustained attention task. *J. Clin. Exp. Neuropsychol.* 25, 561–570.

Snyder, A. C., and Foxe, J. J. (2010). Anticipatory attentional suppression of visual features indexed by oscillatory alpha-band power increases: a high-density electrical mapping study. *J. Neurosci.* 30, 4024–4032.

Stanislaw, H., and Todorov, N. (1999). Calculation of signal detection theory measures. *Behav. Res. Methods Instrum. Comput.* 31, 137–149.

Teasdale, J. D., Dritschel, B. H., Taylor, M. J., Proctor, L., Lloyd, C. A., Nimmo-Smith, I., and Baddeley, A. D. (1995). Stimulus-independent thought depends on central executive resources. *Mem. Cognit.* 23, 551–559.

Teasdale, J. D., Proctor, L., Lloyd, C. A., and Baddeley, A. D. (1993). Working memory and stimulus-independent thought: effects of memory load and presentation rate. *Eur. J. Cogn. Psychol.* 5, 417–433.

Thut, G., Nietzel, A., Brandt, S. A., and Pascual-Leone, A. (2006). Alpha-band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. *J. Neurosci.* 26, 9494–9502.

van Dijk, H., Schoffelen, J.-M., Oostenveld, R., and Jensen, O. (2008). Prestimulus oscillatory activity in the alpha band predicts visual discrimination ability. *J. Neurosci.* 28, 1816–1823.

Verplanck, W. S., Collier, G. H., and Cotton, J. W. (1952). Nonindependence of successive responses in measurements of the visual threshold. *J. Exp. Psychol.* 44, 273–282.

Wagenmakers, E.-J., Farrell, S., and Ratcliff, R. (2004). Estimation and interpretation of 1/falpha noise in human cognition. *Psychon. Bull. Rev.* 11, 579–615.

Wertheimer, M. (1953). An investigation of the randomness of threshold measurements. *J. Exp. Psychol.* 45, 294–303.

Worden, M. S., Foxe, J. J., Wang, N., and Simpson, G. V. (2000). Anticipatory biasing of visuospatial attention indexed by retinotopically specific alpha-band electroencephalography increases over occipital cortex. *J. Neurosci.* 20, RC63.

Wyart, V., and Tallon-Baudry, C. (2009). How ongoing fluctuations in human visual cortex predict perceptual awareness: baseline shift versus decision bias. *J. Neurosci.* 29, 8715–8725.

Yamagishi, N., Callan, D. E., Goda, N., Anderson, S. J., Yoshida, Y., and Kawato, M. (2003). Attentional modulation of oscillatory activity in human visual cortex. *Neuroimage* 20, 98–113.

Yeung, N., Botvinick, M. M., and Cohen, J. D. (2004). The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychol. Rev.* 111, 931–959.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 11 January 2011; accepted: 16 April 2011; published online: 10 May 2011. Citation: Macdonald JSP, Mathan S and Yeung N (2011) Trial-by-trial variations in subjective attentional state are reflected in ongoing prestimulus EEG alpha oscillations. Front. Psychology ***2****:82. doi: 10.3389/fpsyg.2011.00082*

This article was submitted to Frontiers in Perception Science, a specialty of Frontiers in Psychology.

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 30 March 2011
doi: 10.3389/fpsyg.2011.00051

Increased intra-participant variability in children with autistic spectrum disorders: evidence from single-trial analysis of evoked EEG

Elizabeth Milne*

Sheffield Autism Research Lab, Department of Psychology, The University of Sheffield, Sheffield, UK

Edited by:
Guillaume A. Rousselet, University of Glasgow, UK

Reviewed by:
David R. Simmons, University of Glasgow, UK
Cyril R. Pernet, University of Edinburgh, UK

****Correspondence:***
Elizabeth Milne, Sheffield Autism Research Lab, Department of Psychology, The University of Sheffield, Western Bank, Sheffield, South Yorkshire S10 2TN, UK.
e-mail: e.milne@sheffield.ac.uk

Intra-participant variability in clinical conditions such as autistic spectrum disorder (ASD) is an important indicator of pathophysiological processing. The data reported here illustrate that trial-by-trial variability can be reliably measured from EEG, and that intra-participant EEG variability is significantly greater in those with ASD than in neuro-typical matched controls. EEG recorded at the scalp is a linear mixture of activity arising from muscle artifacts and numerous concurrent brain processes. To minimize these additional sources of variability, EEG data were subjected to two different methods of spatial filtering. (i) The data were decomposed using infomax independent component analysis, a method of blind source separation which un-mixes the EEG signal into components with maximally independent time-courses, and (ii) a surface Laplacian transform was performed (current source density interpolation) in order to reduce the effects of volume conduction. Data are presented from 13 high functioning adolescents with ASD without co-morbid ADHD, and 12 neuro-typical age-, IQ-, and gender-matched controls. Comparison of variability between the ASD and neuro-typical groups indicated that intra-participant variability of P1 latency and P1 amplitude was greater in the participants with ASD, and inter-trial α-band phase coherence was lower in the participants with ASD. These data support the suggestion that individuals with ASD are less able to synchronize the activity of stimulus-related cell assemblies than neuro-typical individuals, and provide empirical evidence in support of theories of increased neural noise in ASD.

Keywords: autism, EEG, variability, perception, noise, synchrony, alpha, phase-locking

INTRODUCTION

Autism spectrum disorder (ASD) is a complex neurodevelopmental disorder that has been estimated to occur in 1.16% of children in the UK (Baird et al., 2006). It is characterized by substantial difficulties in social cognition, interaction, and communication (APA, 1994). In addition to these core deficits, ASD is associated with a wide range of more general impairments in many cognitive domains including, executive function (Hill, 2004), memory (Bennetto et al., 1996), attention (Allen and Courchesne, 2001), and perception (Simmons et al., 2009). An underlying etiology that links impairments across such an array of domains has not yet been identified. The literature on perceptual function is particularly puzzling, as while those with ASD show impaired performance of some tasks, e.g., detecting coherent motion within local motion noise (Milne et al., 2002), they show superior performance on tasks that involve detecting a target within a static array (Plaisted et al., 1998). Furthermore, enhanced and diminished perceptual sensitivity appear to co-occur, as Bertone et al. (2005) have demonstrated enhanced first-order contrast perception and decreased second-order contrast perception within the same group of participants.

A persistent difficulty in identifying the etiology of ASD arises from a high level of inter-participant variability. This is evidenced by the number of studies that discuss participant sub-groups within the ASD sample, and the number of studies in which estimates of variation around the mean/median are larger in the ASD group than in the control group. In addition, although many cognitive functions appear to be impaired in those with ASD, the literature is littered with examples of non-replication, suggesting that there is no specific cognitive impairment that is consistent and universal in ASD (see also Happé et al., 2006). The data on impairments in cognitive task performance in ASD may therefore be better interpreted as arising from a pervasive and generalized impairment rather than a collection of several, different, modality, or task specific impairments. A promising candidate for such a generalized impairment is increased levels of intra-participant variability, as this would lead to increased variability between participants, and represents a parsimonious explanation for the many areas of cognition that appear to be impaired in those with ASD. Furthermore, increased intra-participant variability would lead to reduced test re-test reliability and could therefore explain the high level of inconsistency within the literature.

Variability and fluctuation in behavior and task performance are commonly observed in individuals with developmental disorders (Castellanos et al., 2005). Although most typically associated with ADHD, recent empirical work has demonstrated that individuals with ASD show significantly greater intra-individual response time variability during a simple 2AFC task compared not only to a group of typically developing (TD) matched control participants, but also compared to a group of matched participants with ADHD (Geurts et al., 2008). Intra-individual variability appears to be an important

indicator of pathophysiological processing therefore, and its potential to explain a number of task-related behaviors in those with ASD should not be over-looked. However, there are many potential routes to response time variability, including: variability in higher-order cognitive functions such as initiating or generating motor responses (Deutsch and Newell, 2005); transient lapses of attention; and/or variability in early sensory encoding (e.g., Croner et al., 1993; Arieli et al., 1996). The source of intra-participant variability in ASD is therefore unclear.

A number of authors have suggested that neural noise may be increased in individuals with ASD or that increased levels of noise may contribute to reduced cognitive task performance in those with ASD (e.g., Rubenstein and Merzenich, 2003; Baron-Cohen and Belmonte, 2005; Dakin and Frith, 2005; Simmons et al., 2009). In particular some have suggested cortical hyper-excitation especially in primary sensory cortices, which would lead to increased cortical noise in ASD (Rubenstein and Merzenich, 2003). However, as yet, there is no direct evidence for greater neural noise in those with ASD. Increased neural noise would be evidenced by increased variability across individual trials of EEG, therefore the aim of this study is to compare variability in single-trial EEG in a group of children / adolescents with ASD with a group of TD matched controls.

Unfortunately, single-trial analysis of EEG is seldom performed. The main reason for this is that EEG recorded at the scalp constitutes a mixture of a number of sources signals, therefore, activity associated with a single process, being mixed with signals arising from other processes as well as on-going "background" oscillations, is difficult to identify within each trial. Here, EEG data are decomposed with independent component analysis (ICA), which, as described below, un-mixes the different source signals recorded at the scalp, enabling activity from independent processes to be identified in single-trials (Makeig et al., 1997, 2004; Onton et al., 2006) and variability within individuals to be measured.

Whole-brain dynamic processes are underpinned by the formation of cell assemblies, i.e., groups of cells that oscillate in synchrony, or precisely timed succession, for transient periods (Nunez and Srinivasan, 2006). As numerous cell assemblies may be active at any given time, oscillation synchronicity within specific frequency bands is thought to be the mechanism by which the output of single units is identified as being part of a coherent network (Singer et al., 1997). Performing even a simple experimental task will excite a number of different cell assemblies which will be active alongside numerous other task-unrelated assemblies. A difficulty faced by EEG researchers is the fact that electrical activity generated by these separate assemblies becomes mixed, and, via volume conduction, smeared, across the scalp. That is, each EEG electrode records a mixture of signals arising from multiple cognitive processes and from on-going "background" oscillatory activity. Furthermore, scalp electrodes also record activity from non-brain sources including muscle (eye-movements, blinking, heartbeat) and in some cases from non-physiological electrical sources (e.g., line-noise). Filtering and artifact rejection reduces the influence of some of these unwanted contributions to EEG, however, the spatial mixing of numerous brain-based processes means that the signal of interest, i.e., the signal associated with the cognitive task, is mixed with signals from task-unrelated processes and is therefore difficult to observe and measure on a trial-by-trial basis.

This spatial mixing of EEG has shaped the way that EEG data are analyzed, most notably by leading to a dominance in single-trial averaging to calculate the event-related potential (ERP). The theory behind the ERP technique is that by calculating an average of several time-locked trials, event unrelated activity, being phase- and time-random with respect to the time-locking event, cancels to near zero amplitude, whereas the part of the EEG that is time-locked to the relevant event remains visible in the signal. Single-trial analysis is therefore rejected in favor of "average" event-related analyses. However, given the value of understanding variability across single-trials, just as the SD provides vital information regarding the distribution of values around a mean response time, a number of alternate methods have been put forward for facilitating single-trial analysis of EEG data. These include complex filtering (Salajegheh et al., 2004), maximum likelihood estimation (Jaskowski and Verleger, 1999), parametric modeling (von Spreckelsen and Bromm, 1988), multivariate matching pursuit algorithms (Sieluzycki et al., 2009), general linear model analyses (Pernet et al., 2011), and decomposing data using ICA (Jung et al., 2001). ICA provides an elegant solution to the problems associated with spatial mixing of EEG, and facilitates analysis of single-trials by decomposing EEG data into separate informational components of brain dynamics that closely reflect activity associated with specific cognitive or sensory processes, thus removing the need for time-locked averaging (e.g., see Jung et al., 2001).

Independent component analysis is a method of blind source separation that separates N linear mixtures into N independent informational components (Makeig et al., 1997). It is based on the assumption that source signals are statistically independent whereas signal mixtures are not. Maximizing the joint entropy of the extracted signals gives rise to the "un-mixing matrix" W that, when multiplied by EEG data X, produces the original source signals U, i.e., $U = WX$. The columns of the un-mixing matrix, W, hold coefficients of spatial filters that pass the activity of only one independent source process and suppress all the others. Each IC is represented by the time-course of activation (given by each row of U), and the weights with which the component projects to the electrodes which are given in the inverse of the un-mixing matrix W^{-1}. Plotting these weights onto a schematic head model allows one to visualize the scalp topography of each independent component.

A number of papers have demonstrated the usefulness of decomposing EEG data into ICs, not only for isolating artifactual contributions to scalp EEG, but also for studying on-going oscillatory activity, and event-related activity that contribute to ERP deflections recorded at the scalp (for examples see Debener et al., 2005; Onton et al., 2005; Eichele et al., 2010). Here, ICA is used to identify independent components that represent early (<200 ms) activity evoked by presentation of a simple visual stimulus, in order to compare single-trial EEG variability in those with and without ASD. In addition to decomposing the data with ICA, the EEG epochs were converted to current source density (CSD) models (e.g., Kayser and Tenke, 2006). CSD transforms compute the second spatial derivative of voltage between nearby electrode sites, which enhances local electrical activity while attenuating distal activity. By emphasizing local contributions to the surface map, some of the variability associated with spatial smearing via volume conduction may be reduced. Comparing measures of variability obtained from

the raw channel EEG with measures of variability obtained from two different methods of spatially filtering will illustrate which method produces the least variable data, thus ensuring that the most appropriate source of data is used when comparing variability between the participants with and without ASD.

The result in this paper present a re-analysis of data collected for a previous study (Milne et al., 2009), in which the visual evoked response in children/adolescents with and without ASD was investigated. Previously, we compared amplitude and latency indices of the visual evoked potential (VEP, e.g., the C1 and P1 deflections), and changes in α- and γ-band power associated with presentation of Gabor patches at a range of different spatial frequencies. We found that the time at which spectral power increase following stimulus onset was reduced in the participants with ASD (see also Isler et al., 2010), and that the extent to which the spatial frequency content of the stimuli modulated α- and γ-band power was less in the participants with ASD (see also Jemel et al., 2010). Here, I now investigate intra-participant EEG variability: single-trial variability across the time-course is analyzed by comparing point-by-point amplitude variability across trials, and also by computing the degree of inter-trial phase consistency across the time-course. In addition, variability of P1 amplitude, i.e., the consistency of P1 magnitude, and variability of P1 latency, i.e., the consistency of the time at which the peak occurs will be investigated. The P1 deflection is a positive going deflection within the VEP. It peaks between 100 and 160 ms after stimulus onset, is maximal over posterior leads and is generated within the extra-striate cortex. The P1 was selected for analysis as it is time-locked and phase-locked to stimulus onset. Therefore a "P1-like" increase in amplitude should be visible across single-trials, enabling variability between the two groups to be compared[1].

MATERIALS AND METHOD

PARTICIPANTS

Data were collected from 20 children/adolescents with ASD. Data from two participants were not analyzed as these participants had co-morbid diagnoses of ADHD. Data from another participant were excluded as despite having a clinical diagnosis of ASD, she failed to reach cut-off for ASD on the two screening measures that we used to ensure the homogeneity of our sample (further details are given below). From the remaining 17 participants, data from four participants were excluded as these participants did not show an independent component in their data decomposition that fulfilled the criteria imposed for IC selection (further details are given below). Data are therefore presented from a total of 13 participants with ASD.

Participants in the TD control group were selected from a sample of 37 participants (19 male) who performed this task either as part of another study investigating developmental change of the VEP, or as part of the previously published study (Milne et al., 2009). Data from 12 of these participants is presented here; their inclusion was based on being a suitable match to one of the participants with ASD in terms of gender, age, and IQ and generating data from which an independent component that fulfilled the selection criteria (see below) was identified. The two samples of data were therefore well matched both in terms of participant characteristics and data quality. Participant details are provided in **Table 1**.

The participants with ASD were recruited from a local clinic for children with pervasive developmental disorders. Only participants with a clinically defined diagnosis of autism ($N = 5$), Asperger's syndrome ($N = 5$), or ASD ($N = 3$) were recruited[2], no participant had a co-morbid diagnosis, or a known specific neurological or genetic condition (e.g., Fragile X, Rett Syndrome) that could account for their diagnosis of ASD. The TD children were recruited from an e-mail list of volunteers and were screened for any history of developmental, neurological, or genetic disorder. No participant had taken medication within 24 h of participation. Written informed consent was obtained from the parents of all participants and verbal assent was obtained by the participants prior to testing. The procedures followed were in accordance with the ethical standards of the South Sheffield NHS ethics committee and the Declaration of Helsinki.

PSYCHOMETRIC ASSESSMENT

Degree of ASD symptoms were assessed in all participants by way of an observational measure (the Childhood Autism Rating Scale, CARS, Schopler et al., 1988) and a parental questionnaire

[1]Note, that an index of earlier cortical response to visual stimuli is expressed by the C1 component, which peaks at around 100 ms and is generated in the striate cortex (Di Russo et al., 2001). However, C1 was not analyzed here as not all participants with ASD demonstrated a clear C1 deflection.

[2]In the rest of this manuscript this sample will be collectively referred to as "participants with ASD."

Table 1 | Participant demographics.

	ASD ($N = 13$)	TD ($N = 12$)	*t*- and *p*-Values
Gender	1 Female	1 Female	
CHRONOLOGICAL AGE			
Mean	11 years 8 months	12 years 4 months	$t(23) < 1$, $p = 0.63$
SD	2 years 6 months	2 years 11 months	
Range	8 years 4 months–15 years 5 months	7 years 11 months–16 years 0 months	
FULL SCALE IQ			
Mean	105.9	111.1	$t(23) < 1$, $p = 0.37$
SD	16.3	16.9	
Range	65–134	70–131	
CARS SCORE†			
Mean	31.7		
SD	3.9		
Range	25–39.5		
SCQ SCORE‡			
Mean	23.9		
SD	7.3		
Range	9–34		

CARS, Childhood Autism Rating Scale, SCQ, Social Communication Questionnaire.
†Cut-off for autism is 30.
‡Cut-off for autism is 15.

(the Social Communication Questionnaire, SCQ – lifetime version Berument et al., 1999). The CARS requires the experimenter to rate the participant from a scale of 1–4 on 15 item behavioral rating scale, for example "emotional response" and "fear and nervousness." Scores range from 15 to 60 and the cut-off for ASD is 30. The SCQ consists of 40 "Yes/No" questions asking parents if their child currently displays specific autism-related behaviors or whether those behaviors were present between the ages of 4–5 years. Scores range from 0 to 40 and the cut-off score for ASD is 15. Intellectual ability was measured using the Wechsler Abbreviated Scale of Intelligence (WASI, Wechsler, 1999). This is comprised of four standardized sub-tests that assess expressive language, perceptual organization, abstract verbal reasoning, and non-verbal fluid reasoning abilities. The four sub-tests when considered together yield a "full scale IQ" that provides a composite measure of the participant's general intelligence. All participants had normal, or corrected-to-normal, visual acuity.

STIMULI AND PROCEDURE

Gabor patches were created using Matlab (The Mathworks, Inc.) and the psychophysics toolbox (Brainard, 1997). They were presented on a 17-inch monitor, which refreshed at 75 Hz. Stimuli were centrally presented on a gray background (average luminance = 14.4 cd/m^2). The space-average luminance of each grating was 16.3 cd/m^2, and the Michelson contrast, defined by (Lmax − Lmin)/(Lmax + Lmin) was 68%. The slight difference between the average luminance of the background and the stimuli was not visibly apparent and did not lead to any visible edges around the stimuli. At a viewing distance of 114 cm the patches subtended 6.78° by 6.78° of visual angle. All patches were presented in diagonal (45°), orientation, had a gaussian envelope with SD of 0.68°, and with spatial frequency modulation of 0.5, 1, 4, or 8 cycles/degree. An additional stimulus, a gray-scaled image of a zebra was presented. Participants were instructed to respond by pressing a response button with the index finger of their dominant hand as quickly as possible whenever they saw the zebra.

Each of the four Gabor patches was shown 72 times, the zebra was shown 36 times. The order of stimulus presentation was randomized. Each stimulus remained on screen for 39 refresh cycles (507 ms), with an additional variable inter-stimulus interval of either 24 (312 ms), 39 (507 ms), or 54 (702 ms) refresh cycles. A white fixation cross measuring 0.2° by 0.2° remained in the center of the screen for the duration of the task. Participants were asked to maintain fixation and to limit their blink frequency during the experiment.

DATA RECORDING

EEG was continuously recorded using a high-density array of 128 Ag/AgCl electrodes (Electrical Geodesics Inc., Tucker, 1993). Impedance was kept below 50 kΩ. The signal was amplified (×1000), filtered on line with a band-pass of 0.01–80 Hz, then digitized at a sampling rate of 1 kHz. The electro-oculogram (EOG) was recorded from bipolar electrode pairs located at the outer canthi and above and below the left and right eyes. Data were referenced to the vertex electrode, and were stored on the hard drive of a G4 Macintosh power PC.

DATA PRE-PROCESSING

Data were analyzed off-line using EEGLAB (Delorme and Makeig, 2004, http://www.sccn.ucsd.edu/eeglab), and the CSD toolbox (Kayser and Tenke, 2006; Kayser, 2009) running under Matlab 7.4 (The Mathworks, Inc.). A number of pre-processing steps were performed on the data before applying either ICA or CSD interpolation. First, the data were high-pass filtered (1 Hz) to minimize drift. Then the number of channels was pruned from 128 to 64. Pruning was necessary in order to improve the quality of ICA decomposition, given the relatively small amount of data recorded (~5 min). Initially channels that showed noise artifacts due to poor connection to the scalp were deleted, then channels were removed if they showed high kurtosis, finally, additional channels that showed the smallest inter-electrode distance were removed until 64 relatively evenly spaced electrodes remained. In some cases a small number of additional channels were deleted by the experimenter if any noise artifacts on any particular channel were still visible. (Given the small amount of data recorded there was a bias toward rejecting channels rather than portions of data in order to facilitate ICA.)

After channel pruning, noisy segments of data, i.e., segments that contained gross artifacts such as muscle twitching or swallowing were removed from the data. At this stage, the data were (i) decomposed by extended infomax ICA using the function *runica*, as implemented in EEGLAB, and then segmented into epochs associated with presentation of the high spatial frequency Gabor patch (8 cycles/degree); epochs were 800 ms long (−100 ms pre-stimulus), and baseline corrected by subtracting the mean of the 100 ms pre-stimulus interval, and (ii) segmented into epochs as described above and then transformed into CSD estimates (measured in μV/cm^2) using a spherical spline surface Laplacian (Perrin et al., 1989) as implemented by Kayser and Tenke (2006), Kayser (2009), with a spline-smoothing coefficient (λ) of 1.0^{-5}. Both sets of data were then low-pass filtered (<30 Hz).

Independent component source locations were estimated by creating an equivalent current dipole model for each component using the *dipfit* function from EEGLAB, this function estimates dipole location by applying inverse source modeling methods to a standard boundary element head model. The coordinate system used to specify electrode locations of both sets of data is publicly available as a look-up table ["10-5-System Mastoids EGI129.csd"] via the CSD toolbox. Three types of data were analyzed: (i) the back-projected independent time-course of the selected component (see below), (ii) the time-course of the CSD interpolated data from the selected channel (see below), and (iii) the time-course of the raw, spatially un-filtered data from the selected channel.

DATA SELECTION

IC selection

Only components whose scalp maps had <15% residual variance from the best-fitting forward model scalp projection were considered for further analysis. Any remaining components that reflected muscle activity, electrocardiogram, or eye-movements, on the basis of their dipole location, spectra, and scalp maps, were considered artifacts and excluded from further analysis. From the remaining components, selection of the component that represented activity associated with visual perception was based on two criteria: (i) the back-projected time-course was required to show time-locked

increase in amplitude between 100 and 170 ms after stimulus onset; and (ii) the estimated equivalent current dipole was required to be located in posterior cortex. In all cases the selected component accounted for a larger percentage of the variance of the total EEG than any other component between 100 and 170 ms. The mean percentage variance accounted for by the selected components in this time-window was 70%, indicating that the selected components contributed substantially to the VEP. The mean residual variance of the dipole fit for the selected independent components was reasonably low (6.5%), suggesting that the selected ICs reflect the activity of a compact region in the cortex. The estimated Talairach coordinates of the equivalent current dipole of each component suggested that the neural generators of the component activations were located in extra-striate cortex.

Channel selection

Channel selection was based on an optimized electrode approach (e.g., Foxe and Simpson, 2002), whereby the channel that showed the highest P1 amplitude, from the CSD interpolated data, was selected for analysis. The CSD scalp maps at the latency at which the P1 peak was maximal, with the selected electrode marked by a black dot are presented in **Figure 1** (left column); scalp maps from the IC selected for each participant are presented in **Figure 1** (right column). Despite not being used in the IC selection criteria, the similarity between the weight with which the ICs project to the scalp electrodes and the topography of the CSD maps at the time-point when P1 was maximal, confirms that the selected component represent activity associated with the P1 deflection.

DEFINING THE VARIABLES OF INTEREST

P1 amplitude and latency

A number of variables were calculated from the three sources of data. In line with conventional ERP analyses, P1 amplitude was given by the highest (peak) amplitude between 100 and 170 ms; P1 latency was given by the time of this peak from stimulus onset. For the single-trial analyses, the highest amplitude between 100 and 170 ms was identified in each trial; P1 amplitude was calculated as the median amplitude value, P1 latency was calculated as the median time at which these single-trial peaks occurred.

Variability

Variability in the expression of the P1 amplitude was estimated by calculating the median absolute deviation (MAD) of the P1 amplitude and latency values, normalized by dividing by the median. In addition, across-trial amplitude variability was calculated at each time-point in the epoch. Due to the inherent difficulty of calculating co-efficients of variation when the central tendency is close to 0, amplitude values were normalized prior to computing the MAD estimate by converting all of the data to a *z*-score.

The degree to which each participant showed phase coherence over trials was calculated by computing an inter-trial phase coherence (ITPC) measure. ITPC is a frequency domain measure that indicates the degree to which EEG activity is phase-locked, across trials, to specific experimental events – in this case presentation of the 8 cycles/degree Gabor patch. The power spectrum of each trial was calculated using a hanning tapered fast Fourier transform with a window size of 128 ms. α-Band power was dominant in the signal following stimulus presentation, therefore inter-trial α phase coherence, specifically at 11.7 Hz, was calculated using the *newtimef* function in EEGLAB. The highest ITPC value within the 100–170 ms post stimulus onset time-window was extracted for each participant to give an indication of the maximum level of ITPC, a high level (up to 1) indicates strong phase coherence, i.e., low variability; a low level (down to 0) indicates weak phase coherence, i.e., high variability. In addition, ITPC was calculated across the time-series.

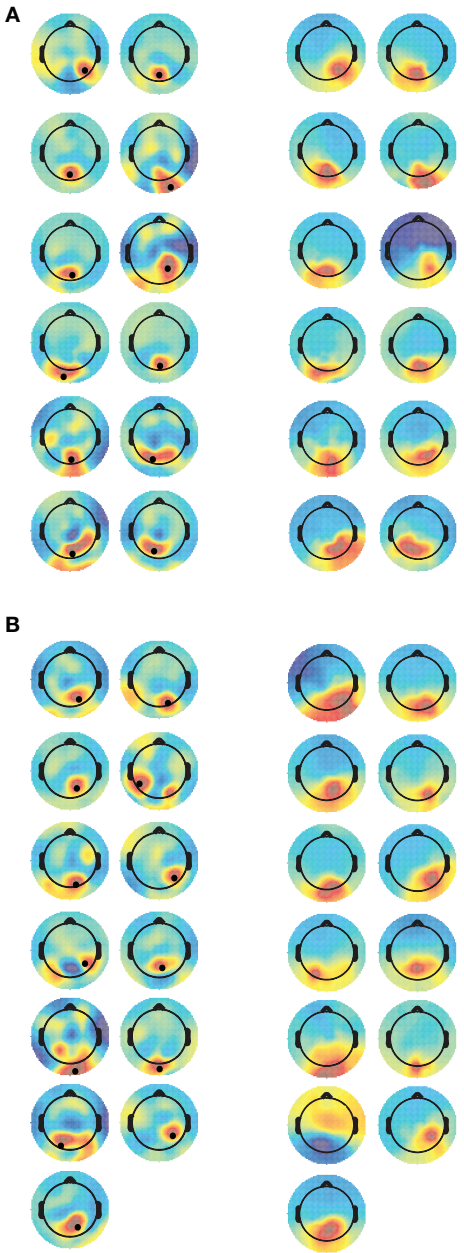


FIGURE 1 | Current source density (CSD) and IC scalp maps. CSD maps (left panel) are plotted at the time when P1 amplitude was highest. The electrode selected for analysis is indicated by the black dot. IC scalp maps (right panel) are stationary, and therefore constant across the time-course. [**(A)** = TD, **(B)** = ASD].

In sum, three single variables were extracted for analysis – variability of P1 amplitude, variability of P1 latency and maximum ITPC. The amplitude and ITPC time-courses were also compared on a point-by-point basis.

RESULTS

A positive going deflection was seen in the single-trials from all three sources of data. This is illustrated in **Figure 2**, which shows the single-trial amplitude across the epoch from a TD participant and a participant with ASD as measured from each source of data.

WHICH SOURCE OF DATA SHOWS THE HIGHEST VARIABILITY?

The three measures of peak variability (see above), from all subjects, were entered into separate one-way ANOVAs with a within-subject factor of data source (raw EEG, CSD interpolated data, and back-projected IC). All three ANOVAs showed a significant effect of data source (all $F > 3.4$, all $p < 0.05$) and were therefore followed-up with paired-samples t-tests. These analyses indicated that there was no difference in variability when measured from the CSD interpolated data and the back-projected IC data (all $t < 1$, all $p > 0.1$), but the variables computed from the raw channels were significantly more variable than the variables computed from the spatially filtered sources. The data are presented in **Table 2**. Based on these results, the following analyses of group differences in variability were performed using the spatially filtered data rather than the raw EEG.

GROUP COMPARISON: INTRA-PARTICIPANT VARIABILITY

The extent to which the participants with ASD showed increased variability when compared with the neuro-typical participants was investigated with three 2 × 2 repeated measures ANOVAs with a within-subject factor of data source (CSD or ICA) and a between subjects factor of group (ASD or TD). The results of these analyses

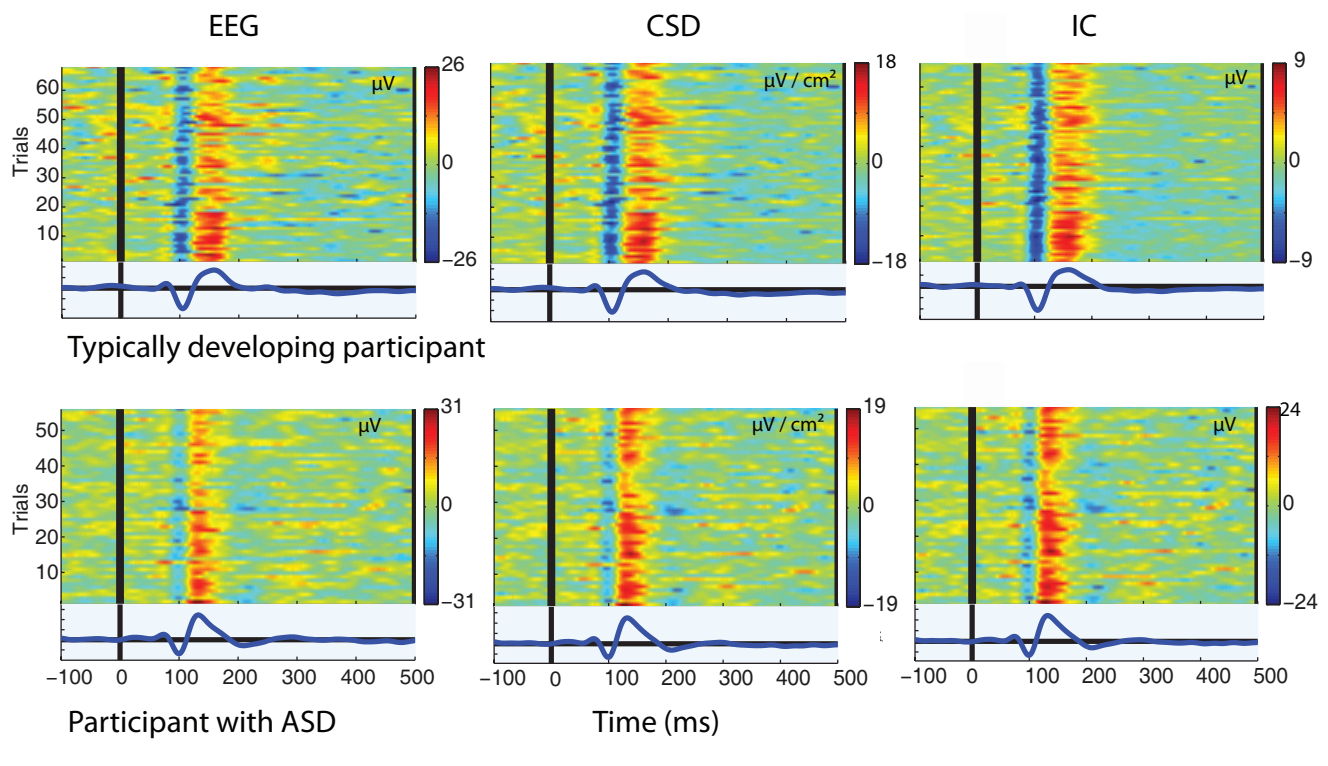


FIGURE 2 | Single-trial amplitude from the raw EEG, CSD interpolated data, and back-projected IC data. Each colored trace on the *y*-axis indicates amplitude across the time-course of a single-trial (red = +ve, blue = –ve as shown on colormap). The blue trace below each figure is the average of the single-trials.

Table 2 | Mean measures of variability as calculated from the three sources of data.

	Mean value per data source			Statistical comparison	Statistical comparison
	EEG	CSD	ICA	EEG vs CSD	EEG vs ICA
P1 amplitude variability	0.38 [0.34 0.42]	0.34 [0.30 0.37]	0.34 [0.29 0.39]	$t = 3.85$, $p = 0.001$ [0.022 0.072]	$t = 1.89$, $p = 0.072$ [0.004 0.083]
P1 latency variability	0.07 [0.05 0.08]	0.06 [0.04 0.07]	0.06 [0.04 0.07]	$t = 2.98$, $p = 0.007$ [0.004 0.020]	$t = 2.58$, $p = 0.016$ [0.002 0.022]
Max. α-band ITPC	0.77 [0.70 0.84]	0.81 [0.74 0.87]	0.81 [0.74 0.87]	$t = 4.4$, $p < 0.001$ [0.019 0.053]	$t = 2.1$, $p = 0.046$ [0.001 0.072]

95% confidence for the variable mean and for the difference between the two variables are given in square brackets.

are shown in **Table 3**, and the data are plotted in **Figure 3A**. The participants with ASD showed, on average, greater variance than the neuro-typical participants. This was the case for both measures of P1 variability: amplitude variability, ASD mean = 0.38 [0.33 0.43], TD mean = 0.30 [0.25 0.35]; latency variability, ASD mean = 0.07 [0.05 0.08], TD mean = 0.04 [0.03 0.06]; and also for the maximum α-band ITPC, ASD mean = 0.74 [0.66 0.83], TD mean = 0.87 [0.78 0.96].

Figure 3B illustrates amplitude variability in the two groups of participants at each time-point. **Figure 3C** illustrates α-band ITPC in the two groups of participants at each time-point. In order to compare these group time-courses, independent-samples *t*-tests were computed at each time-point. ITPC was significantly reduced ($p < 0.05$) in the participants with ASD between 160 and 271 ms after stimulus onset. Amplitude variability was significantly greater ($p < 0.05$) in the participants with ASD between 97 and 114 ms, between 185 and 194 ms, and between 283 and 300 ms after stimulus onset.

GROUP COMPARISON – P1 AMPLITUDE AND LATENCY

Although not the main area of interest in this article, P1 latency and amplitude were compared between groups as this is the first study to report P1 amplitude calculated from single-trials rather than the average ERP in individuals with ASD. It has been suggested that individuals with ASD may show hyper-sensitivity in posterior cortical areas associated with early visual perception (Mottron et al., 2006). As such, significantly greater P1 amplitude could be predicted[3]. Given that individuals with ASD show increased latency jitter compared to TD participants (see above), the amplitude of the ERP could be artificially reduced in this population, masking any such effect. By contrast, computing P1 from the single-trials creates a variable that is immune to the effects of latency jitter. However, there was no group difference in P1 amplitude from any data source when calculated from either the ERP peak or the median of single-trials (all $t < 1$. all $p > 0.4$). **Figures 3D,E** illustrate P1 amplitude when calculated as the peak of the ERP (**Figure 3D**) and the median amplitude of the single-trial peaks (**Figure 3E**). In addition these figures show variability as the SD of the single-trial peaks (3D) and the MAD of the single-trial peaks (3E). For comparison with previously published data, data from the raw EEG is presented in order to illustrate the difference in amplitude when calculated from the ERP and from the median of the single-trials.

Latency to peak was compared in the two groups using a repeated measures ANOVA with three levels: data source (EEG, CSD, or ICA), data type (ERP or median of single-trials) and group (ASD or TD). The only significant effect was the main effect of group, $F(1,23) = 9.4$, $p < 0.01$ which reflected the fact that latency to peak was faster, on average 18 ms faster, in the ASD group than the TD group, regardless of the data source or the data type. **Table 4** presents P1 amplitude and latency values in each group as calculated from the different sources of data. The table illustrates that, as expected, when calculated from the median of single-trials rather than the peak of the ERP, P1 amplitude is larger, whereas the latency of the peak does not differ.

Under some circumstances, ERP amplitude has been shown to decrease over multiple trials (e.g., Maurer et al., 2008). Therefore, in order to evaluate whether increased variability in the participants with ASD may reflect a larger habituation-effect, linear regressions of trial number against (IC) peak amplitude were computed for each participant. Mean slope size did not differ between groups however, indicating that this was not the case (TD mean β = –0.04 [–0.13 0.05], ASD mean β = –0.06 [–0.15.03], $t(23) < 1$, $p > 0.1$.

DATA INTEGRITY

Differences in data integrity between the two groups could conflate estimates of variability, therefore a number of analyses were carried out to ensure that data from both groups were of similar quality. The mean number of epochs analyzed did not differ between the two groups, TD = 58.6 [52 65], ASD = 53.3 [47 60], $t(23) = -1.2$, $p = 0.241$, suggesting that any differences in variability between the two groups are unlikely to arise from unequal number of epochs being submitted to analysis. There was no significant difference in the percentage variance accounted for by the selected ICs within the P1 time-window, TD = 70.7% [59.8 81.6], ASD group = 69% [58.6 79.5], $t(23) < 1$, $p = 0.822$. Furthermore, there was no significant group difference in residual variability of the dipole fit, TD = 7.1% [5.7 8.6], ASD = 5.9% [4.5 7.3], $t(23) = -1.28$, $p = 0.213$), suggesting that the quality of the ICA decomposition was similar between the two groups.

CORRELATION WITH REACTION TIME

The experiment from which these data are taken required participants to press a response key whenever they saw an oddball stimulus – a zebra. The groups did not differ in the number of responses made, TD = 34.3 [33.0 36.9], ASD = 34.5 [32.2 36.3], $t(20) < 1$, $p = 0.657$, in mean reaction time TD = 393.3 [363 413], ASD = 412.4 [378 431], $t(20) = 1.37$, $p = 0.307$, nor in the mean MAD estimate of reaction time variability, TD = 45.1 [34.4 52.5], ASD = 38.2 [28.2 48.1], $t(20) < 1$, $p = 0.423$. However, there was a significant relationship between reaction time variability and peak amplitude variability, $r_s(22) = 0.479$, $p = 0.024$. (Behavioral data from three participants with ASD was unavailable due to a technical fault.)

[3]Although, as discussed in the introduction, the late P1 complex analyzed here most likely reflects activity in extra-striate cortex rather than the primary visual cortex.

Table 3 | Results of statistical tests comparing measures of variability in the two groups of participants.

	Data source	Group	Interaction	95% CI of the group difference
P1 amplitude variability	$F < 1$, $p = 0.792$	$F = 5.37$, $p = 0.030$	$F = 3.83$, $p = 0.063$	[0.008 0.148]
P1 latency variability	$F < 1$, $p = 0.975$	$F = 5.40$, $p = 0.029$	$F < 1$, $p = 0.747$	[0.003 0.047]
Max. α-band ITPC	$F < 1$, $p = 0.977$	$F = 4.67$, $p = 0.041$	$F < 1$, $p = 0.418$	[0.006 0.256]

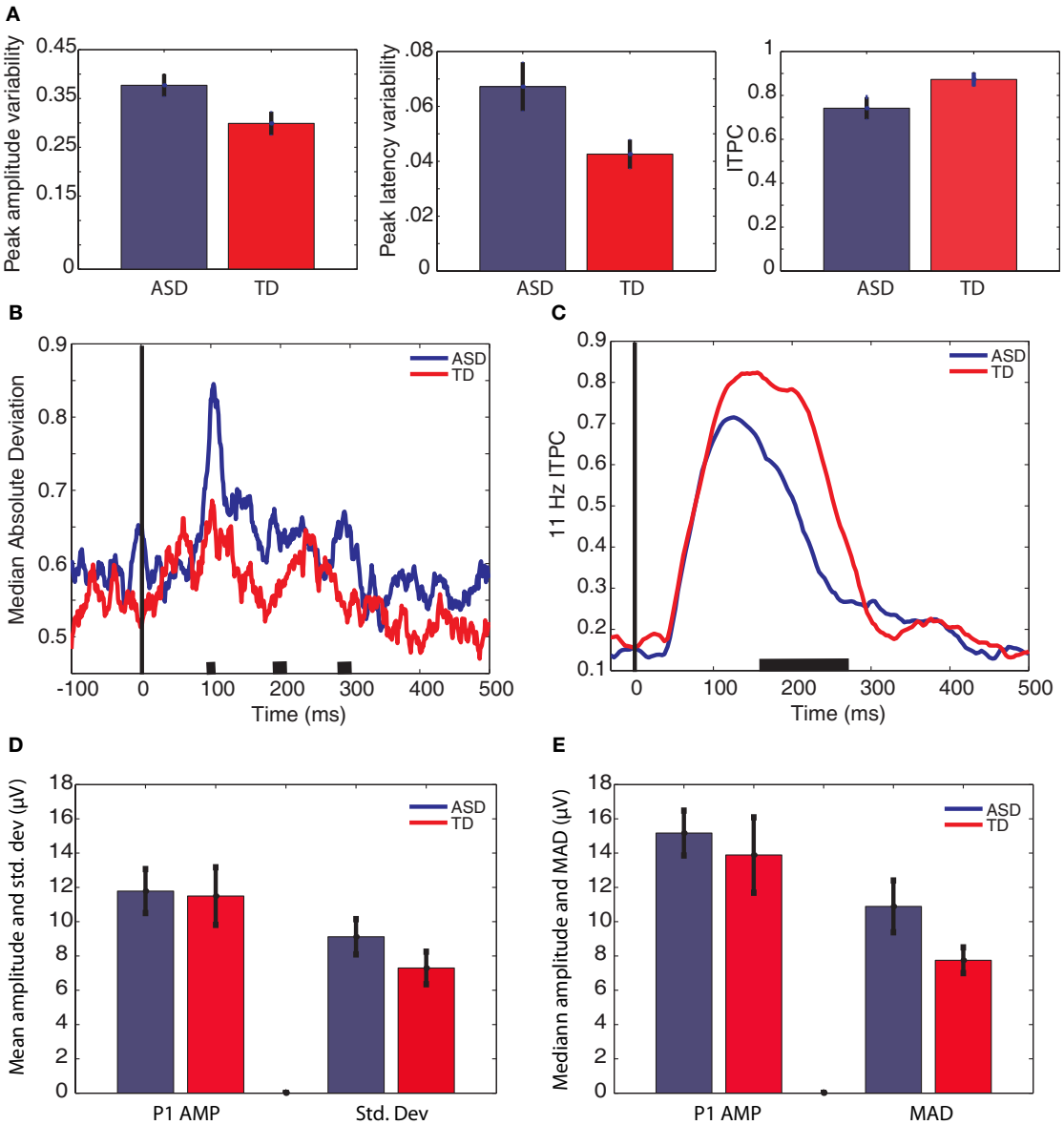


FIGURE 3 | Estimates of variability averaged across group. [**(A)** (normalized)] measures of peak variability. The left graph shows mean variation in the amplitude of the P1 peak, the middle graph shows mean variation in the latency of the P1 peak, and the right graph shows the mean maximum ITPC between 100 and 170 ms. [**(B)** (normalized)] median absolute deviation in amplitude across trials at each time-point, averaged across participant groups. **(C)** ITPC at each time-point, averaged across participant groups. The black lines on the *x*-axis of plots **(B)** and **(C)** indicate time-points of group difference ($p < 0.05$). P1 amplitude is shown in plots **(D)** and **(E)**. **(D)** P1 amplitude calculated from the ERP peak, and the SD of the single-trial P1 peaks. **(E)** P1 amplitude calculated as the median of the single-trial P1 peaks, and the median absolute deviation of the single-trial P1 peaks. Bars represent ±1 SE.

DISCUSSION

The aim of this work was to establish whether those with ASD show greater variability across single-trial evoked EEG compared with neuro-typical individuals. A second aim was to compare single-trial EEG variability when extracted from spatially filtered data and from raw-scalp EEG data in order to select the most appropriate variables for group comparison. All three measures of peak variability – P1 amplitude, P1 latency, and maximum α-band phase coherence – were smaller when analyzed from the spatially filtered data than from the scalp EEG data, highlighting the benefits of applying spatial filtering techniques to EEG. Having validated the use of CSD and ICA in this study, measures of single-trial variability were compared between the participants with and without ASD, with the finding that intra-participant variability was significantly greater in the participants with ASD than in the control group.

These data suggest that previous reports of increased response time variability in those with ASD (Geurts et al., 2008) may be underpinned by variability within cortical dynamics associated with the ability to

Table 4 | Mean P1 amplitude and latency in the two groups of participants.

	ASD	TD
P1 AMPLITUDE		
Back-projected IC data (μV)		
ERP Peak	3.17 [2.4 3.9]	3.06 [2.3 3.8]
Median of single-trial peak	3.86 [3.1 4.7]	3.62 [2.8 4.5]
CSD interpolated data (μV/cm^2)		
ERP Peak	7.34 [5.3 9.4]	7.62 [5.5 9.8]
Median of single-trial peak	8.97 [6.7 11.2]	8.92 [6.5 11.3]
Raw EEG data (μV)		
ERP Peak	11.79 [8.9 14.8]	11.49 [8.4 14.6]
Median of single-trial peak	15.17 [11.6 18.7]	13.89 [10.2 17.6]
P1 LATENCY		
Back-projected IC data (ms)		
ERP Peak	134 [124 144]	154 [143 163]
Median of single-trial peak	134 [125 142]	152 [142 161]
CSD interpolated data (ms)		
ERP Peak	132 [122 141]	153 [142 162]
Median of single-trial peak	133 [124 141]	152 [143 161]
Raw EEG data (ms)		
ERP Peak	135 [126 144]	154 [144 162]
Median of single-trial peak	136 [127 143]	153 [144 161]

synchronize the activity of stimulus-related cell assembly(ies) consistently across trials. The experimental paradigm used here did not elicit a significant difference between response time variability in those with and without ASD. This may be because only a small number of trials (36) were available to ascertain variability, in contrast to Geurts et al. (2008) who used nearly twice as many (64 trials), or it may be because of the small group sizes and consequently reduced power of the analyses performed here. Nevertheless, there was a significant relationship between response time variability and P1 amplitude variability. Note that P1 and the behavioral data were extracted from separate trials (the P1 was extracted from trials in which the eliciting stimulus was a Gabor patch with a spatial frequency content of 8 cycles/degree and the behavioral data were derived from trials in which the eliciting stimulus was a zebra), therefore it is not the case that specific trial-by-trial variations in ERP amplitude are driving the variability in response time, rather it appears as though a common mechanism may underpin both behavioral variability and ERP amplitude variability.

As described in the Introduction, neurocortical dynamics result from the activation of partially distinct and interacting cell assemblies; the mechanism of communication within these cell assemblies is synchronous oscillations. A number of authors have suggested that ASD may be characterized by reduced neural synchrony, especially of high-frequency (γ-band) oscillations (e.g., see Brock et al., 2002), although evidence to support this position is mixed. While some studies have shown lower levels of evoked γ-band power in those with ASD (Wilson et al., 2007), more recent data indicates that while evoked γ-band power may be reduced in those with ASD, induced γ-band power is increased, and inter-trial γ-band phase coherence (ITPC) is reduced (Rojas et al., 2008). The concept of "evoked" or "induced" EEG is defined by whether or not single-trial activity is time- and phase-locked to a stimulus (evoked activity),or whether it is perturbed by the stimulus, but neither time- nor phase-locked to it (induced activity). However, as the data presented above, and numerous other estimations of ITPC (e.g., Tallon-Baudry and Bertrand, 1999) illustrate, complete phase-locking across trials (i.e., ITPC = 1) is physiologically unrealistic. Therefore the boundary for defining whether stimulus-related activity should be considered to be evoked or induced is unclear. Rojas et al. (2008) point out that their data fit a model in which total (evoked + induced) stimulus-related γ-band power is equivalent in the participants with and without ASD, and that reduced inter-trial phase consistency, computationally, leads to a reduction of what is classed as evoked activity and an increase in what is classed as induced activity. Thus these authors conclude that the production of γ-band oscillations in response to external stimulation is no different in those with and without ASD, rather their data point toward dysfunction in the timing of γ-band oscillations in the participants with ASD.

The data reported here provide evidence of reduced ITPC in the α-band in ASD. Together with the result of Rojas et al. (2008), these data indicate widespread dysfunction of neural timing in ASD, rather than a specific deficit of high-frequency γ-band oscillations as some authors have predicted. Reduced ITPC in ASD indicates that those with ASD are less able to synchronize the activity of stimulus-related cell assembly(ies) consistently across trials, and provide evidence for temporal disruptions in the organization and recruitment of cell assemblies. It is not clear whether this temporal disruption underpins, is caused by, or is unrelated to, postulated neural de-synchrony in ASD.

A number of possible etiologies of atypical neural oscillations in ASD have been suggested, including: a surfeit in local connectivity – especially in primary sensory areas (Belmonte et al., 2004); smaller and more dispersed cortical mini-columns leading to a reduction in inhibitory inter-neuronal activity (Casanova et al., 2002); an imbalance of cortical excitation and inhibition due to increased glutamergic/reduced GABAergic signaling (Rubenstein and Merzenich, 2003); and impairment in the inferior olive – a structure that that mediates electrical synapses and that drives neural synchrony, and has been found to be structurally atypical in some individuals with ASD (Welsh et al., 2005). No theory has yet linked any of these putative impairments with increased intra-participant variability in those with ASD. However within the literature on ADHD, intra-participant variability has been theoretically linked with inconsistent and inefficient neuronal transmission, which may arise from impairment in astrocytes, a type of glial cell that plays a critical role in fueling neuronal oscillations (Russell et al., 2006). Astrocyte impairment in ASD could therefore account for a range of features of ASD including neural de-synchrony, EEG single-trial variability, and behavioral (response time) variability. Given the important role of glia in synapse formation and maintenance (Bolton and Eroglu, 2009) the suggestion that astrocyte impairment may be a critical factor in ASD is a tantalizing one. It is important to note however that, in addition to the theoretically formulated suggestions described above, a variety of neuronal characteristics (e.g., synaptic transmission, channel gating, fluctuation in transmitter release, postsynaptic receptor activation, ion concentrations, membrane conductance) may contribute to variability of evoked EEG response (Sannita, 2006), therefore it is not currently possible to identify the precise source(s) of EEG variability.

Regardless of the precise source, increased EEG variability in those with ASD is evidence of increased intrinsic neural "noise" (Sannita, 2006). Increased neural noise in ASD has been predicted by a number of authors (see Simmons et al., 2009), however, the data reported here represent the first empirical demonstration of increased neural noise in ASD. Increased neural noise has the potential to influence behavior in a variety of ways, and its impact on different levels of function, e.g., perception, cognition, and behavior, may not be consistent. Whilst an increased noise-to-signal ratio leads to reduced perceptual sensitivity in many cases, one type of noise – stochastic resonance – can amplify a signal, leading to increased sensitivity. Increased levels of neural noise have therefore been discussed in relation to atypical perception in ASD, and offered as a parsimonious explanation of data in which those with ASD show both hyper- and hypo-reactivity to perceptual stimuli and enhanced and impaired perceptual sensitivity measured with psychophysical tasks (Simmons et al., 2009).

Increased neural noise is less likely to have an advantageous effect on cognitive task performance however, as it may lead to a number of sub-optimal outcomes including a general decrease in response times and greater response time variability, more errors in tasks with more than one possible response, and the need for increased repetitions of a task to achieve the same outcome as those with lower levels of noise. Furthermore, noise-related reduction in task performance would be evidenced by impairments across many domains and tasks, rather than in isolated tasks, and it would also lead to increased inter-participant variability. This description of data is very similar to that represented by the literature on cognitive function in ASD. Increased neural noise is therefore a plausible, and parsimonious, explanation both for the array of cognitive tasks in which participants with ASD have been shown to perform more poorly than those without ASD, and for the significant inter-individual variability present in those with ASD. In support of this position are two demonstrations where reduced task performance can be accounted for by what may be termed "noise." For example, thresholds for detecting coherent motion can be artificially inflated by transient lapses of attention (McAnally et al., 2001), and intra-individual response variability is a strong predictor of success in the Go No-Go task (Bellgrove et al., 2004), suggesting that lower sensitivity to coherent motion and failure to inhibit prepotent responses, both of which have been reported in those with ASD (see Ozonoff et al., 1994; Milne et al., 2002 respectively), may arise due to increased neural noise rather than reflecting a specific impairment in either motion perception or in response inhibition, as is the current interpretation of these data (see also Baron-Cohen and Belmonte, 2005 for a similar argument).

Some authors have suggested that cortical hyper-excitability in ASD may be restricted to/more pronounced in, primary sensory areas (e.g., Rubenstein and Merzenich, 2003; Mottron et al., 2006). Therefore in order to evaluate these results in light of current theories it is necessary to consider where the neural generators of the P1 deflection analyzed here might be. The location of the electrodes selected from the CSD data, and the estimated location of the equivalent current dipole of the ICs suggests that the P1 analyzed here is generated in extra-striate cortex. This is commensurate with a number of papers that have localized the neural generators of the P1 deflection to the extra-striate cortex (e.g., Di Russo et al., 2001; Ales et al., 2010). Therefore, these data provide evidence for variability in extra-striate cortex rather than primary visual cortex. As noted in the *Introduction* the earlier C1 deflection would be a more appropriate deflection with which to investigate variability in primary visual cortex, however this was not analyzed here as a number of participants with ASD failed to show a clear C1 deflection, either in the (averaged) ERP or in the single-trials. The specific reason for this is unclear, but is being addressed by on-going studies by our research group. Therefore, although the P1 deflection reported here does not tap the earliest stage visual processing, it was the most robust early deflection in the data, making it the best available candidate to investigate increased noise in the visual cortex in ASD. Future studies are required to establish whether areas of primary sensory cortex, generating earlier ERP deflections, show similar, or possibly greater, levels of variability in participants with ASD.

In addition to increased variability, the P1 peak in the participants with ASD occurred significantly sooner than in the control group. This finding, from these data, was reported previously (Milne et al., 2009), so will not be dwelt on here. Nevertheless, reduced latency to peak is commensurate with the suggestion of local hyper-connectivity in ASD which has been predicted by some to lead to increased cortical noise (see for example, Belmonte et al., 2004).

There are a number of implications from this work for existing EEG studies in those with ASD. Consistent with previous reports, these data provide no evidence for difference in P1 amplitude in participants with ASD. Although there is one report of reduced P1 amplitude in children/adolescents with pervasive developmental disorder, including ASD (Hoeksma et al., 2004), this may have been due to latency jitter, as ERP amplitude (when calculated from the peak of the averaged single-trials) is intrinsically related to latency variability. Conversely, the suggestion that individuals with ASD may have hyper-responsive visual cortices would predict increased P1 amplitude in those with ASD, and increased latency jitter may mask this potential outcome. However, the data did not support this prediction, as when P1 amplitude was calculated as the median of the single-trial peaks, there was still no group difference in P1 amplitude. The data reported here indicate that establishing degree of latency jitter within each participant is possible, and should be an essential part of ERP analysis if conclusions are to be drawn to regarding the origin of observed group differences. Before leaving this point, it is important to point out that a number of physiological factors contribute to ERP amplitude. Although ICA was able to isolate the signal associated with perceptual encoding from the total EEG, and therefore facilitate comparison of within-participant variables such as variability, it cannot address the potentially confounding factors of individual differences that may lead to differences in ERP amplitude between groups including differences in cortical convolution, position of the calcarine sulcus, and/or conductivity of underlying tissue, etc. Given that there is some evidence of cortical folding abnormalities in children with ASD (Nordahl et al., 2007), direct comparison of EEG amplitude or EEG power between experimental groups, without first normalizing the data, may not be a reliable technique.

When analyzing the data presented here, significant attempts were made to minimize potential confounds that could artificially inflate the estimates of variability in one group or another. Note that the within-subject estimates of variability reported here are normalized, thus validating group comparisons; note also that the two groups of participants were well matched as regards to

age, IQ, and gender, and that the data from the two groups was matched in terms of quality. It is therefore unlikely that methodological factors contributed to the group differences reported here. However, as is outlined in the methodology, data is not reported from all participants who participated in the study. This was driven primarily by the goal of ensuring well matched samples, but was also a consequence of the fact that not all participants generated an IC that was considered reliable enough to be analyzed. The most likely reason for this is that only a small amount of data (approximately 51/2 min of data per participant) were recorded and available for ICA decomposition. The experiment was necessarily short given the age of the participants, but the quality of ICA decomposition would be greatly improved with longer recordings, therefore future work should aim to replicate the findings reported here with larger groups of participants and with longer data recordings.

Behavioral variability is not unique to those with ASD (Castellanos et al., 2005), therefore future research is required to establish the universality of increased EEG variability in ASD and in other developmental disorders (such as ADHD), and to establish whether increased variability is a general characteristic of brain pathology, or whether distinctive features of variability occur in different developmental disorders. Furthermore it is necessary to establish the extent to which increased EEG variability is an enduring endophenotype of ASD, or whether it related either to external factors such as context or particular task requirements, or to internal factors such as cognitive state (e.g., awake, asleep, tired, alert). The presence of a significant correlation between EEG variability and response time variability provides preliminary evidence that response time variability and EEG variability are related, albeit in a small sample of participants. This relationship should be tested more rigorously in future studies in which larger groups of participants are tested and different types of behavioral response tasks (such as simple reaction time, choice reaction time, response inhibition, etc.) are performed. In addition, more detailed single-trial analyses should be performed in order to examine the temporally dynamic patterns of EEG fluctuations, and the relationship between EEG variability and cognitive task performance and both inter- and intra-participant variability needs to be clarified.

Note that the mean P1 peak amplitude variability, measured as the MAD estimate of the P1 peak amplitude from the raw EEG data in the TD group, was 0.37. This broadly concurs with existing data in which the coefficient of variation of VEP amplitude recorded in 100 healthy adults from electrodes positioned above the occipital cortex was reported to be 0.41 (Klistorner and Graham, 2001). However, given that the MAD estimator is less influenced by outlying data points than the coefficient of variation, estimates of variability from this statistic tend to be lower than from the coefficient of variation, so a direct comparison between these two statistics cannot be made. For comparison, the co-efficient of variation of P1 amplitude in the TD group, based on the SD of these data was 0.58, i.e., higher than that reported by Klistorner and Graham (2001) – possibly reflecting developmental change in amplitude variability.

CONCLUSION

In conclusion, these data illustrate that analysis of single-trial EEG activity is less variable when the data are spatially filtered, either using ICA or CSD, prior to analysis. Therefore, when comparing data between different groups of participants, more accurate results are likely to be obtained if indices obtained from ICs of CSD interpolated EEG rather than raw channel indices are compared. Further to this, EEG variability across single-trials was significantly greater in the participants with ASD as compared to the TD control group. These data provide the first empirical demonstration of increased neural noise in those with ASD. Increased variability in neural activity may result in a number of negative consequences for individuals with ASD and may contribute to the substantial inter-individual variability that characterizes the literature on cognitive function in those with ASD.

ACKNOWLEDGMENTS

I would like to thank all the participants and their families, who generously gave up their time to participate in this research. I am indebted to Scott Makeig, Arnaud Delorme, and others at SCCN for numerous, and invaluable, discussions about ICA. I would also like to thank Alison Scope for collecting the data presented here, Suzanna Laycock for stimulating discussions about neural noise and developmental disorders, and Mike X. Cohen for introducing me to CSD interpolation.

REFERENCES

Ales, J. M., Yates, J. L., and Norcia, A. M. (2010). V1 is not uniquely identified by polarity reversals of responses to upper and lower visual field stimuli. *Neuroimage* 52, 1401–1409.

Allen, G., and Courchesne, E. (2001). Attention function and dysfunction in autism. *Front. Biosci.* 16, D105–D119.

APA. (1994). *Diagnostic and Statistical Manual of Mental Disorders*. Washington, DC: American Psychiatric Press.

Arieli, A., Sterkin, A., Grinvald, A., and Aertsen, A. (1996). Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. *Science* 273, 1868–1871.

Baird, G., Simonoff, E., Pickles, A., Chandler, S., Loucas, T., Meldrum, D., and Charman, T. (2006). Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the special needs and autism project (SNAP). *Lancet* 368, 210–215.

Baron-Cohen, S., and Belmonte, M. K. (2005). Autism: a window onto the development of the social and the analytic brain. *Annu. Rev. Neurosci.* 28, 109–126.

Bellgrove, M. A., Hester, R., and Garavan, H. (2004). The functional neuroanatomical correlates of response variability: evidence from a response inhibition task. *Neuropsychologia* 42, 1910–1916.

Belmonte, M. K., Cook, E. H., Anderson, G. M., Rubenstein, J. L., Greenough, W. T, Beckel-Mitchener, A., Courchesne, E., Boulanger, L. M., Powell, S. B., Levitt, P. R., Perry, E. K., Jiang, Y. H., DeLorey, T. M., and Tierney, E. (2004). Autism as a disorder of neural information processing: directions for research and targets for therapy. *Mol. Psychiatry* 9, 646–663.

Bennetto, L., Pennington, B. F., and Rogers, S. J. (1996). Intact and impaired memory functions in autism. *Child Dev.* 67, 1816–1835.

Bertone, A., Mottron, L., Jelenic, P., and Faubert, J. (2005). Enhanced and diminished visuo-spatial information processing in autism depends on stimulus complexity. *Brain* 128, 2430–2441.

Berument, S. K., Rutter, M., Lord, C., Pickles, A., and Bailey, A. (1999). Autism screening questionnaire: diagnostic validity. *Br. J. Psychiatry* 175, 444–451.

Bolton, M. M., and Eroglu, C. (2009). Look who is weaving the neural web: glial control of synapse formation. *Curr. Opin. Neurobiol.* 19, 491–497.

Brainard, D. H. (1997). The psychophysics toolbox. *Spat. Vis.* 10, 433–436.

Brock, J., Brown, C., Boucher, J., and Rippon, G. (2002). The temporal binding deficit hypothesis of autism. *Dev. Psychopathol.* 14, 209–224.

Casanova, M. F., Buxhoeveden, D. P., Switala, A. E., and Roy, E. (2002). Minicolumnar pathology in autism. *Neurology* 58, 428–432.

Castellanos, F. X., Sonuga-Barke, E. J. S., Scheres, A., Di Martino, A., Hyde, C., and Walters, J. R. (2005). Varieties of

attention-deficit/hyperactivity disorder-related intra-individual variability. *Biol. Psychiatry* 57, 1416–1423.
Croner, L. J., Purpura, K., and Kaplan, E. (1993). Response variability in retinal ganglion-cells of primates. *Proc. Natl. Acad. Sci. U.S.A.* 90, 8128–8130.
Dakin, S. C., and Frith, U. (2005). Vagaries of visual perception in autism. *Neuron* 48, 497–507.
Debener, S., Makeig, S., Delorme, A., and Engel, A. K. (2005). What is novel in the novelty oddball paradigm? Functional significance of the novelty P3 event-related potential as revealed by independent component analysis. *Cogn. Brain Res.* 22, 309–321.
Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21.
Deutsch, K. M., and Newell, K. M. (2005). Noise, variability, and the development of children's perceptual-motor skills. *Dev. Rev.* 25, 155–180.
Di Russo, F., Martinez, A., Sereno, M. I., Pitzalis, S., and Hillyard, S. A. (2001). Cortical sources of the early components of the visual evoked potential. *Hum. Brain Mapp.* 15, 95–111.
Eichele, H., Juvodden, H. T., Ullsperger, M., and Eichele, T. (2010). Mal-adaptation of event-related EEG responses preceding performance errors. *Front. Hum. Neurosci.* 4:65. doi: 10.3389/fnhum.2010.00065
Foxe, J. J., and Simpson, G. V. (2002). Flow of activation from V1 to frontal cortex in humans. *Exp. Brain Res.* 142, 139–150.
Geurts, H. M., Grasman, R. P. P. P., Verté, S., Oosterlaan, J., Roeyers, H., van Kammen, S. M., and Sergeant, J. A.. (2008). Intra-individual variability in ADHD, autism spectrum disorders and Tourette's syndrome. *Neuropsychologia* 46, 3030–3041.
Happé, F., Ronald, A., and Plomin, R. (2006). Time to give up on a single explanation for autism. *Nat. Neurosci.* 9, 1218–1220.
Hill, E. L. (2004). Evaluating the theory of executive dysfunction in autism. *Dev. Rev.* 24, 189–233.
Hoeksma, M. R., Kemner, C., Verbaten, M. N., and van Engeland, H. (2004). Processing capacity in children and adolescents with pervasive developmental disorders. *J. Autism Dev. Disord.* 34, 341–354.
Isler, J. R., Martien, K. M., Grieve, P. G., Stark, R. I., and Herbert, M. R. (2010). Reduced functional connectivity in visual evoked potentials in children with autism spectrum disorder. *Clin. Neurophysiol.* 121, 2035–2043.
Jaskowski, P., and Verleger, R. (1999). Amplitudes and latencies of single-trial ERPs estimated by a maximum-likelihood method. *IEEE Trans. Biomed. Eng.* 46, 987–993.
Jemel, B., Mimeault, D., Saint-Amour, D., Hosein, A., and Mottron, L. (2010). VEP contrast sensitivity responses reveal reduced functional segregation of mid and high filters of visual channels in autism. *J. Vis.* 10, 13.
Jung, T. P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., and Sejnowski, T. J. (2001). Analysis and visualization of single-trial event-related potentials. *Hum. Brain Mapp.* 14, 166–185.
Kayser, J. (2009). *Current Source Density (CSD) Interpolation Using Spherical Splines – CSD Toolbox (version 1.1)*. Available at: http://psychophysiology.cpmc.columbia.edu/Software/CSDtoolbox
Kayser, J. R., and Tenke, C. E. (2006). Principal components analysis of Laplacian waveforms as a generic method for identifying ERP generator patterns: I. Evaluation with auditory oddball tasks. *Clin. Neurophysiol.* 117, 348–368.
Klistorner, A. I., and Graham, S. L. (2001). Electroencephalogram-based scaling of multifocal visual evoked potentials: effect on intersubject amplitude variability. *Invest. Ophthalmol. Vis. Sci.* 42, 2145–2152.
Makeig, S., Debener, S., Onton, J., and Delorme, A. (2004). Mining event-related brain dynamics. *Trends Cogn. Sci.* 8, 204–210.
Makeig, S., Jung, T.-P., Bell, A. J., Ghahremani, D., and Sejnowski, T. J. (1997). Blind separation of auditory event-related brain responses into independent components. *Proc. Natl. Acad. Sci. U.S.A.* 94, 10979–10984.
Maurer, U., Rossion, B., and McCandliss, B. D. (2008). Category specificity in early perception: face and word N170 responses differ in both lateralization and habituation properties. *Front. Hum. Neurosci.* 2:18. doi: 10.3389/neuro.09.018.2008
McAnally, K. I., Stuart, G., and Castles, A. (2001). Can contrast sensitivity functions in dyslexics be explained by inattention rather than a magnocellular deficit? *Vision Res.* 24, 3205–3211.
Milne, E., Scope, A., Pascalis, O., Buckley, D., and Makeig, S. (2009). Independent component analysis reveals atypical EEG activity during visual perception in individuals with autism. *Biol. Psychiatry* 65, 22–30.
Milne, E., Swettenham, J., Hansen, P., Campbell, R., Jeffries, H., and Plaisted, K. (2002). High motion coherence thresholds in children with autism. *J. Child Psychol. Psychiatry* 43, 255–263.
Mottron, L., Dawson, M., Soulieres, I., Hubert, B., and Burack, J. A. (2006). Enhanced perceptual functioning in autism: an update and eight principles of autistic perception. *J. Autism Dev. Disord.* 36, 27–43.
Nordahl, C. W., Dierker, D., Mostafavi, I., Schumann, C. M., Rivera, S. M., Amaral, D. G., and Van Essen, D. C. (2007). Cortical folding abnormalities in autism revealed by surface-based morphometry. *J. Neurosci.* 27, 11725–11735.
Nunez, P. L., and Srinivasan, R. (2006). *Electric Fields of the Brain*. Oxford: Oxford University Press.
Onton, J., Delorme, A., and Makeig, S. (2005). Frontal midline EEG dynamics during working memory. *Neuroimage* 27, 341–356.
Onton, J., Westerfield, M., Townsend, J., and Makeig, S. (2006). Imaging human EEG dynamics using independent component analysis. *Neurosci. Biobehav. Rev.* 30, 808–822.
Ozonoff, S., Strayer, D. L., McMahon, W. M., and Filloux, F. (1994). Executive function abilities in autism and Tourette syndrome: an information processing approach. *J. Child Psychol. Psychiatry* 35, 1015–1032.
Pernet, C. R., Chauveau, N., Gaspar, C., and Rousselet, G. A. (2011). Limo EEG: a toolbox for hierarchical linear modeling of electroencephalographic data. *Comput. Intell. Neurosci.* doi: 10.1155/2011/831409
Perrin, F., Pernier, J., Bertrand, O., and Echallier, J. F. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalogr. Clin. Neurosci.* 72, 184–187.
Plaisted, K., O'Riordan, M., and Baron-Cohen, S. (1998). Enhanced visual search for a conjunctive target in autism: a research note. *J. Child Psychol. Psychiatry* 39, 777–783.
Rojas, D. C., Maharajh, K., Teale, P., and Rogers, S. J. (2008). Reduced neural synchronization of gamma-band meg oscillations in first-degree relatives of children with autism. *BMC Psychiatry* 8, 66. doi: 10.1186/1471-244X-8-66
Rubenstein, J. L. R., and Merzenich, M. M. (2003). Model of autism: increased ratio of excitation/inhibition in key neural systems. *Genes Brain Behav.* 2, 255–267.
Russell, V. A., Oades, R. D., Tannock, R., Killeen, P. R., Auerbach, J. G., Johansen, E. B., and Sagvolden, T. (2006). Response variability in attention-deficit/hyperactivity disorder: a neuronal and glial energetics hypothesis. *Behav. Brain Funct.* 2, 30.
Salajegheh, A., Link, A., Elster, C., Burghoff, M., Sander, T., Trahmsb L., and Poeppeland, D. (2004). Systematic latency variation of the auditory evoked m100: from average to single-trial data. *Neuroimage* 23, 288–295.
Sannita, W. G. (2006). Individual variability, end-point effects and possible biases in electrophysiological research. *Clin. Neurophysiol.* 117, 2569–2583.
Schopler, E., Reichler, R. J., and Renner, B. R. (1988). *The Childhood Autism Rating Scale*. Los Angeles: Western Psychological Services.
Sieluzycki, C., Konig, R., Matysiak, A., Kus, R., Ircha, D., and Durka, P. J. (2009). Single-trial evoked brain responses modeled by multivariate matching pursuit. *IEEE Trans. Biomed. Eng.* 56, 74–82.
Simmons, D. R., Robertson, A. E., McKay, L. S., Toal, E., McAleer, P., and Pollick, F. E. (2009). Vision in autism spectrum disorders. *Vision Res.* 49, 2705–2739.
Singer, W., Engel, A. K., Kreiter, A. K., Munk, M. H. J., Neuenschwander, S., and Roelfsema, P. R. (1997). Neuronal assemblies: necessity, signature and detectability. *Trends Cogn. Sci.* 1, 252–260.
Tallon-Baudry, C., and Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends Cogn. Sci.* 3, 151–162.
Tucker, D. (1993). Spatial sampling of head electrical fields: the geodesic sensor net. *Electroencephalogr. Clin. Neurophysiol.* 87, 154–163.
von Spreckelsen, M., and Bromm, B. (1988). Estimation of single-evoked cerebral potentials by means of parametric modeling and kalman filtering. *IEEE Trans. Biomed. Eng.* 35, 691–700.
Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. San Antonio, TX: The Psychological Corporation.
Welsh, J. P., Ahn, E. S., and Placantonakis, D. G. (2005). Is autism due to brain desynchronization? *Int. J. Dev. Neurosci.* 23, 253–263.
Wilson, T. W., Rojas, D. C., Reite, M. L., Teale, P. D., and Rogers, S. J. (2007). Children and adolescents with autism exhibit reduced MEG steady-state gamma responses. *Biol. Psychiatry* 62, 192–197.

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 28 December 2010; accepted: 19 March 2011; published online: 30 March 2011.
Citation: Milne E (2011) Increased intra-participant variability in children with autistic spectrum disorders: evidence from single-trial analysis of evoked EEG. Front. Psychology ***2****:51. doi: 10.3389/fpsyg.2011.00051*
This article was submitted to Frontiers in Perception Science, a specialty of Frontiers in Psychology.

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 23 June 2011
doi: 10.3389/fpsyg.2011.00137

Modeling single-trial ERP reveals modulation of bottom-up face visual processing by top-down task constraints (in some subjects)

Guillaume A. Rousselet[1]*, Carl M. Gaspar[1], Kacper P. Wieczorek[1] and Cyril R. Pernet[2]

[1] Centre for Cognitive Neuroimaging, Institute of Neuroscience and Psychology, University of Glasgow, Glasgow, UK
[2] Brain Research Imaging Centre, SINAPSE Collaboration, University of Edinburgh, Edinburgh, UK

Edited by:
Paul Sajda, Columbia University, USA

Reviewed by:
Marios G. Philiastides, University of Nottingham, UK
Michael Tangermann, Technischen Universität, Berlin

***Correspondence:**
Guillaume A. Rousselet, Centre for Cognitive Neuroimaging, Institute of Neuroscience and Psychology, College of Medical, Veterinary and Life Sciences, 58 Hillhead street, Glasgow, G12 8QB, UK.
e-mail: guillaume.rousselet@glasgow.ac.uk

We studied how task constraints modulate the relationship between single-trial event-related potentials (ERPs) and image noise. Thirteen subjects performed two interleaved tasks: on different blocks, they saw the same stimuli, but they discriminated either between two faces or between two colors. Stimuli were two pictures of red or green faces that contained from 10 to 80% of phase noise, with 10% increments. Behavioral accuracy followed a noise dependent sigmoid in the identity task but was high and independent of noise level in the color task. EEG data recorded concurrently were analyzed using a single-trial ANCOVA: we assessed how changes in task constraints modulated ERP noise sensitivity while regressing out the main ERP differences due to identity, color, and task. Single-trial ERP sensitivity to image phase noise started at about 95–110 ms post-stimulus onset. Group analyses showed a significant reduction in noise sensitivity in the color task compared to the identity task from about 140 ms to 300 ms post-stimulus onset. However, statistical analyses in every subject revealed different results: significant task modulation occurred in 8/13 subjects, one showing an increase and seven showing a decrease in noise sensitivity in the color task. Onsets and durations of effects also differed between group and single-trial analyses: at any time point only a maximum of four subjects (31%) showed results consistent with group analyses. We provide detailed results for all 13 subjects, including a shift function analysis that revealed asymmetric task modulations of single-trial ERP distributions. We conclude that, during face processing, bottom-up sensitivity to phase noise can be modulated by top-down task constraints, in a broad window around the P2, at least in some subjects.

Keywords: ERP, GLM, single-trial analyses, bootstrap, N170, individual differences, faces, noise

INTRODUCTION

Following the first reports of larger scalp responses to faces compared to objects (Bötzel and Grüsser, 1989; Jeffreys, 1989; Jeffreys and Tukmachi, 1992; Jeffreys et al., 1992; Seeck and Grüsser, 1992), there have been hundreds of studies on the early event-related potentials (ERPs) to faces and objects. The vast majority of these studies used (i) averaged ERP, (ii) group statistics, and (iii) categorical designs. Their findings can be summarized shortly: sometime between 100 and 200 ms after stimulus onset, ERPs to different object categories tend to differ from each other, and faces are most of the time associated with larger N170 peaks than other object categories (Rossion and Jacques, 2008).

Recently, several research groups have started to study these early preferential responses to faces in individual subjects (Schyns et al., 2003, 2007; Philiastides and Sajda, 2006; Philiastides et al., 2006; Rousselet et al., 2007a, 2008a,b, 2009, 2010; Smith et al., 2007; Liu et al., 2009; Ratcliff et al., 2009; van Rijsbergen and Schyns, 2009). Individual subjects' ERPs, show, not surprisingly, systematic differences between faces and objects consistent with group effects reported so far (Rousselet et al., 2008a). These studies have also revealed inter-subject differences: despite coarse agreement between group and individual subject statistical analyses, individual subjects show reliable ERPs to faces and noise, which differ reliably across subjects, for reasons yet to be discovered (Rousselet et al., 2010; Gaspar et al., in press Reliability of ERP and single-trial analyses).

Individual differences in early visual processing have been largely ignored in the face literature, and implicitly treated as measurement errors that can be filtered out by averaging data across subjects. Although understanding the average brain is a worthy goal, only the single-trial approach, in conjunction with parametric designs, will allow us to understand brain mechanisms and the information content of brain states (Schyns et al., 2009; Schyns, 2010). In the single-trial framework, timing is essential. Indeed, how fast the visual system can discriminate among object categories provides strong constraints on possible computational implementations (Thorpe and Fabre-Thorpe, 2001; Rousselet et al., 2004; Thorpe, 2009). In particular, the timing of task modulations might help us tease apart periods of mostly bottom-up, stimulus driven activity, from time-windows engaging flexible neuronal populations that might be tuned to certain tasks. Thus, task modulations are key to understand brain mechanisms (Schyns, 1998; Pernet et al., 2007; Schyns et al., 2009).

For over 10 years, the ERP face literature has been debating the existence of task modulations of the N170 face preferential response. Several studies used targets vs. non-targets manipulations, in which faces at fixation are attended or ignored, and found no evidence of task modulations on the N170 (Séverac-Cauquil et al., 2000; Carmel and Bentin, 2002; Rousselet et al., 2007b), and its magnetic analog, the M170 (Lueschow et al., 2004; Furey et al., 2006; Okazaki et al., 2008). Similarly, intracranial recordings failed to reveal top-down modulation of the N200 to faces (Puce et al., 1999). One exception is found in a recent study, which reported an effect of category expectation on the N170 (Aranda et al., 2010). However, the effect seems to be weak and in a direction opposite to the one expected, so it might be a type I error. In contrast to target vs. non-target task manipulations, the N170 can be modulated by spatial attention (Jacques and Rossion, 2008; Crist et al., 2008) or by directing attention away from faces, in conditions in which letters superimposed on faces have to be discriminated (Eimer, 2000; Mohamed et al., 2009). Effects of language interference (Landau et al., 2010) and working memory (Sreenivasan et al., 2007) have also been suggested.

Hence, at least in some conditions, early face processing seems to be modulated by spatial attention and other factors. However, the modulations of face ERPs reported so far tend to be ill defined because it is unclear what aspect of face processing is modulated by the task. It remains also unclear how and when task demands affect the processing of a face presented at fixation. Very few studies have tackled this fundamental question by using a design in which the same stimulus is presented but processed differently because task requirements change the diagnosticity of input information (Pernet et al., 2007; Schyns, 1998). Schyns and his colleagues used reverse correlation techniques and large number of trials to reveal changes in single-trial information sensitivity (Schyns et al., 2007; Smith et al., 2007; van Rijsbergen and Schyns, 2009). However, although these studies show that ERPs are sensitive to different information from the same stimuli in different tasks, they do not provide a quantification of how task requirements affect the brain sensitivity to the same information.

One earlier study aimed at answering this question and reported larger N170 amplitude in a gender task compared to an identification task, but only for coarse, not fine scale information (Goffaux et al., 2003). This result suggests the use of certain face spatial scale information when it is relevant for the task at hand, an interaction between task demands and available information essential to reveal the information content of brain activity (Pernet et al., 2007; Schyns et al., 2009). However, the effects reported by Goffaux et al. (2003) seem very small and there was no report of task effects in individual subjects. It is also unclear if the effects might not be due to task modulations of the ERP sensitivity to the structured noise added to filtered images.

A more recent study reported one of the most striking and interpretable task effects on ERP face sensitivity. In Philiastides et al. (2006), experiment 2, a cue indicated on each trial how subjects were to process a subsequent stimulus: either discriminate its color (red vs. green) or its category (face vs. car). Colored pictures of faces and cars had two noise levels, created by altering the Fourier phase spectrum, which contains most of the information about object identity (Gaspar and Rousselet, 2009). In the categorization task, single-trial activity discriminated between face and car trials. In the same task, ERPs to faces and cars were sensitive to the level of image phase noise roughly in the time period 100–300 ms after stimulus onset. Among several important results in this paper, the authors show that in the color task, in which the noise dimension becomes task irrelevant, noise sensitivity was strongly reduced shortly after 200 ms. This is an important result because it suggests a timing for task effects in situations in which subjects discriminate stimuli presented at fixation: the first 200 ms of brain activity is mostly bottom-up, not modulated by task constraints, followed by a second period of brain activity which is modulated by top-down, task related influences.

In previous studies, similarly to Philiastides et al. (2006) we described phase noise sensitivity of face ERPs in the time-window 100–300 ms (Rousselet et al., 2008b, 2009, 2010). Although in their second experiment (Philiastides et al., 2006) used only two noise levels, 30 and 45% phase coherence, their results suggest that when we use a larger range of noise levels, as in our previous experiments, noise sensitivity should be strongly reduced after 200 ms when it is made task irrelevant. We tested this hypothesis by asking subjects to perform two tasks: the same face identity discrimination task (face 1 vs. face 2) we used in previous studies (Rousselet et al., 2008b, 2009, 2010) and the same color discrimination task (red vs. green) used by Philiastides et al. (2006). We performed both group level and single-subject analyses to reveal the detailed time-course of the task effects.

MATERIALS AND METHODS

Square brackets indicate the boundaries of 95% confidence intervals (CIs) constructed using a percentile bootstrap with 1000 samples (Wilcox, 2005).

SUBJECTS

We recruited 13 subjects, including the second and third authors, and 11 subjects from the Glasgow Psychology subject pool. Subjects' mean age was 24 years old (min = 20, max = 32); eight were females, 11 were right handed. Their mean high-contrast 63 cm decimal acuity was 104 (min = 99, max = 110); their low-contrast 63 cm decimal acuity was 96 (min = 89, max = 103). All subjects had a Pelli-Robson contrast sensitivity of 1.95 and successfully passed the Ishihara color blindness test for red-green color deficiencies. On average subjects had 19 years of education (min = 15, max = 23). All subjects except the two authors received £6/hour for their participation and all subjects gave written informed consent. The research ethics board from the University of Glasgow approved the research protocol.

STIMULI

We used two front-view male face photographs cropped within a common 4.3° × 6.3° oval frame and pasted on a uniform 9° × 9° gray background (**Figure 1**). These faces were selected from a set of 10 faces, which are described in detail in previous publications (Gold et al., 1999; Husk et al., 2007). All stimuli had the same global amplitude spectrum. We added noise to their phase spectra so that their percentage of global phase coherence ranged from 10 to 80%, with 10% increments. Noise was random on each trial, which means that subjects never saw the exact same image twice. We also colorized the faces with red and green tones by manipulating

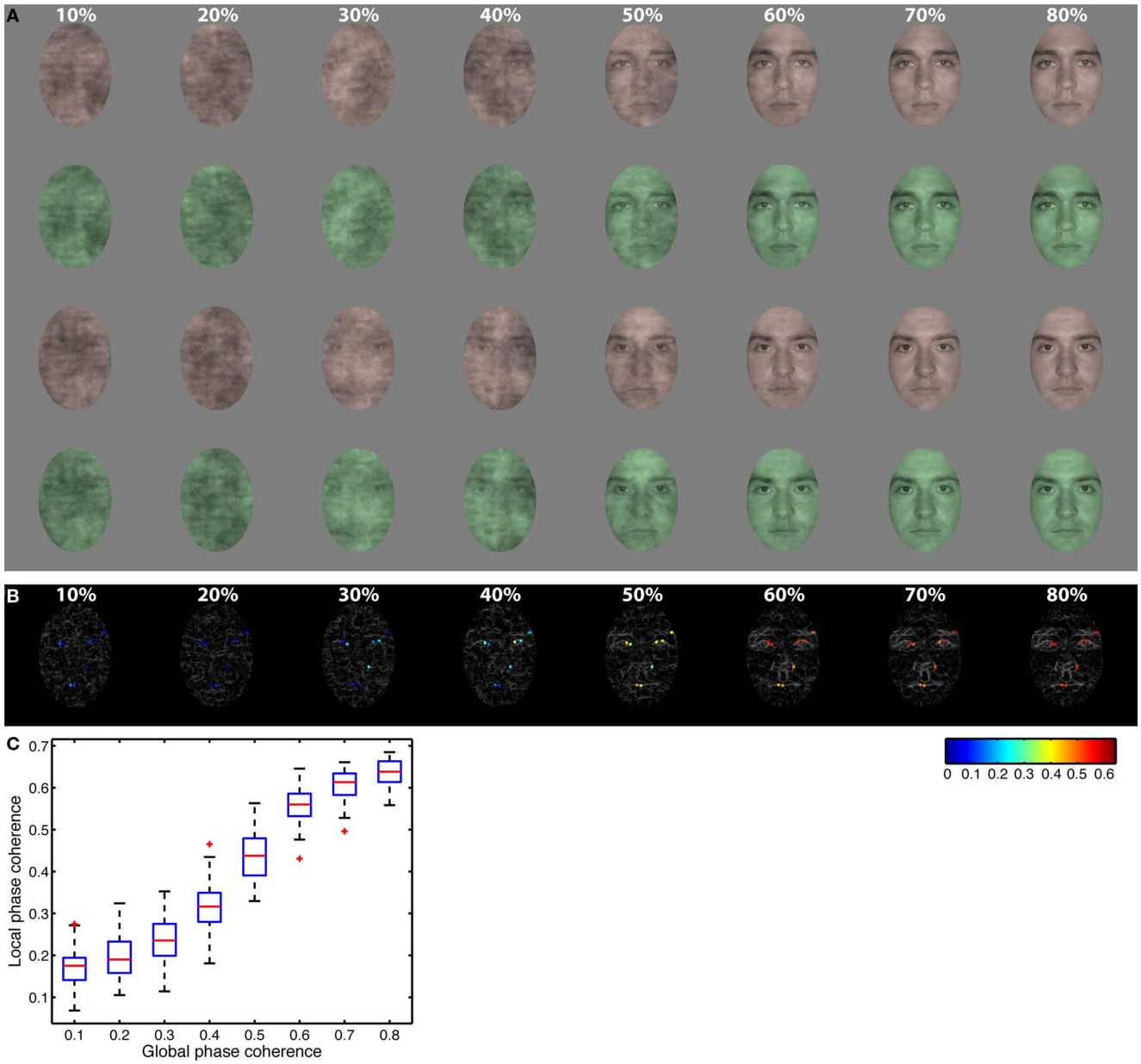


FIGURE 1 | Stimuli. (A) All observers saw the same two faces presented in red or green at eight levels of global phase coherence. Rows 1 and 2: face identity 1; rows 3 and 4: face identity 2; rows 1 and 3: red tones; rows 2 and 4: green tones. **(B)** Gray lines show edges identified using Kovesi's local phase coherence algorithm in eight face examples ranging from 10% phase coherence (left) to 80% phase coherence (right). Superimposed on each edge map, local phase coherence is color coded at the 10 pixels with the highest local phase coherence. These 10 pixels were identified in the two original faces at 100% global phase coherence. Local phase coherence was maximal (red) at 80% global phase coherence, and approached zero (blue) at 10%. **(C)** Boxplots of local phase coherence at each level of global phase coherence for all the images seen by one subject. There is a non-linear relationship between the two variables.

the hue (H), saturation (S), and value (V) of the original images (red: $H = 0.04$, $S = 0.17$, V = unchanged; green: $H = 0.34$, $S = 0.23$, V = unchanged). The value (V) was normalized so that, on average, each face regardless of color or phase coherence had the same average luminance (about 33 cd/m^2) and RMS contrast (0.1). We colorized only the face itself and not the background, and used relatively small images to ensure that subjects paid attention to the face in the two tasks. This manipulation increases the likelihood that task effects are due to a change in the task relevance of one stimulus dimension while subjects attempt to discriminate the same stimulus. A Dell Precision 390 workstation with Nvidia Quadro FX 3450/4000 graphics card and MATLAB Psychophysics Toolbox controlled the stimulus display. Images were displayed on a SAMSUNG SyncMaster 1100MB CRT monitor with a resolution of 800 × 600 pixels and a 85-Hz refresh rate. The screen was 28° × 21°of visual angle.

EXPERIMENTAL DESIGN

Testing was conducted in a sound-attenuated booth in which the monitor was the only source of light. An 80-cm viewing distance was maintained with a chinrest. We tested subjects in two experimental sessions. The first day was a practice behavioral session; the second day consisted of both behavioral tasks and simultaneous EEG recordings. Each day, subjects performed two interleaved tasks. On half of the blocks they performed a one-interval, two-alternative forced choice task discriminating between two faces. On the other half of the blocks, they discriminated between two colors. Identity and color tasks were blocked so subjects could focus on one task for an entire block of trials, without having to prepare to switch task on each trial (Johnson and Olshausen, 2003), in an attempt to increase the likelihood of finding strong task effects. The same stimuli were presented in the two tasks. In both tasks, on each trial, one face appeared briefly (36 ms), and subjects had to indicate which of two possible faces or two possible colors was presented by pressing 1 or 2 on the numerical pad of the keyboard. The association between button and identity/color was assigned randomly for all subjects. Subjects were given unlimited time to respond, and were told to emphasize response accuracy, not speed. All subjects performed the task with the same single pair of male faces throughout the experiment. Subjects saw eight conditions along a noise-signal continuum, from 10 to 80% phase coherence, with increments of 10% (**Figure 1**).

There were 10 blocks of 96 trials: 960 trials in total, with 120 trials per level of phase coherence. Within each block, there were equal repetitions of each face, each color and each phase coherence level. Each block was preceded by 10 practice trials that allowed subjects to learn the stimulus-key association. Practice trials were used to ensure a high level of performance in older subjects, whose data are not reported here. In a regular trial, a small fixation cross – a 0.3° " + " in the middle of the screen – appeared for 500 ms, after which a blank screen was presented for a random duration ranging from 500 to 700 ms (**Figure 2A**). Then a test stimulus was presented for 36 ms, followed by a blank screen that stayed on until subjects provided their response. Practice trials were very similar, except that immediately after stimulus presentation, a choice screen appeared that showed each face in grayscale (identity task) or red and green noise textures (color task) simultaneously, one above each other with a corresponding label below each item. Auditory feedback was provided after the subject pressed a response key, with low- and high-pitched tones indicating incorrect and correct responses. Feedback was provided only during practice trials.

EEG RECORDING AND PREPROCESSING

We acquired EEG data with a 128-channel Biosemi Active Two EEG system (BioSemi, Amsterdam, Netherlands). We recorded from four additional electrodes – UltraFlat Active BioSemi electrodes – below and at the outer canthi of both eyes. Analog signal was digitized at 512 Hz and band-pass filtered online between 0.1 and 200 Hz. Electrode offsets were kept between ±20 μV.

Offline, data were average-referenced. Then, we removed bad channels without interpolation, applied a 40-Hz low-pass filter and epoched the data between −300 and 1,200 ms. An ICA (Makeig et al.,

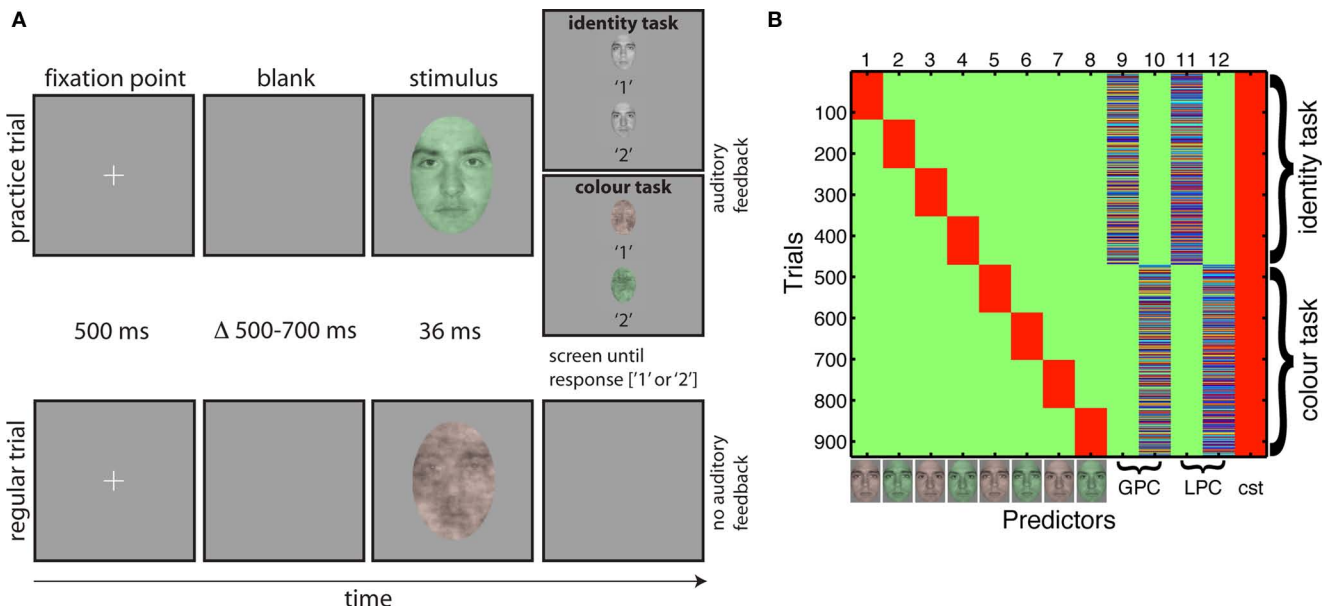


FIGURE 2 | Tasks and design matrix. (A) Organization of practice trials and regular trials in the two tasks. A trial started with the presentation of a fixation point for 500 ms. Then, after a random delay ranging from 500 to 700 ms, a stimulus was presented for about 36 ms. During practice trials (top row), a choice screen appeared immediately after the stimulus, showing the two targets of the task and their associated response keys. The screen stayed on until the subject's response, which was followed by auditory feedback, before the trial sequence resumed. During regular trials (bottom row), a blank screen appeared immediately after the stimulus, and remained on until the subject's response. No feedback was provided during regular trials. Stimuli are not drawn to scale. **(B)** Example of a design matrix in one subject (color scale: green = 0; red = 1). The first eight predictors were categorical: they indicate the stimulus type (i.e., red or green, face 1 or 2) and the task. The next four predictors were continuous: global phase coherence (GPC) and local phase coherence (LPC) in the identity and color tasks. Each continuous predictor was *z*-scored independently before insertion in the design matrix. The last column was a constant term (cst = 1 s).

2004), as implemented in the *runica* EEGLAB function (Delorme and Makeig, 2004; Delorme et al., 2007) was then computed and we removed components corresponding to blink activity, identified by visual inspection of their scalp topographies, time-courses and activity spectra. Subsequently, we re-epoched the data between −300 and 500 ms, and subtracted the average baseline activity from each time point. Trials with abnormal activities were excluded based on a ±100-μV threshold for extreme values. An epoch was rejected for abnormal trend if it had a slope larger than 75 μV/epoch and a regression R^2 larger than 0.3. All remaining trials were included in the analyses, whether they were associated with correct or incorrect behavioral responses. After epoch rejection, the average number of trials per subject was 904 (min = 849, max = 958).

GENERAL LINEAR MODELING OF EEG DATA

Subjects' epoched data were modeled using LIMO EEG, an open source Matlab toolbox for hierarchical GLM, compatible with EEGLAB: https://gforge.dcn.ed.ac.uk/gf/project/limo_eeg/ (Pernet et al., 2011). The general linear model was used to express single-trial ERP amplitudes, in microvolt, independently at each time point and each electrode, using the model:

$$\begin{aligned} ERP_{t,e} &= XB + \varepsilon \\ ERP_{t,e} &= \beta_0 + cat_1\beta_1 + cat_2\beta_2 + ... + cat_8\beta_8 + \phi_{G\text{-}ID}\beta_9 + \phi_{G\text{-}CO}\beta_{10} \\ &\quad + \phi_{L\text{-}ID}\beta_{11} + \phi_{L\text{-}CO}\beta_{12} + \varepsilon \end{aligned} \quad (1)$$

In Eq. 1, all trials for each time frame t and electrode e ($ERP_{t,e}$ dimension $1 \times n$) were modeled as the sum of a constant term β_0, the eight experimental conditions (each combination of stimulus identity, color and task – $2 \times 2 \times 2 = 8$, Cat_{1-8} – corresponding to the first 8 columns of the design matrix), the global phase coherence in the identity and the color tasks ($\varphi_{G\text{-}ID}$, $\varphi_{G\text{-}CO}$), the local phase coherence in the identity and the color tasks ($\varphi_{L\text{-}ID}$, $\varphi_{L\text{-}CO}$), and an error term ε. All predictors formed the design matrix X of dimension $n \times p$ (**Figure 2B**, $p = 13$). The beta parameters (dimension $p \times 1$) were found using an ordinary least square solution.

Global phase coherence was our image noise manipulation. Kovesi's (1999, 2003) local phase coherence is a measure of wavelet phase alignment across spatial frequencies, which is independent of image contrast and luminance. Local phase coherence may predict subjects' behavior in a natural scene classification task (Gaspar and Rousselet, 2009) and seems to provide a good representation of non-linear changes in local image structure imposed by the linear global phase coherence manipulation (Rousselet et al., 2008b, 2009, 2010). In our stimuli, pixels with high local phase coherence corresponded to local edges around the eyes, nose, and mouth (**Figure 1**).

The design matrix represents a typical ANCOVA model with categorical and continuous predictors. However, whereas in ANCOVA one is usually interested in the categorical effects whilst controlling for covariates, here we were interested in the covariate effects: we looked at the relationship between image phase coherence and single-trial ERP amplitude in the identity and the color tasks whilst accounting for the main effects of identity, color, and task. For instance, the average ERP in one condition (e.g., identity discrimination of green face 1) could differ from the average ERP in another condition (e.g., color discrimination of green face 1) because subjects might be less attentive in the easier color task compared to the more challenging identity task. These confounding mean ERP differences were accounted for in the design matrix, thus allowing us to measure how single-trial ERPs were modulated by image noise in the two tasks. We used linear contrasts to combine the beta weights associated with the global and local phase coherence predictors in the identity task (column 9 + column 11 in **Figure 2B**) and in the color task (column 10 + column 12) to study the time-course of the overall ERP noise sensitivity:

$$d_{identity} = \beta_9 + \beta_{11} \quad (2)$$

$$d_{color} = \beta_{10} + \beta_{12} \quad (3)$$

We did not look at task modulations separately for global phase coherence and local phase coherence because these two predictors were strongly correlated (identity: mean angle = 17.4 [17.3, 17.6], mean correlation = 95.4 [95.3, 95.5]; color: mean angle = 17.6 [17.3, 17.9], mean correlation = 95.3 [95.1, 95.5]; mean angle differences between the two tasks = −0.1727 [−0.5705, 0.2544]). High correlation between regressors may lead to unstable beta parameter estimates, whereas their linear combination remains stable, hence our analysis of the combination of global and local phase coherences. We refer to this summary statistics as noise sensitivity in the rest of the paper and explored task effects by contrasting $d_{identity}$ and d_{color}.

GROUP LEVEL ANALYSES

Group analyses of noise sensitivity task modulations were computed using a bootstrap-t technique for paired samples with 1000 resamples (Wilcox, 2005). Although full scalp analyses are possible in LIMO EEG, we performed the analyses at only one electrode for two reasons: first because we observed in previous studies that noise sensitivity is localized at few posterior electrodes that display redundant information (Rousselet et al., 2008b, 2009, 2010); second because we wanted to compare different group analyses to single-subject analyses. We analyzed group results using four different ways to pull data together.

A popular way to do group analyses is to average the data across subjects, find the best electrode in this group average, and make a measurement at that same electrode in all subjects –group defined best electrode. Here, we averaged across subjects the R^2 maps of the ANCOVA model fit to the data, and selected the electrode showing the largest mean R^2. Bootstrap paired *t*-tests were then computed between noise sensitivity contrasts in the identity and color tasks at all time points at this electrode.

A potentially more fruitful way to do group statistics is to optimize the electrode by selecting the best electrode independently in each subject (Foxe and Simpson, 2002; Liu et al., 2002; Rousselet et al., 2010). We thus took the electrode at which the model provided the best fit independently in each subject, i.e., where R^2 was the largest for each subject –R^2 optimized electrode. Then we computed paired *t*-tests between noise sensitivity contrasts from these potentially spatially different electrodes. The signal at R^2 optimized electrodes was the most sensitive to image and task parameters as described by the design matrix and therefore constitutes the most likely candidate for reflecting the activity of cortical sources sensitive to image information. Hence, this kind of optimized

averaging tends to average signals that reflect common processing across subjects, whereas using the same spatial electrode may lead to averaging signals reflecting different processes.

Yet another way to optimize electrodes across subjects consists in selecting for each subject the electrode with the largest noise sensitivity task difference –task effect optimized electrode. In this case, instead of taking the electrode where the ANCOVA model provided the best fit, we selected for each subject the electrode showing the strongest noise sensitivity task effect. The paired *t*-test was then computed between noise sensitivity contrasts from these potentially spatially different electrodes.

Finally, we used the maximum absolute beta coefficients across electrodes computed at each time point (the envelope), to ensure our analyses did not miss local maxima at electrodes other than the one showing the largest R^2. For every subject, we computed a paired *t*-test between noise sensitivity contrasts from the envelopes.

For both group and single-subject analyses, task modulations at one electrode were quantified by normalizing the maximum absolute task difference in noise sensitivity by the maximum absolute noise sensitivity in the identity task –the maxima were defined across time frames:

$$\text{Task modulation} = 100 \times \max\left(\left|d_{\text{identity}} - d_{\text{color}}\right|\right) \Big/ \max\left(\left|d_{\text{identity}}\right|\right) \quad (4)$$

CONTROL FOR MULTIPLE COMPARISONS

We controlled for multiple comparisons using bootstrap and the clustering technique as implemented in the Matlab Fieldtrip toolbox, with a minimum of two neighboring channels per cluster (Maris and Oostenveld, 2007). As described in (Pernet et al., 2011), the clustering technique in LIMO EEG works for analyses both at single electrodes (temporal clustering) and at multiple electrodes (spatial–temporal clustering). For group analyses, because only one electrode or equivalent electrode was considered, we employed temporal clustering to control for multiple comparisons. For single-subject analyses, because the whole scalp was analyzed, we employed spatial–temporal clustering (familywise error rate = 0.05).

For *t*-tests and ANOVAs the validated bootstrap technique includes centering the empirical distributions of each between-subject and within-subject levels so that the null-hypothesis of no difference in means is true (Berkovits et al., 2000; Wilcox, 2005; Seco et al., 2006). Thus, for the group paired *t*-tests, noise sensitivity contrasts across subjects were centered for each condition separately and paired *t*-tests were computed 1000 times by sampling subjects with replacement (Wilcox, 2005). However, this technique is not appropriate to our ANCOVA single-subject analyses because the continuous covariates can potentially have as many levels as trials. We used therefore a different strategy to derive an estimate of the sampling distribution of our *F* statistics under the null-hypothesis. For each subject, epoched single-trials were sampled with replacement and fitted to the original design matrix, thus breaking the link between the data and the design –an estimation of the data distribution under the null-hypothesis H0.

For both the bootstrap paired *t*-test (group analyses) and the bootstrap ANCOVA (single-subject analyses), in each bootstrap loop we first computed the sum of each temporal or spatial–temporal cluster of contiguous significant *F* values (univariate $p < 0.05$), separately for each predictor, each linear contrast, and in the case of the ANCOVA, for the global fit of the entire model (R^2). Second, we saved the maxima across these cluster sums – one maximum for each *F* test (familywise correction). After performing these steps 1000 times for group statistics and 600 times for single-subject analyses (as recommended for various linear models, Wilcox, 2005), we used the 95th percentiles of the bootstrapped maximum *F* cluster sums to threshold the original *F* cluster sums. For each test, the significant original *F* values (univariate $p < 0.05$) were clustered and if their sum were larger than the corresponding bootstrapped maximum cluster sum threshold, the cluster was significant.

NOISE SENSITIVITY CLUSTER STATISTICS

For each subject, we used a percentile bootstrap rather than using an *F* test of noise sensitivity (sum of beta coefficients for global and local phase coherence). Bootstrap distributions were used to compute 95% CIs under H0, during the same simulation that was used to estimate the *F* distributions of the ANCOVA parameters. These thresholds were then applied to each bootstrap to mark significant noise sensitivity. Significant effects were then clustered and a maximum sum of absolute noise sensitivity was saved for each bootstrap. The bootstrap distributions of maximum sum of absolute noise sensitivity computed under H0 were used to cluster the observed noise sensitivity in each task.

SHIFT FUNCTION ANALYSES OF THE DECILES OF SINGLE-TRIAL ERP DISTRIBUTIONS

We also used the shift function to measure how single-trial ERP distributions changed from the identity task to the color task. The shift function compares entire distributions instead of relying exclusively on one point estimate such as the mean or the median. In our application of the shift function, the *x*-axis is the Harrell–Davis (hd) estimator of quantiles one to nine of the single-trial ERPs in the identity task (see Wilcox, 2005, pp. 71–73 and 139–141). The *y*-axis is the difference, Delta, between the hd estimators of the quantiles of the identity and color ERP distributions. Hence, the shift function represents how much the data from one task must be shifted to be comparable to the data from another task at each quantile. Task differences were estimated by a bootstrap procedure, and corrected for multiple comparisons such that the simultaneous probability coverage of the 9 CIs remained close to the nominal 0.05 alpha level (see Wilcox, 2005, pp. 151–155). The analyses were performed on modeled single-trial ERP data at the max R^2 electrode (i.e., the electrode at which the model explained best the data); they included all the significant time points that contained the maximum noise sensitivity task difference. Modeled ERP are more meaningful to analyze because they are reconstructed after removing the error term, the part of variance that the model cannot explain.

RESULTS

We consider first the group analyses, second the single-trial analyses, third the comparison of group and single-trial analyses and fourth the shift function analyses of the deciles of the single-trial ERP distributions.

GROUP ANALYSES

Group defined best electrode

If the best electrode is defined as the electrode showing the largest mean R^2 across subjects, we obtain the results in **Figure 3A**. This best electrode was right posterior–lateral (B8 in the Biosemi system, between PO8 and PO10) and had a maximum mean R^2 of 0.23 [0.17, 0.32] that peaked at 141 ms. The mean ERPs at the eight global phase noise levels started to diverge shortly after 100 ms in the identity and the color tasks. The parametric ERP modulation by image noise can be better appreciated by looking at the time-course of the group-averaged noise sensitivity, which peaked at the same electrode and time point as R^2 did. Noise sensitivity was reduced in the color task compared to the identity task in a single cluster, between 139 and 277 ms after stimulus onset (**Table 1**). At the latency of the maximum task effect, 242 ms, there was 20.7% noise sensitivity reduction compared to the maximum sensitivity in the identity task.

R^2 optimized electrode

The electrode with the largest R^2 was also the electrode with maximum noise sensitivity or was part of the same cluster as the electrode with maximum noise sensitivity and behaved similarly to it. Across subjects, max R^2 electrodes were all located in a cluster of lateral posterior electrodes, as reported in previous experiments and as expected from the face literature. R^2 averaged across subjects was stronger over the right hemisphere. This pattern was also found in individual subjects (**Figures 5–8**): eight subjects had a maximum R^2 at right hemisphere electrodes; two subjects at left hemisphere electrodes; three subjects at midline electrodes. The right hemisphere electrodes of maximum model fit included B8 (one subject) or one of its neighbors (seven subjects). The maximum mean R^2 at the optimized electrode was 0.27 [0.2, 0.35], and peaked at 139 ms (**Figure 3B**; **Table 1**). As expected if R^2 results were sufficiently consistent across subjects, this optimized maximum average R^2 was larger than at the group defined best electrode. There was a significant task effect in a single cluster between 172 to 275 ms post-stimulus onset, with 18.7% noise sensitivity reduction at the latency of the maximum task difference, 213 ms.

Task effect optimized electrode

Results of this analysis are presented in **Figure 3C**. Although we selected for each subject the electrode showing the largest task difference in noise sensitivity, no differences could be observed at the group level. Indeed, taking the largest effect can be misleading because certain predictors can be significant at electrodes and time frames at which the overall model does not explain the data significantly. This is indeed what we found: except for 3 subjects for whom maximum task effects occurred at electrodes that were part of the cluster of electrodes with the maximum R^2, for the other 10 subjects R^2 was lower and early noise sensitivity (<200 ms) was weak at the electrodes of maximum task differences (**Figure 3C**). Hence, across subjects, noise sensitivity and task effects were not significant at the electrode optimized based on maximum task differences.

Maximum absolute betas

Analyses on the beta coefficient *envelopes* gave results similar to those obtained on the group defined best electrode and the R^2 optimized electrode (**Figure 3D**; **Table 1**). Two significant clusters were observed: a first task effect occurred between 154 to 254 ms post-stimulus onset, with 18.2% noise sensitivity reduction at the latency of the maximum task difference, 197 ms; a second effect of similar size occurred after 400 ms. This analysis suggests that we did not miss the big picture by defining the electrode to analyze based on the group-averaged R^2 or the single-subject max R^2. The statistical tests might suggest that the group defined best electrode did a better job because it showed a significant task effect earlier than the one observed in the R^2 optimized test. However, picking the best group electrode to show group effects is circular because what the result ought to be is unknown. By contrast, selecting the best electrode separately in each subject takes into account inter-subject variability and leads to group results more sensitive to individual differences.

In addition to group analyses at all time points, we compared noise sensitivity across subjects at the latency of the P1, N170, and P2 peaks, for the sake of comparisons with previous studies (**Figure 4**). For each subject, the latency of the peaks was measured at the max R^2 electrode in the 80% coherence condition (face ERP). Measurements at electrode B8 gave similar results. Then, the mean sensitivity was measured in time-windows encompassing five time points on either side of the peak latency, hence about 21.5 ms. This analysis revealed, across subjects, weak to no noise sensitivity around the P1 (~96 ms), and stronger noise sensitivity around the N170 (~146 ms) and the P2 (~207 ms). Importantly, only the P2 showed task modulations of noise sensitivity. Similar P2 results were obtained at the latency of the P2 defined at the group level.

SINGLE-SUBJECT ANALYSES

Figures 5–8 provide, for each subject, a detailed description of their R^2, noise sensitivity, task effects and behavioral results. The time-course of the R^2 functions and of the beta coefficients for noise sensitivity are similar to those reported previously in young subjects (Rousselet et al., 2008b, 2009, 2010). Because the main purpose of our study was to quantify task modulations of early ERP noise sensitivity, we focus the report of the single-trial analyses on the electrode showing the maximum R^2 for each subject. All these electrodes were found at posterior–lateral locations. No comparable fits were observed at frontal electrodes. Thus, our analyses seem to capture task modulations of evoked noise sensitivity from the visual system, rather than electrophysiological correlates of the top-down modulation signal itself. Figures at the electrodes showing the maximum noise sensitivity in the identity task or the color task were almost identical to those presented here (max R^2) because these electrodes were either the same electrodes or part of the same cluster of electrodes.

Single-trial analyses revealed an inter-subject variability hidden behind the seemingly simple group averages and statistics. Individual subjects differed widely in the shape of their ERPs, R^2 functions, scalp topographies, nature and time-course of the task effects (**Figures 5–8**). The mean of the maximum R^2 measured in each subject was 0.31 [0.24, 0.38], min = 0.17, max = 0.57; it peaked at 141 ms [136, 149], min = 133, max = 186. Image noise sensitivity started at about 100 ms in both tasks. In the identity task, the median onset was 100 ms [98, 105], min = 92, max = 115; in the color task it was 100 ms [96, 111], min = 86, max = 133. The median difference between the two tasks was 0 ms [−7.8, 4],

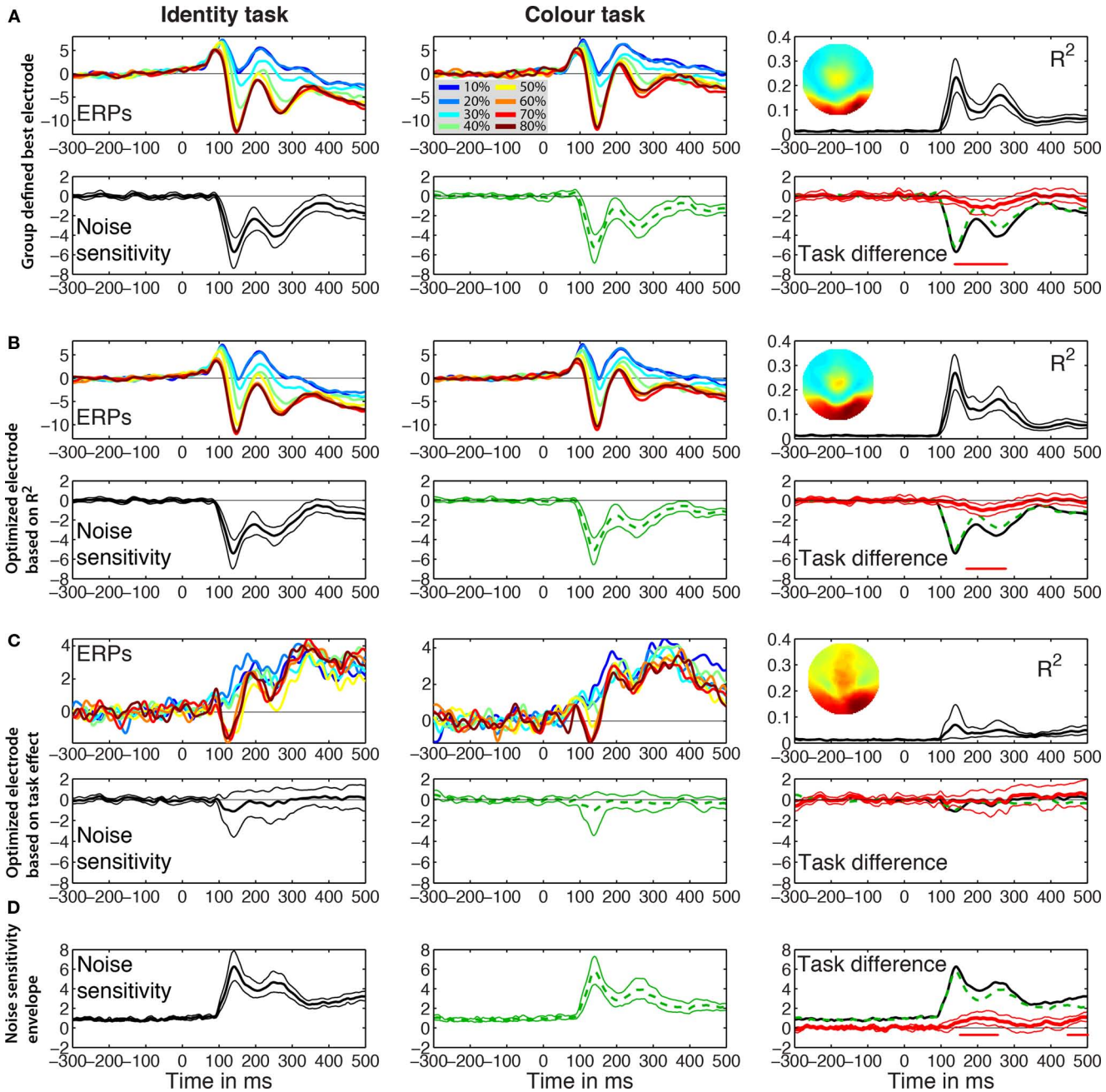


FIGURE 3 | Event-related potentials group results. The left and middle columns contain the results for the identity and the color tasks. The right column shows the R^2 and the noise sensitivity task differences. **(A)** Group defined best electrode. Mean ERPs are shown color coded at each level of global phase coherence in the two tasks. In the R^2 plot, the inset shows the topographic map of the interpolated R^2 values at the latency of the maximum R^2. Noise sensitivity is the sum of the global and local phase coherence beta weights in μV/std of the predictor. Thick lines represent averaged data, surrounded by thin lines for the 95% percentile bootstrap CI. The red horizontal bar under the zero line indicates time points of significant effects, based on bootstrap *t*-test with temporal cluster correction for multiple comparisons. The task difference in red is identity (black continuous line) minus color (green dashed line). At the latency of the maximum task difference observed within the first 300 ms after stimulus onset, noise sensitivity in the identity task was –4 μV/std, color = –2.8 μV/std, difference = –1.2 μV/std. **(B)** R^2 optimized electrode. The topographic map was obtained by averaging the maps from individual subjects. At the latency of the maximum task difference, noise sensitivity in the identity task was –2.8 μV/std, color = –1.8 μV/std, difference = –1 μV/std. **(C)** Task effect optimized electrode. The R^2 bump between 100 and 200 ms was mostly due to three subjects (S9–S11) who had maximum task effects at electrodes that were part of the cluster of electrodes with the maximum R^2. **(D)** Noise sensitivity envelope. The maximum across electrodes of the absolute noise sensitivity was used for each subject. At the latency of the maximum task difference, noise sensitivity in the identity task was 3.9 μV/std, color = 2.8 μV/std, difference = 1 μV/std.

min = −21, max = 16. Results using the mean were similar and a shift function analyses failed to reveal significant differences at any deciles of the distribution of onset differences.

Only eight subjects showed a significant task modulation of noise sensitivity. Noise sensitivity decreased in the color task compared to the identity task in seven subjects: S_1 = 29%, S_3 = 46%, S_4 = 35%, S_6 = 19%, S_7 = 39%, S_{10} = 29%, S_{11} = 39%, mean = 33.7%. Subject S_2 showed an effect in the opposite direction (**Figure 5**), with significantly stronger noise sensitivity in the color task than in the identity task (40.7% sensitivity increase). All subjects but S_{11} (**Figure 7**) had a single cluster of significant task differences.

There was no significant link between task effects and behavioral thresholds: mean behavioral 75% correct threshold was 0.37 global phase coherence for the five subjects who did not show a significant ERP task modulation, it was 0.39 global phase coherence for subjects who did show a reduction in the color task (difference = −0.02 [−0.07; 0.04]), and 0.37 global phase coherence for S_2, who showed opposite ERP results. Analyses with the median gave similar results.

The max R^2 electrodes almost never showed significant differences between the two faces (identity task) or the two colors (color task), in keeping with previous reports on faces, cars, and words (Nobre et al., 1998; Philiastides et al., 2006; Rousselet et al., 2008b, 2010). One subject showed both significant identity sensitivity in the identity task and a significant task modulation of identity sensitivity at few time points around 450 ms after stimulus onset. Five subjects showed either identity or color sensitivity in one or the other task, but without significant task modulation, or significant task modulations but without significant identity or color sensitivity.

Table 1 | Task effects.

	Group analyses			Single-trial analyses
	Best group electrode (B8)	**R^2 optimized electrode**	**Beta envelope**	**Mean of eight subjects with significant task effects**
Onset	139 ms	172 ms	154 ms	214 ms [155, 271]
Peak latency	242 ms	213 ms	197 ms	297 ms [246, 352]
Effect size	20.7%	18.7%	18.2%	34.6% [28.9, 39.5]
Duration	139 ms	104 ms	101 ms	186 ms [142, 234]

Peak latency is the latency of the maximum absolute task effect. Effect size is defined in Section "Group Level Analyses," Eq. 4.

GROUP VS. SINGLE-TRIAL ANALYSES

Group analyses suggest a decrease in noise sensitivity in the color task compared to the identity task around 140–300 ms post-stimulus onset. This task modulation was observed at the group defined best electrode and the R^2 optimized electrode (**Figure 9**). In the spatial–temporal clusters containing these electrodes, single-trial analyses revealed a different picture: 8 subjects out of 13 had a significant task modulation of noise sensitivity; 1 had an increased sensitivity in the color task and 7 had a decreased sensitivity. In the time-window of the group effect, only six subjects showed a significant effect; a maximum of five subjects showed an effect simultaneously, including the subject who had an effect in the opposite direction. Thus, in our sample, at any time point showing a significant group effect, there were at most 5 subjects out of 13 showing a significant effect, 4 of which in the same direction as the group effect (31%).

Onsets and durations of task effects also revealed discrepancies between group and single-trial analyses (**Table 1**). At the R^2 optimized electrode, the group task effects started at 172 ms and lasted 103 ms. In contrast, for the eight subjects that showed significant effects, the average task effect onset was 214 ms [155, 271] min = 86, max = 332; the average task effect duration was the 186 ms [142, 234], min = 92, max = 320.

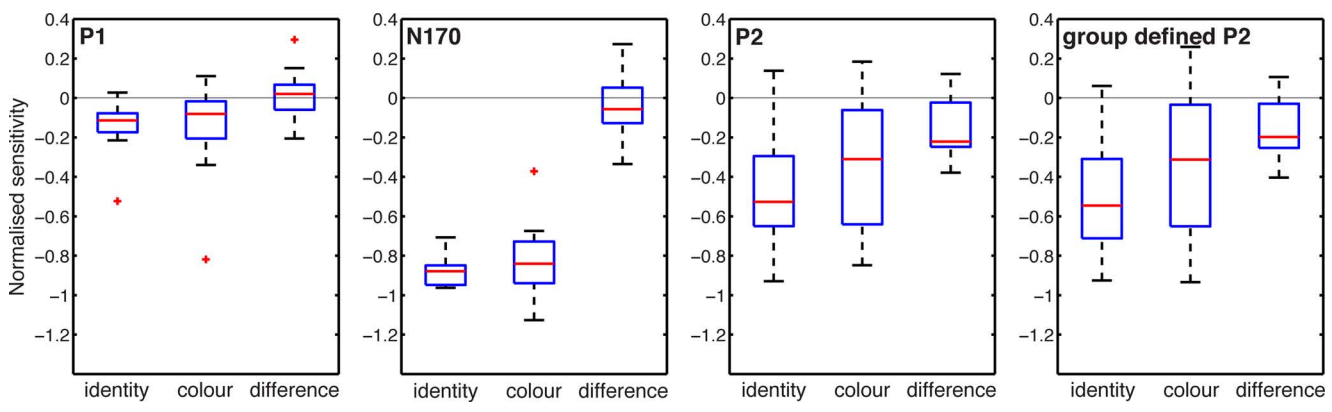


FIGURE 4 | Task effects around peak time-windows. Noise sensitivity in the two tasks was normalized by dividing by the maximum absolute noise sensitivity in the identity task. The three first ERP peaks were measured at these latencies across subjects: P1 (median = 96, min = 82, max = 119), N170 (median = 146, min = 131, max = 184), P2 (median = 207, min = 180, max = 256). The P2 defined from the group-averaged data had a latency of 207 ms.

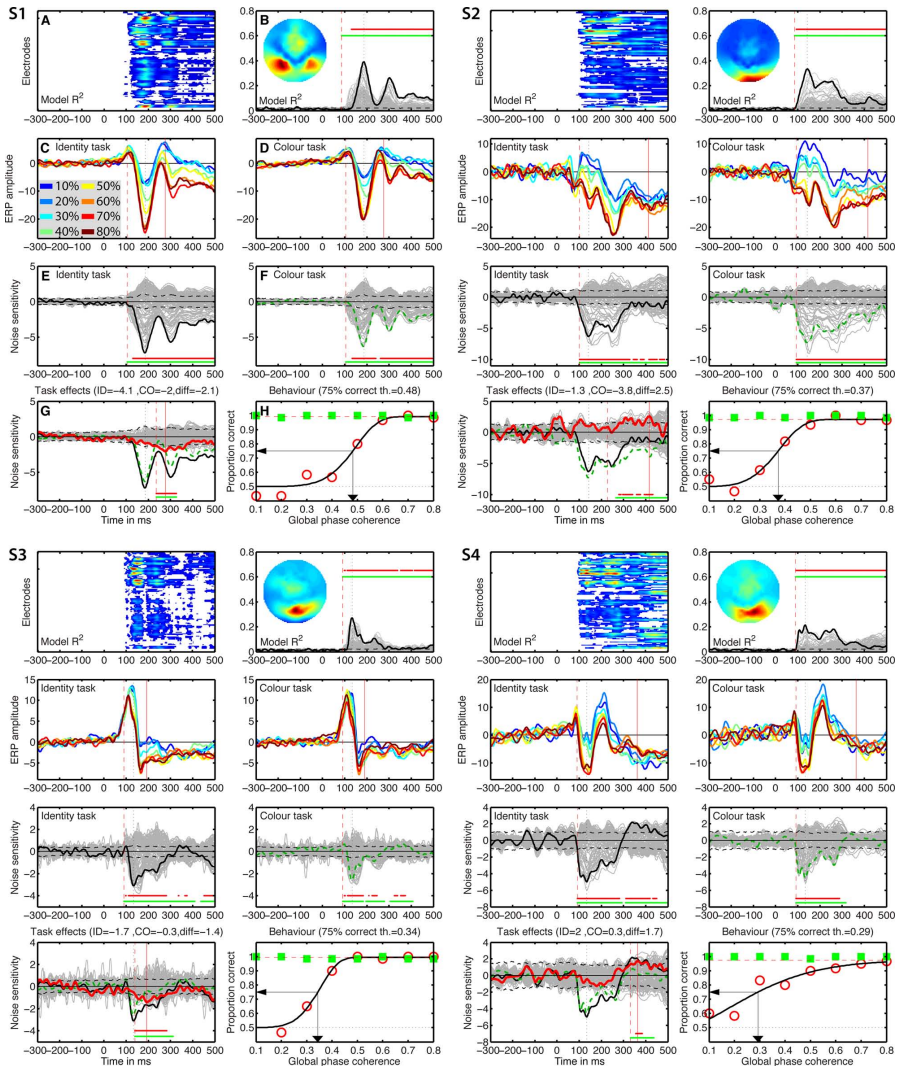


FIGURE 5 | Individual results for subjects S1 to S4. (A) Statistically significant model R^2 at all the electrodes and time frames from –300 to 500 ms after stimulus onset. Electrodes are stacked up along the *y*-axis. The tick on the *y*-axis marks the electrode at which the maximum R^2 was recorded. This electrode is plotted as a continuous black line in **(B)**. R^2 ranges from near zero in blue to the maximum for that subject in red. **(B)** Model R^2 at all the electrodes and time frames. The electrode at which the maximum R^2 was recorded is plotted in black. The other electrodes are plotted in gray. The inset shows the topographic distribution of the R^2 at the latency of the maximum, indicated by a vertical black dotted line. This vertical line is also plotted in all the other panels for comparisons. For subject S_1, R^2 had a bilateral occipital–lateral distribution, with a maximum over the left hemisphere (left bottom red cluster). The red vertical dashed line indicates the time frame of the earliest significant R^2 across all electrodes. Near the top of the panel, the upper horizontal line (red) marks significant time frames at the maximum R^2 electrode. The lower horizontal line (green) marks significant time frames of the spatial–temporal cluster to which the maximum R^2 electrode belonged. For subject S_1, this horizontal line starts at the latency of the earliest significant model fit (red vertical dashed line), indicating that the maximum R^2 electrode is part of a spatial–temporal cluster that captures the earliest effects. The horizontal dashed line is the univariate one-sided 95% CI of the R^2 under H0, at the maximum R^2 electrode. Although this is for illustration only, because the actual statistical test was based on spatial–temporal clusters, it gives a good indication of the R^2 values expected by chance. **(C)** Mean ERPs in the identity task. The red vertical dashed line indicates the time frame of the earliest significant noise sensitivity in the identity task, across all electrodes. This line is also plotted in **(E)**. The red continuous vertical line indicates the latency of the maximum task difference and is also plotted in **(D,G)**. **(D)** Mean ERPs in the color task. The red vertical dashed line indicates the time frame of the earliest significant noise sensitivity in the color task, across all electrodes. This line is also plotted in **(F)**. **(E)** Noise sensitivity beta coefficients in the identity task. Noise sensitivity at the maximum R^2 electrode is plotted in black, the other electrodes in gray. Units are μV/std of the predictor. Near the bottom of the panel, the upper horizontal line (red) marks significant noise sensitivity time frames at the maximum R^2 electrode. The lower horizontal line (green) marks significant noise sensitivity time frames of the spatial–temporal cluster to which the maximum R^2 electrode belonged. The black horizontal dashed lines show the univariate two-sided 95% confidence interval of noise sensitivity under H0, at the maximum R^2 electrode. **(F)** Noise sensitivity beta coefficients in the color task. Noise sensitivity at the maximum R^2 electrode is plotted as a green dashed line, the other electrodes in gray. **(G)** Noise sensitivity beta coefficient task differences. Noise sensitivity differences at the maximum R^2 electrode are plotted as a thick red line, the other electrodes in gray. The black continuous line and the green dashed line are the same as those in **(E,F)**. The red continuous vertical line indicates the latency of the maximum task difference. At that latency, the title indicates the amplitude of the noise sensitivity in the identity task (ID), in the color task (CO), and the difference between the two tasks (diff). **(H)** Proportion correct as a function of global phase coherence, in red circles for the identity task, in green squares for the color task. Data from the identity task were fitted with a cumulative Weibull function. The vertical arrow points to the 75% correct threshold in the identity task. The threshold appears in bracket in the title. The red horizontal dashed line marks the maximum proportion correct in the identity task obtained from the fit.

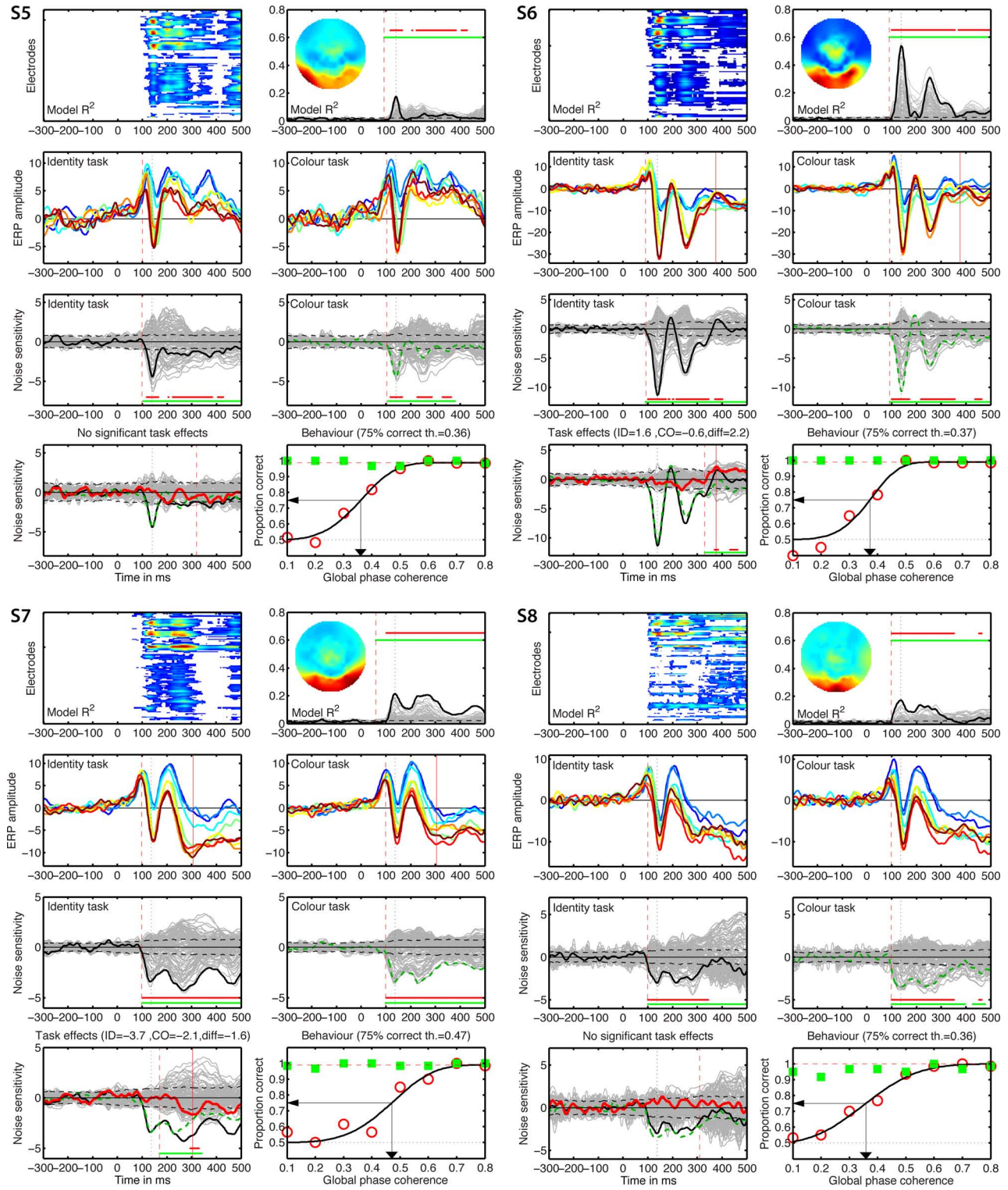


FIGURE 6 | Individual results for subjects *S5–S8*. See **Figure 5** caption for details.

Single-trial results can be misleading too. The results at the group defined best electrode showed significant effects of longer duration than results at the R^2 optimized electrode (**Figure 9**). However, results at the group defined best electrode mix together electrodes that do not necessarily provide the best fit in all subjects. This means that some of the effects at this electrode are

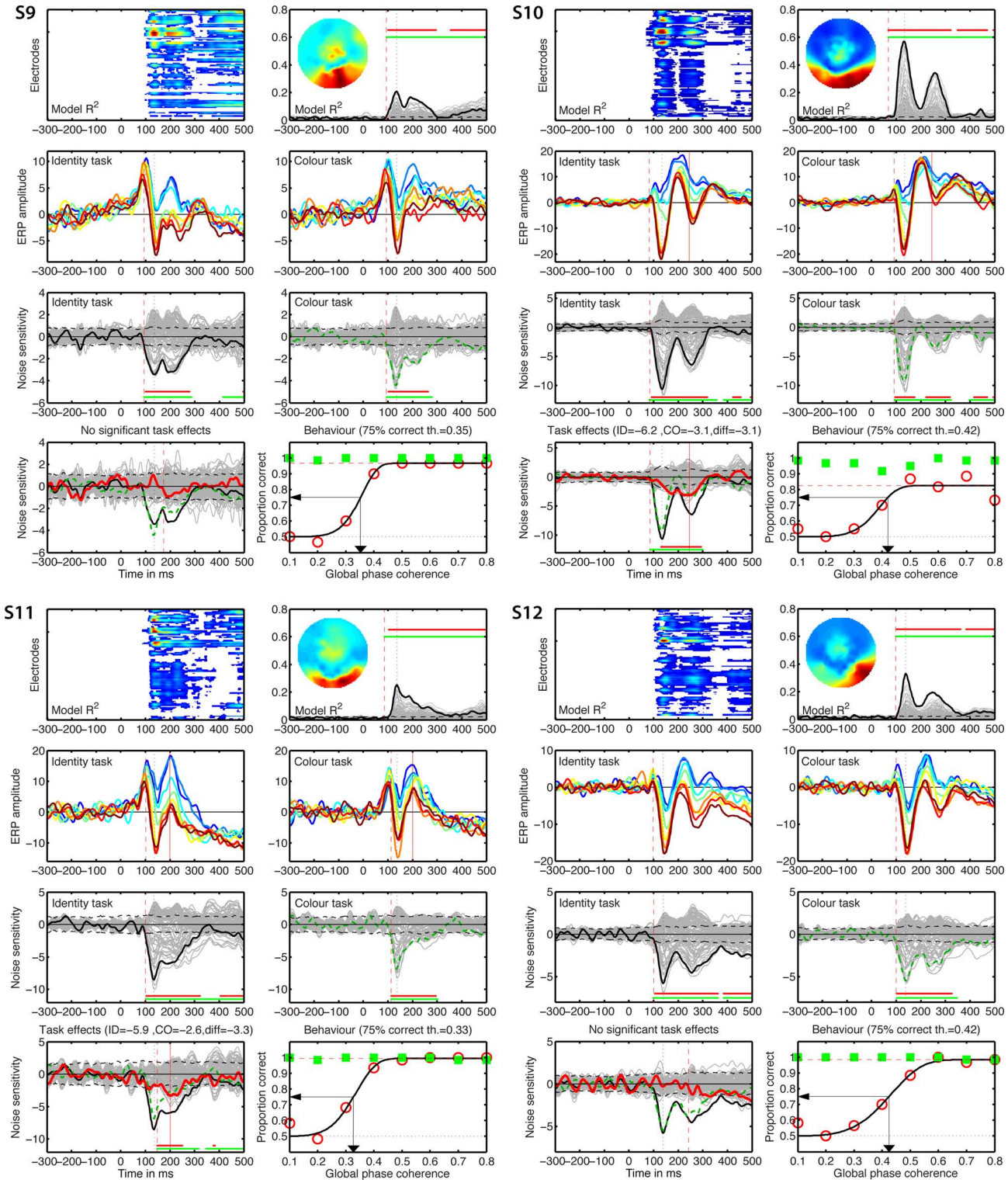


FIGURE 7 | Individual results for subjects *S9*–*S12*. See **Figure 5** caption for details.

not as meaningful as the results observed at the best fitting electrodes. In particular, the late task modulations observed in the left column of **Figure 9**, correspond to clusters of electrodes and periods of time showing noise sensitivity much weaker than that observed in the 100- to 300-ms time-window at the R^2 optimized electrodes.

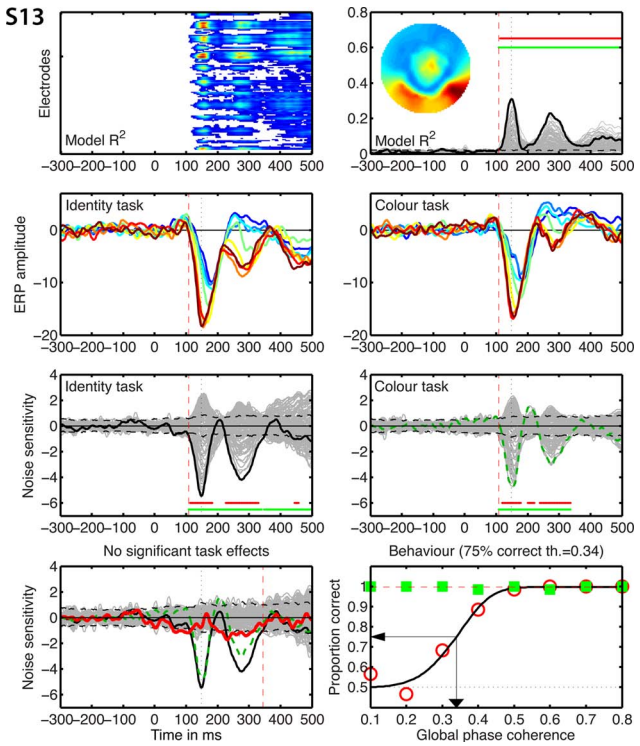


FIGURE 8 | Individual results for subject *S*13. See **Figure 5** caption for details.

In addition to the analyses performed independently at each time point (**Figures 5–8**), we also provide a continuous measure of integration time in the two tasks. This was achieved by measuring the time it takes to integrate 50% of noise sensitivity during the first half-second following stimulus onset (Rousselet et al., 2010). Noise sensitivity in the two tasks was normalized by the maximum absolute noise sensitivity in the identity task, defined across time frames. Then the absolute noise sensitivity in each task was integrated over time (**Figure 10**). At the group level, noise sensitivity integration increased sharply after 100 ms and started to differ significantly between the two tasks at 227 ms after stimulus onset. The 50% integration threshold was reached 16 ms earlier in the identity task compared to the color task. About 14% less noise sensitivity was integrated in the color task relative to the identity task. Analyses performed in each subject individually provided a somewhat different picture. In keeping with group analyses, cumulated noise sensitivity started to rise at about 100 ms in most subjects. However, onset of task effects, 50% integration times and total cumulated noise sensitivity differed markedly across subjects and from the group analyses (**Figure 10**).

Given the discrepancy between group and individual subject analyses, it is important to consider weak power as a potential explanation for the absence of task effect in some subjects. Indeed, lack of significant effects might be due to a real absence of effects or the presence of relatively weak effects that our statistical test might miss. Although lack of power cannot be completely ruled out, it appears that subjects with significant task effects at the R^2 optimized electrode had substantial effect sizes, with maximum F cluster sums at least 1.6 times larger than the largest bootstrap F cluster sums obtained by chance (**Figure 11**: $S_1 = 2.7$, $S_2 = 2.6$, $S_3 = 3.2$, $S_6 = 1.9$, $S_7 = 2.1$, $S_{10} = 2.3$, $S_{11} = 2.4$). One subject had lower effect size than the other subjects, with a cluster sum 0.8 larger than that obtained by chance (**Figure 11**: S_4). Subjects with no significant task effects had no significant cluster whatsoever (S_5), relatively low cluster sums (S_9 and S_{12}) or cluster sums so low that they fell at the bottom of the bootstrap cluster sum distributions (S_8 and S_{13}).

SHIFT FUNCTION ANALYSES OF THE DECILES OF THE SINGLE-TRIAL ERP DISTRIBUTIONS

Changes in task constraints could affect noise sensitivity by modulating preferentially single-trial ERPs to noise textures or to faces. Alternatively, these changes could be a uniform compression or expansion of the distribution. In our design, noise levels are artificial. Therefore, we studied the nature of the task effects using the shift function, a technique that assumes data follow a continuum. The shift function analyses revealed that the modulation in noise sensitivity in the color task could be attributed to a modulation of a particular type of stimuli. In five subjects (**Figure 12**: S_1, S_3, S_7, S_{10}, S_{11}), noise sensitivity reduction in the color task was due primarily to increased amplitudes of face ERPs, which became closer to that of noise trials. In two subjects (**Figure 12**: S_4 and S_6), noise sensitivity reduction was due mostly to an increase in amplitude of the noise trials. Finally, in the only subject who showed increased noise sensitivity in the color task (S_2), the effect was also due to an amplitude increase of ERPs to noise trials. In addition, in S_2, S_4, S_6, S_7, and S_{10}, there was an overall increase in ERP amplitude in the color task compared to the identity task. Thus, task constraints had non-uniform effects on ERP distributions, with most modulations being an increase in amplitude of face trials.

DISCUSSION

Using identical colored face stimuli in two tasks, and a parametric noise manipulation, we observed a significant reduction in ERP noise sensitivity when noise level was task irrelevant. Overall, following (Philiastides et al., 2006) we conclude that task effects on noise sensitivity are weak before 200 ms, in the window of the N170, and mainly present around the P2. However, task effects were highly variable across subjects, with individual differences in onsets, durations and effect sizes. These idiosyncrasies will need to be addressed in future studies.

Based on the work of (Philiastides et al., 2006), we tested the hypothesis that there is a clear boundary, at about 200 ms after stimulus onset, between bottom-up face processing and brain activity that depends on task demand. Our group results were qualitatively similar to those of Philiastides et al. (2006) with weak task effects before 200 ms and stronger differences beyond 200 ms. Changing the task requirements did not abolish noise sensitivity altogether, but reduced it by about 19–46% in individual subjects. Results were also inconsistent across subjects, with a minority of subjects showing effects consistent with group analyses, and several subjects showing no significant effects whatsoever, despite similar behavioral performances. The nature of the task effects also differed among subjects, as revealed by analyses of single-trial ERP distributions. These results points to the existence of idiosyncratic modulations of brain activity depending on task requirements.

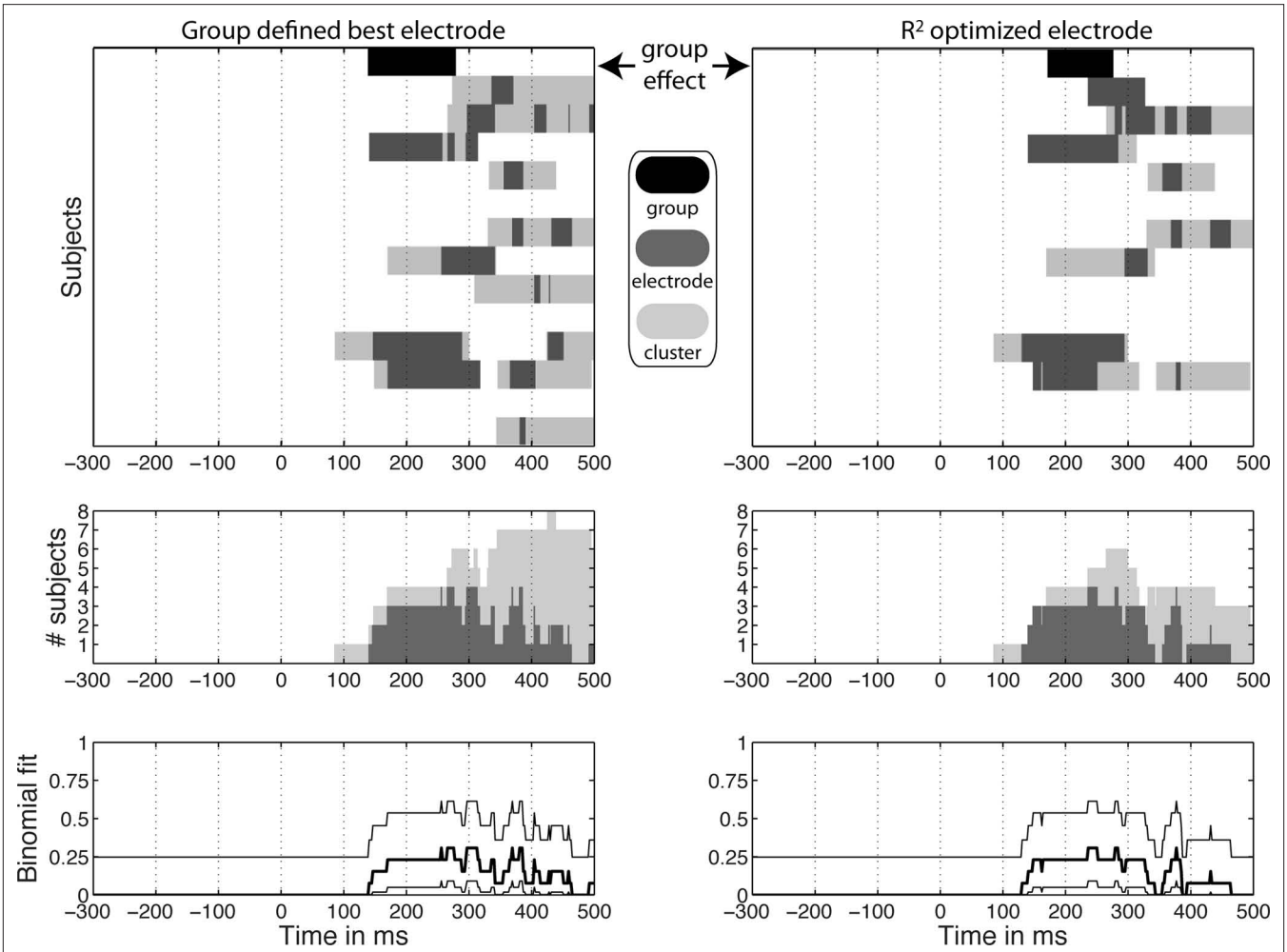


FIGURE 9 | Significant task effects at the group level and in individual subjects. In the top graphs, significant group effects appear at the top of each column in black, above significant results in individual subjects in gray. Dark gray indicates significant effects at the group defined best electrode (left column) and at the R^2 optimized electrode (right column). Light gray indicates significant effects at the spatial–temporal cluster to which an electrode belonged. The middle graphs show the number of subjects with a significant task modulation at each time point at one electrode (light gray) or in the cluster to which it belonged (dark gray). Remember that subject 2 was the only one with stronger noise sensitivity in the color task compared to the identity task. The bottom graphs show the percentage of subjects showing a significant task modulation at one electrode (thick black line) with a 95% binomial confidence interval around it (thin black lines). This percentage provides an indication of the probability of observing a group difference given the single-subject results. These binomial confidence intervals place an upper limit of about 61% probability of observing a group effect given the single-subject data. In comparison, the group results presented at the top of **Figure 9** depend on the probability of the data under the null-hypothesis.

It remains unclear whether the task dependent noise sensitivity we observed is related to differences in task difficulty between the color and the identity tasks, or if it is due to changes in the diagnosticity of stimulus phase information (Banko et al., 2011; Philiastides et al., 2006; Philiastides and Sajda, 2007). More generally, noise sensitivity between 100 and 300 ms after stimulus onset probably reflects the activity of object and face processing areas that are sensitive to stimulus evidence (Philiastides and Sajda, 2007; Rousselet et al., 2008b; Tjan et al., 2006). Noise sensitivity however does not reflect activity from a general discrimination mechanism because it is not present for color and identity. Similarly, Philiastides et al. (2006) found that early single-trial visual ERPs did not discriminate between red and green or between two motion directions. However, these other stimulus dimensions can be studied by using different techniques, such as adaptation (Vizioli et al., 2010), ICA and filtering (Snyder and Foxe, 2010), and frequency tagging (Quigley et al., 2010).

Contrary to several ERP studies described in the introduction, some of our subjects did show moderate task modulations in the time-window of the N170. The absence of task effects in previous face ERP studies is difficult to interpret because of the use of group statistics. One thing that most studies have in common is the use of relatively high-contrast stimuli. Because the effect of attention is contrast dependent (Reynolds and Heeger, 2009), attention effects on face ERPs might be more likely to be observed at low-contrast. A systematic study of attention modulations as a function of face contrast remains to be

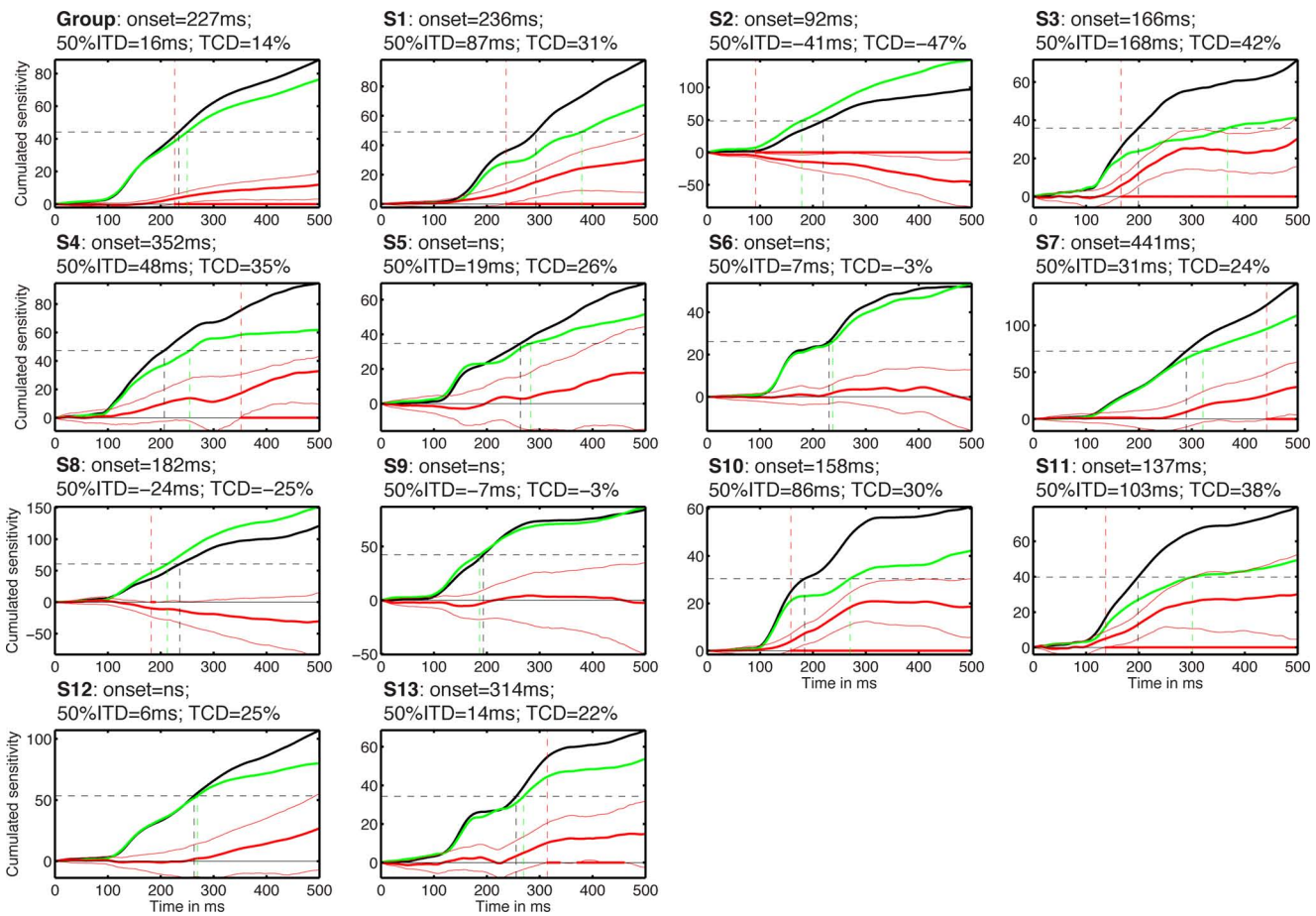


FIGURE 10 | Cumulated task effects. The subject number is indicated by S#. Each cell shows the cumulated normalized sensitivity in the identity task (black) and in the color task (green). The difference between the two tasks is shown with thick red lines, with a 95% confidence interval around it (thin red lines). Red dots along the zero horizontal line mark time points of significant task differences, with no correction for multiple comparisons. The vertical red dashed line that crosses the entire cell marks the onset of significant task effects. The horizontal black dashed line marks the value corresponding to 50% of the total cumulated sensitivity in the identity task. The two vertical lines that originate from the 50% line and terminate on the *x*-axis mark the time to reach that 50% value in the two tasks. The title of each cell contains the onset of the task effects; the 50% integration time difference (50% ITD) between the color and the identity tasks; the task cumulated difference (TCD) between the identity and the color tasks, expressed in proportion of the maximum cumulated sensitivity in the identity task.

carried out. In our experiment, contrast and luminance were quite low, suggesting even weaker task modulations in more realistic circumstances. Another problem with previous reports of null effects in group statistics is the absence of statistical analyses in individual subjects, as well as poor data description (Rousselet and Pernet, 2011). Group statistics pull out effects that are consistent across subjects, even when these effects are not significant in individual subjects. Although this might seem like a good property, given the number of face ERP experiments carried out each year and the common belief that it is satisfactory to test more subjects to achieve significance, group statistics might be responsible for many false positives in the literature (Wagenmakers, 2007). By definition, group statistics are also insensitive to single-subject significant effects that are inconsistent across subjects, for instance if timing and topographies differ. Thus, group statistics can be misleading because of increased chances of false positives and false negatives, at least in theory. Yet, it is at the moment difficult to evaluate if our results, showing a large discrepancy between group and single-trial analyses constitute a unique curiosity or if our results reveal a pervasive problem in the ERP literature. Indeed, typical face ERP studies are mostly concerned with group statistics of peak measurements, with little concern for reliability and quantification of the effects. In fact, most studies are content with the discussion of any effect $p < 0.05$ (Rousselet and Pernet, 2011). Current practice in the ERP literature tends to hide the rich inter-subject variability that we ought to explain: we perform perceptual tasks as individuals, not as a collective brain. Moreover, many studies report weak effect sizes, unexpected results and do not control for multiple comparisons properly. One is left wondering what proportion of ERP results will ever be replicated (Miller, 2009). In many studies, beyond the recurrent fundamental flaws of null-hypothesis significance testing (Wagenmakers, 2007), the lack of robustness of *t*-tests and ANOVAs, and the lack of proper

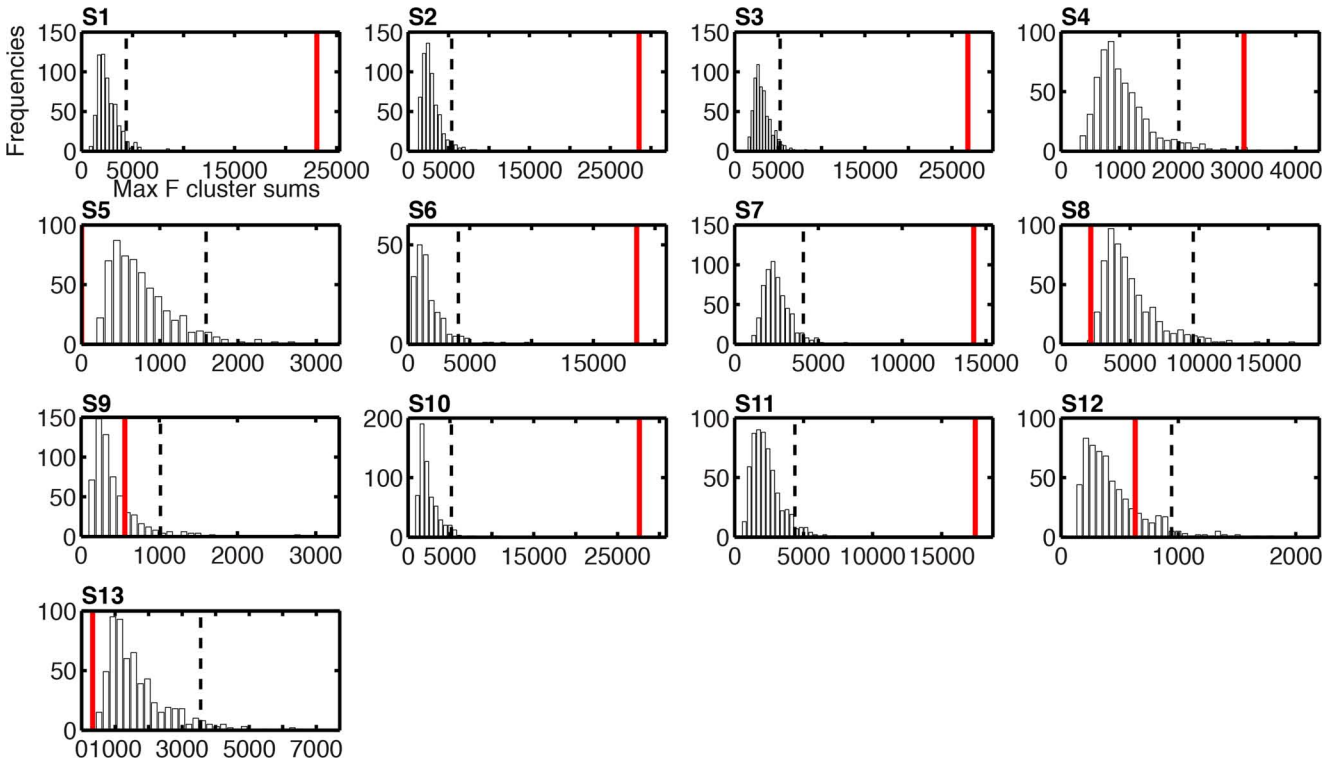


FIGURE 11 | Histograms of the bootstrap distributions of maximum *F* cluster sums for task effects on single-trial ERP noise sensitivity. The subject number is indicated by S# in bold font. These bootstrap distributions were calculated under the null-hypothesis H0, as described in Methods; hence they reflect the size of spatial–temporal task effects that can be expected by chance, due to random sampling, across the entire search space. For each subject, the vertical black dashed line marks the 95th percentile of the H0 bootstrap distribution. The vertical red continuous line indicates the maximum sum of *F* values across the spatial–temporal clusters that contained the maximum R^2 electrode. For subject S_5, the cluster sum is equal to zero because no cluster passed the two-electrode threshold: they were present at this electrode only.

control for multiple comparisons (Wilcox, 2005), readers are too often left with so little evidence that it is impossible to judge the importance of their results. Here we've tried to provide a richer set of descriptions than is usually available in face ERP papers.

Of course, the lack of significant task effects in some subjects, and the lack of consistency across subjects who did show significant effects, might be attributable to different sources of variance, including differences in scalp thickness and electrode application, rather than individual differences in visual processing. These differences could lead to differences in statistical power across subjects. Although subjects who did show task effects had relatively large effect sizes, it is possible that more trials or better regression analyses, or both, would be necessary to reveal significant effects at different time points and in subjects showing null results. We are exploring the possibility of using smooth variance estimators, weighted models, and adjusting statistical thresholds based on empirical distributions to increase statistical power. However, our data driven estimates of effects expected by chance suggest that some subjects had indeed no task modulation of noise sensitivity whatsoever (**Figure 11**).

Finally, null or inconsistent effects might not reveal the absence of an effect but our failure to quantify changes in a multidimensional space. For instance, we reported most of our single-trial analyses at one electrode only. Based on extensive inspection of the current dataset and previous datasets, the electrode at which the model provides the best fit seems to capture most of the effects (Rousselet et al., 2008b, 2009, 2010). Because of spatial blurring, neighboring electrodes contain redundant information, so pooling results across electrodes as is often done in group analyses would be of no benefit for univariate single-subject analyses. Of course, there might be extra information available in a multivariate space containing a large number of electrodes (Philiastides and Sajda, 2006). Hence, it will be worth extending our univariate model to measure multivariate relationships between single-trial ERP amplitude, stimulus evidence, and task demand. Finally, as discussed by (Liu et al., 2009), a single-trial linear classifier has the advantage over a GLM approach to provide a measure of information. However, it is not clear how linear classifiers can be applied to more complicated designs such as our ANCOVA.

To conclude, all these considerations about group and single-subject analyses are rather circular, because it is not clear what ought to be found. The ERP community relies mostly on group analyses, and therefore most readers might be biased to conclude that discrepancies between group and single-subject analyses reflect problems in single-subject analyses. This point of view is misguided because a significant group effect does not provide any guarantee that even 50% of the subjects will show the group effect (**Figure 9**).

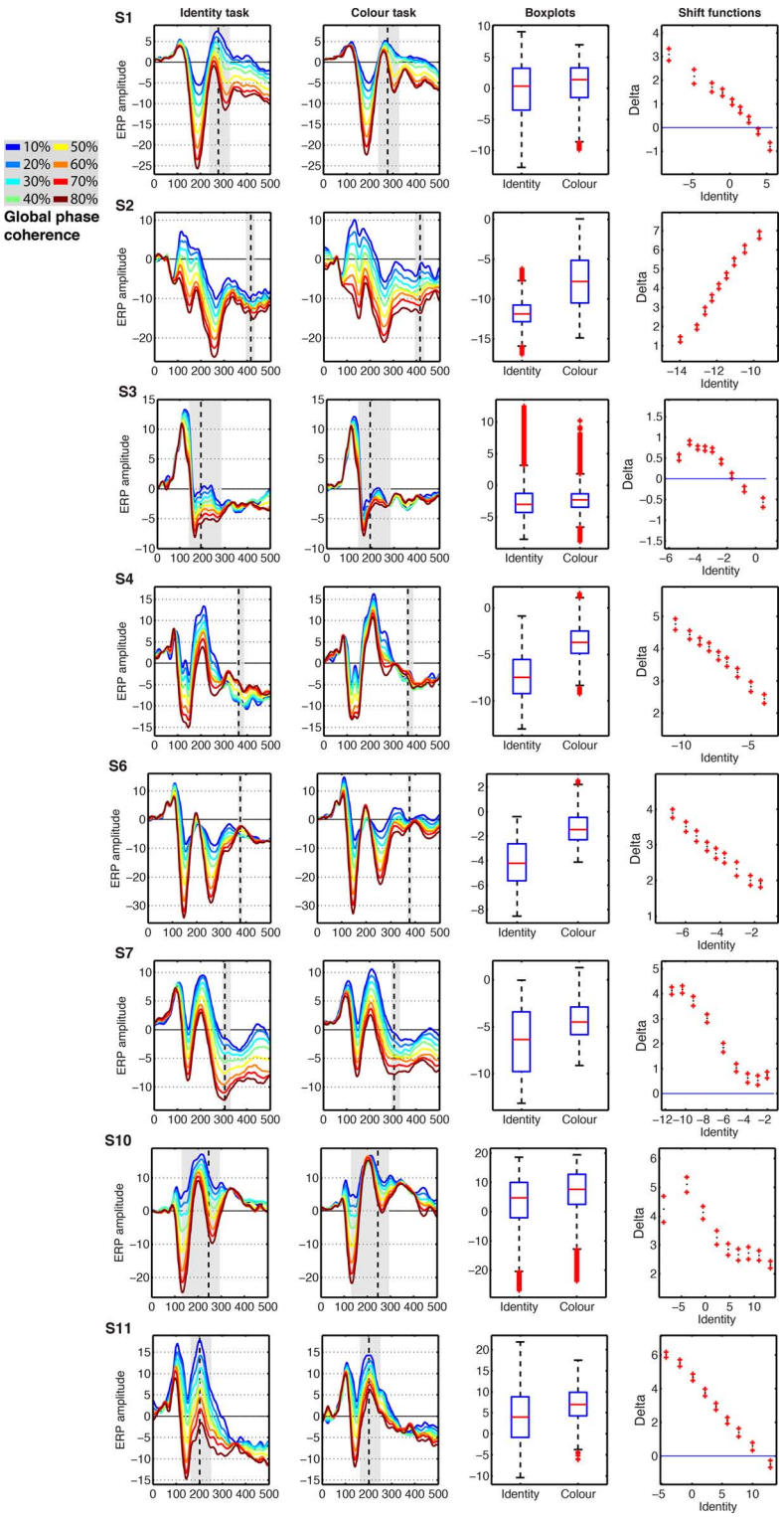


FIGURE 12 | Event-related potentials results from the eight subjects showing significant task effects. The subject number is indicated by S# in bold font. The first two columns show the modeled ERPs in the identity and the color tasks. The vertical dashed line marks the latency of the largest task difference. The vertical gray shaded area marks all the continuous time frames at which a significant effect was observed, and which contained the time frame of maximum effect. The third column shows boxplots of the single-trial modeled $ERP_{(t,ei)}$ amplitudes in the identity and color tasks, summed across the time frames marked by gray areas in columns 1 and 2. The fourth column shows the shift function between the distributions in column three. The *x*-axis shows the estimated deciles in the identity task. The *y*-axis shows the estimated difference deciles between the identity and the color task, marked as nine dots, with the ends of the confidence intervals marked by plus signs.

The inter-individual differences we describe in detail in this paper might well be differences in visual processing. If this is true, these differences must be explained, and should not be simply smoothed out by using group statistics.

CONCLUSION

At the group level, ERP sensitivity to phase noise was reduced between about 140 and 300 ms when stimulus phase information was task irrelevant. This result suggests that sensitivity to image structure can be modulated by the task at hand, but only after an initial period of essentially bottom-up visual processing, from about 90 to 140 ms. In contrast to what the group results might suggest, we observed a significant task effect in only 60% of subjects, and at any time point only 31% of subjects showed results consistent with group analyses. Interestingly, the origin of the reduction in ERP sensitivity to phase information differed among subjects, being due to an increase in ERP amplitude to the noisiest stimuli (noise textures) in some subjects and an increase in ERP amplitude to the least noisy stimuli (faces) in others. Overall, our analyses demonstrate the usefulness of single-trial analyses and parametric designs to study information processing. Our results also suggest that, in some situations, group statistics can be so misleading that their use, without complementary individual subject analyses, is questionable.

ACKNOWLEDGMENTS

ESRC grant RES-000-22-3209 supported Guillaume A. Rousselet. Cyril R. Pernet is funded by the SINAPSE collaboration – http://www.sinapse.ac.uk, a pooling initiative funded by the Scottish Funding Council and the Chief Scientist Office of the Scottish Executive. A Biomedical Vacation Scholarship from the Wellcome Trust and a grant from the University of Glasgow Settlement charity supported Kacper P. Wieczorek. We thank Sarah Driscoll for her help collecting data.

REFERENCES

Aranda, C., Madrid, E., Tudela, P., and Ruz, M. (2010). Category expectations: a differential modulation of the N170 potential for faces and words. *Neuropsychologia* 48, 4038–4045.

Banko, E. M., Gal, V., Kortvelyes, J., Kovacs, G., and Vidnyanszky, Z. (2011). Dissociating the effect of noise on sensory processing and overall decision difficulty. *J. Neurosci.* 31, 2663–2674.

Berkovits, I., Hancock, G. R., and Nevitt, J. (2000). Bootstrap resampling approaches for repeated measure designs: relative robustness to sphericity and normality violations. *Educ. Psychol. Meas.* 60, 877–892.

Bötzel, K., and Grüsser, O.-J. (1989). Electric brain potentials evoked by pictures of faces and non-faces: a search for "face-specific" EEG-potentials. *Exp. Brain Res.* 77, 349–360.

Carmel, D., and Bentin, S. (2002). Domain specificity versus expertise: factors influencing distinct processing of faces. *Cognition* 83, 1–29.

Crist, R. E., Wu, C.-T., Karp, C., and Woldorff, M. G. (2008). Face processing is gated by visual spatial attention. *Front. Hum. Neurosci.* 1:10. doi: 10.3389/neuro.09.010.2007

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

Delorme, A., Sejnowski, T., and Makeig, S. (2007). Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. *Neuroimage* 34, 1443–1449.

Eimer, M. (2000). Effects of face inversion on the structural encoding and recognition of faces. Evidence from event-related brain potentials. *Brain Res. Cogn. Brain Res.* 10, 145–158.

Foxe, J. J., and Simpson, G. V. (2002). Flow of activation from V1 to frontal cortex in humans. A framework for defining "early" visual processing. *Exp. Brain. Res.* 142, 139–150.

Furey, M. L., Tanskanen, T., Beauchamp, M. S., Avikainen, S., Uutela, K., Hari, R., and Haxby, J. V. (2006). Dissociation of face-selective cortical responses by attention. *Proc. Natl. Acad. Sci. U.S.A.* 103, 1065–1070.

Gaspar, C. M., Rousselet, G. A., and Pernet, C. R. (in press). Reliability of ERP and single-trial analyses. *NeuroImage*.

Gaspar, C. M., and Rousselet, G. A. (2009). How do amplitude spectra influence rapid animal detection? *Vision Res.* 49, 3001–3012.

Goffaux, V., Jemel, B., Jacques, C., Rossion, B., and Schyns, P. G. (2003). ERP evidence for task modulations on face perceptual processing at different spatial scales. *Cogn. Sci.* 27, 313–325.

Gold, J., Bennett, P. J., and Sekuler, A. B. (1999). Identification of band-pass filtered letters and faces by human and ideal observers. *Vision Res.* 39, 3537–3560.

Husk, J. S., Bennett, P. J., and Sekuler, A. B. (2007). Inverting houses and textures: investigating the characteristics of learned inversion effects. *Vision Res.* 47, 3350–3359.

Jacques, C., and Rossion, B. (2008). Electrophysiological evidence for temporal dissociation between spatial attention and sensory competition during human face processing. *Cereb. Cortex* 17, 1055–1065.

Jeffreys, D. A. (1989). A face-responsive potential recorded from the human scalp. *Exp. Brain Res.* 78, 193–202.

Jeffreys, D. A., and Tukmachi, E. S. (1992). The vertex-positive scalp potential evoked by faces and by objects. *Exp. Brain Res.* 91, 340–350.

Jeffreys, D. A., Tukmachi, E. S., and Rockley, G. (1992). Evoked potential evidence for human brain mechanisms that respond to single, fixated faces. *Exp. Brain Res.* 91, 351–362.

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

Kovesi, P. (1999). Image features from phase congruency. *Videre J. Comput. Vis. Res.* 1, 2–26.

Kovesi, P. (2003). "Phase congruency detects corners and edges," in *Paper presented at the The Australian Pattern Recognition Society Conference: DICTA 2003*, Sydney.

Landau, A. N., Aziz-Zadeh, L., and Ivry, R. B. (2010). The influence of language on perception: listening to sentences about faces affects the perception of faces. *J. Neurosci.* 30, 15254–15261.

Liu, H., Agam, Y., Madsen, J. R., and Kreiman, G. (2009). Timing, timing, timing: fast decoding of object information from intracranial field potentials in human visual cortex. *Neuron* 62, 281–290.

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

Lueschow, A., Sander, T., Boehm, S. G., Nolte, G., Trahms, L., and Curio, G. (2004). Looking for faces: attention modulates early occipitotemporal object processing. *Psychophysiology* 41, 350–360.

Makeig, S., Delorme, A., Westerfield, M., Jung, T. P., Townsend, J., Courchesne, E., and et al. (2004). Electroencephalographic brain dynamics following manually responded visual targets. *PLoS Biol.* 2, e176. doi: 10.1371/journal.pbio.0020176

Maris, E., and Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* 164, 177–190.

Miller, J. (2009). What is the probability of replicating a statistically significant effect? *Psychon. Bull. Rev.* 16, 617–640.

Mohamed, T. N., Neumann, M. F., and Schweinberger, S. R. (2009). Perceptual load manipulation reveals sensitivity of the face-selective N170 to attention. *Neuroreport* 20, 782–787.

Nobre, A. C., Allison, T., and McCarthy, G. (1998). Modulation of human extrastriate visual processing by selective attention to colours and words. *Brain* 121, 1357–1368.

Okazaki, Y., Abrahamyan, A., Stevens, C. J., and Ioannides, A. A. (2008). The timing of face selectivity and attentional modulation in visual processing. *Neuroscience* 152, 1130–1144.

Pernet, C., Schyns, P. G., and Demonet, J. F. (2007). Specific, selective or preferential: comments on category specificity in neuroimaging. *Neuroimage* 35, 991–997.

Pernet, C. R., Chauveau, N., Gaspar, C., and Rousselet, G. A. (2011). LIMO EEG: a toolbox for hierarchical LInear MOdeling of EletroEncephaloGraphic data. *Comput. Intell. Neurosci.* [Article ID 831409]. doi: 10.1155/2011/831409

Philiastides, M. G., Ratcliff, R., and Sajda, P. (2006). Neural representation of task difficulty and decision making during perceptual categorization: a timing diagram. *J. Neurosci.* 26, 8965–8975.

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

Philiastides, M. G., and Sajda, P. (2007). EEG-informed fMRI reveals spatiotemporal characteristics of perceptual decision making. *J. Neurosci.* 27, 13082–13091.

Puce, A., Allison, T., and McCarthy, G. (1999). Electrophysiological studies of human face perception. III: effects of top-down processing on face-specific potentials. *Cereb. Cortex* 9, 445–458.

Quigley, C., Andersen, S. K., Schulze, L., Grunwald, M., and Müller, M. M. (2010). Feature-selective attention: evidence for a decline in old age. *Neurosci. Lett.* 474, 5–8.

Ratcliff, R., Philiastides, M. G., and Sajda, P. (2009). Quality of evidence for perceptual decision making is indexed by trial-to-trial variability of the EEG. *Proc. Natl. Acad. Sci. U.S.A.* 106, 6539–6544.

Reynolds, J. H., and Heeger, D. J. (2009). The normalization model of attention. *Neuron* 61, 168–185.

Rossion, B., and Jacques, C. (2008). Does physical interstimulus variance account for early electrophysiological face sensitive responses in the human brain? Ten lessons on the N170. *Neuroimage* 39, 1959–1979.

Rousselet, G. A., Gaspar, C. M., Pernet, C. R., Husk, J. S., Bennett, P. J., and Sekuler, A. B. (2010). Healthy aging delays scalp EEG sensitivity to noise in a face discrimination task. *Front. Psychology* 1:19. doi: 10.3389/fpsyg.2010.00019

Rousselet, G. A., Husk, J. S., Bennett, P. J., and Sekuler, A. B. (2007a). Single-trial EEG dynamics of object and face visual processing. *Neuroimage* 36, 843–862.

Rousselet, G. A., Macé, M. J.-M., Thorpe, S. J., and Fabre-Thorpe, M. (2007b). Limits of ERP differences in tracking object processing speed. *J. Cogn. Neurosci.* 19, 1–18.

Rousselet, G. A., Husk, J. S., Bennett, P. J., and Sekuler, A. B. (2008a). Time course and robustness of ERP object and face differences. *J. Vis.* 8:3, 1–18.

Rousselet, G. A., Pernet, C. R., Bennett, P. J., and Sekuler, A. B. (2008b). Parametric study of EEG sensitivity to phase noise during face processing. *BMC Neuroscience* 9, 98. doi: 10.1186/1471-2202-9-98

Rousselet, G. A., Husk, J. S., Pernet, C. R., Gaspar, C. M., Bennett, P. J., and Sekuler, A. B. (2009). Age-related delay in information accrual for faces: evidence from a parametric, single-trial EEG approach. *BMC Neurosci.* 10, 114. doi: 10.1186/1471-2202-10-114

Rousselet, G. A., and Pernet, C. R. (2011). Quantifying the time course of visual object processing using ERPs: it's time to up the game. *Front. Psychology* 2:107. doi: 10.3389/fpsyg.2011.00107

Rousselet, G. A., Thorpe, S. J., and Fabre-Thorpe, M. (2004). How parallel is visual processing in the ventral pathway? *Trends Cogn. Sci.* 8, 363–370.

Schyns, P. G. (1998). Diagnostic recognition: task constraints, object information, and their interactions. *Cognition* 67, 147–179.

Schyns, P. G. (2010). Grand challenges in perception science: modeling the future. *Front. Psychology* 1:10. doi: 10.3389/fpsyg.2010.00010

Schyns, P. G., Gosselin, F., and Smith, M. L. (2009). Information processing algorithms in the brain. *Trends Cogn. Sci. (Regul. Ed.)* 13, 20–26.

Schyns, P. G., Jentzsch, I., Johnson, M., Schweinberger, S. R., and Gosselin, F. (2003). A principled method for determining the functionality of brain responses. *Neuroreport* 14, 1665–1669.

Schyns, P. G., Petro, L. S., and Smith, M. L. (2007). Dynamics of visual information integration in the brain for categorizing facial expressions. *Curr. Biol.* 17, 1580–1585.

Seco, G. V., Izquierdo, M. C., García, M. P. F., and Díez, F. J. H. (2006). A comparison of the Bootstrap-F, improved general approximation, and Brown-Forsythe multivariate approaches in a mixed repeated measures design. *Educ. Psychol. Meas.* 66, 35–62.

Seeck, M., and Grüsser, O.-J. (1992). Category-related components in visual evoked potentials: photographs of faces, persons, flowers and tools as stimuli *Exp. Brain Res.* 92, 338–349.

Séverac-Cauquil, A., Edmonds, G. E., and Taylor, M. J. (2000). Is the face-sensitive N170 the only ERP not affected by selective attention? *Neuroreport* 11, 2167–2171.

Smith, M. L., Gosselin, F., and Schyns, P. G. (2007). From a face to its category via a few information processing states in the brain. *Neuroimage* 37, 974–984.

Snyder, A. C., and Foxe, J. J. (2010). Anticipatory attentional suppression of visual features indexed by oscillatory alpha-band power increases: a high-density electrical mapping study. *J. Neurosci.* 30, 4024–4032.

Sreenivasan, K. K., Katz, J., and Jha, A. P. (2007). Temporal characteristics of top-down modulations during working memory maintenance: an event-related potential study of the N170 component. *J. Cogn. Neurosci.* 19, 1836–1844.

Thorpe, S. J. (2009). The speed of categorization in the human visual system. *Neuron* 62, 168–170.

Thorpe, S. J., and Fabre-Thorpe, M. (2001). Seeking categories in the brain. *Science* 291, 260–263.

Tjan, B. S., Lestou, V., and Kourtzi, Z. (2006). Uncertainty and invariance in the human visual cortex. *J. Neurophysiol.* 96, 1556–1568.

van Rijsbergen, N. J., and Schyns, P. G. (2009). Dynamics of trimming the content of face representations for categorization in the brain. *PLoS Comput. Biol.* 5, e1000561. doi: 10.1371/journal.pcbi.1000561

Vizioli, L., Rousselet, G. A., and Caldara, R. (2010). Neural repetition suppression to identity is abolished by other-race faces. *Proc. Natl. Acad. Sci. U.S.A.* 107, 20081–20086.

Wagenmakers, E. J. (2007). A practical solution to the pervasive problems of p values. *Psychon. Bull. Rev.* 14, 779–804.

Wilcox, R. R. (2005). *Introduction to Robust Estimation and Hypothesis Testing,* 2nd Edn. New York, NY: Elsevier Academic Press.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 03 January 2011; accepted: 09 June 2011; published online: 23 June 2011. Citation: Rousselet GA, Gaspar CM, Wieczorek KP and Pernet CR (2011) Modeling single-trial ERP reveals modulation of bottom-up face visual processing by top-down task constraints (in some subjects). Front. Psychology ***2****:137. doi: 10.3389/fpsyg.2011.00137*

This article was submitted to Frontiers in Perception Science, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

REVIEW ARTICLE
published: 09 April 2011
doi: 10.3389/fpsyg.2011.00060

Ongoing EEG phase as a trial-by-trial predictor of perceptual and attentional variability

R. VanRullen[1,2]*, N. A. Busch[3,4], J. Drewes[1,2] and Julien Dubois[1,2,5]

[1] *Université de Toulouse, Centre de Recherche Cerveau et Cognition, Université Paul Sabatier, Toulouse, France*
[2] *Centre National de la Recherche Scientifique, UMR 5549, Faculté de Médecine de Rangueil, Toulouse, France*
[3] *Institute of Medical Psychology, Charité–Universitätsmedizin Berlin, Berlin, Germany*
[4] *Berlin School of Mind and Brain, Humboldt Universität, Berlin, Germany*
[5] *Division of Biology, California Institute of Technology, Pasadena, CA, USA*

Edited by:
Guillaume A. Rousselet, University of Glasgow, UK

Reviewed by:
Peter Lakatos, Hungarian Academy of Sciences, Hungary
Gregor Thut, University of Glasgow, UK

****Correspondence:***
R. VanRullen, Université de Toulouse, Centre de Recherche Cerveau et Cognition, Université Paul Sabatier, 31062 Toulouse, France.
e-mail: rufin.vanrullen@cerco.ups-tlse.fr

Even in well-controlled laboratory environments, apparently identical repetitions of an experimental trial can give rise to highly variable perceptual outcomes and behavioral responses. This variability is generally discarded as a reflection of intrinsic noise in neuronal systems. However, part of this variability may be accounted for by trial-by-trial fluctuations of the phase of ongoing oscillations at the moment of stimulus presentation. For example, the phase of an electro-encephalogram (EEG) oscillation reflecting the rapid waxing and waning of sustained attention can predict the perception of a subsequent visual stimulus at threshold. Similar ongoing periodicities account for a portion of the trial-by-trial variability of visual reaction times. We review the available experimental evidence linking ongoing EEG phase to perceptual and attentional variability, and the corresponding methodology. We propose future tests of this relation, and discuss the theoretical implications for understanding the neuronal dynamics of sensory perception.

Keywords: EEG, oscillation, phase, pre-stimulus, spontaneous, ongoing, perception, attention

INTRODUCTION

Run a computer program twice with the same inputs: chances are, you should get the same output twice. As any experimenter knows, it is not so with the human brain. This unreliable device persistently fails to provide a consistent outcome: reaction times (RTs) vary by a factor of two or more, perception sometimes gets distorted and sometimes does not occur at all – even though the external world has been carefully controlled and equated, trial after trial. This variability gets in the way of any serious scientific measurement, and therefore scientists have dubbed it "noise" and found ways to discount it, generally by considering the mean response over several hundreds of trials as the true standard of brain function. Oftentimes, however, one comes across a signal in the brain that tells a lot about the subject's perception on a given trial, or that can explain hitherto unexplained differences between individual trials. Such signals are the focus of the Special Topic to which this article belongs. More specifically, in this review we will consider situations in which the *phase* of ongoing brain oscillations (i.e., whether the oscillation is currently at its peak, its trough, or any particular point in between), even before any stimulus is actually presented to the subject, can inform us about their subsequent perception. Beyond the obvious implication that the brain has little to do with modern computers, these recent findings reveal much about its processing strategies.

OSCILLATORY PHASE INFLUENCES NEURAL RESPONSES

Neurons in the brain communicate by sending electrical pulses or "spikes," which create electric potential differences at synapses and cell bodies. In turn, these voltage differences are responsible for the opening and closing of membrane channels and the subsequent flow of electrically charged ions in and out of the cell bodies. Altogether, these processes induce large variations in local and long-distance electrical voltages at different temporal scales that can be considered as signatures of neuronal communication. These signatures are picked up by experimenters, for example using depth electrodes recording the extra-cellular potential and the local field potential (LFP), or using surface electrodes recording the electro-encephalogram (EEG).

The net effect of spikes and synaptic transmission at the level of neuronal populations often takes the form of an oscillation of the electric potential, in which the extra-cellular voltage increases and decreases at regular intervals. The responsiveness of single neurons to the same input intensity (i.e., the same number of spikes received) can vary greatly depending on the neurons' present state (i.e., their membrane potential) as well as on whether the extra-cellular voltage oscillation is in its lower or higher stage. This influence of spontaneous oscillatory phase on neuronal processing has long been recognized *in vitro* (Calvin and Stevens, 1967; Levitan et al., 1968; Stern et al., 1997) but it is only recently that the potential effects of oscillatory phase on sensory processing have started to be investigated *in vivo*. For example, Fries et al. (2001) reported that the phase of pre-stimulus gamma (40–70 Hz) oscillations in cat visual cortex determined the latency of subsequent neuronal firing. Montemurro et al. (2008) found that the precise phase of an ongoing delta (1–4 Hz) oscillation at which neurons in primary visual cortex fired carried information about the visual stimulus that could not be extracted based on firing rate alone. In fact, the firing phase within each gamma oscillatory cycle is a reliable indicator of neuronal activation (Vinck et al., 2010a). These results support previously published theories proposing that the phase of spike

firing relative to an ongoing oscillatory signal could constitute a meaningful neural coding scheme (VanRullen et al., 2005a; Fries et al., 2007).

The same relation that exists between oscillatory phase at the moment a neuron receives its inputs, and this neuron's responsiveness to those inputs, can also be observed over larger-scale neuronal populations comprising entire brain areas. Indeed, past studies have also reported an influence of pre-stimulus EEG phase on the magnitude of various subsequent event-related potential (ERP) components – which represent a sensory system's response to its visual or auditory inputs (Jansen and Brandt, 1991; Brandt, 1997; Kruglikov and Schiff, 2003). Because neuronal firing ultimately generates subjective perception, and because ERPs are often regarded as external markers of this perception, the literature reviewed so far seems to point, albeit indirectly, to a possible relation between ongoing oscillatory phase and sensory perception. The direct measurement of this relation will be the topic of this review.

Our focusing on oscillatory phase does not imply, of course, that the *amplitude* of ongoing oscillations has no impact on perception. For one thing, the phase of an oscillatory signal can only be reliably computed when this signal has significant power. This is not only true in a mathematical sense, but also at the biophysical level: if membrane potential fluctuations were not synchronous over a reasonably large population of neurons, any influence of phase existing for individual neurons would average out at the population level. Furthermore, it is well accepted now that oscillatory power in various frequency bands bears significant relations to sensory perception and attention (Klimesch, 1999; Tallon-Baudry and Bertrand, 1999; Engel and Singer, 2001; Varela et al., 2001; Ergenoglu et al., 2004; Hanslmayr et al., 2005, 2007; Thut et al., 2006; van Dijk et al., 2008). Our motivation for concentrating on ongoing oscillatory phase is, simply, that this variable has been largely overlooked, at least until recent years. Similarly, we will restrict this review to cases of truly spontaneous oscillatory activity, even though numerous recent studies in human and non-human primates have reported an entrainment of the phase of brain oscillations to rhythmic stimulus presentation sequences, accompanied by periodic modulations of perception, attention, and RTs (Large and Jones, 1999; Lakatos et al., 2008; Schroeder and Lakatos, 2009; Mathewson et al., 2010).

Measuring the influence of ongoing oscillatory phase on perception cannot be performed using conventional methods, but requires single-trial analyses, which – one way or another– relate the variability of spontaneous brain signals to the changes in response variables across trials – rather than discarding this variability by averaging across trials or computing correlations across subjects. Before surveying the recent experimental advances in this area, we shall detail a few of the data analysis methods that have been used to uncover this relation.

METHODS FOR LINKING ONGOING EEG PHASE TO TRIAL-BY-TRIAL VARIABILITY

Our general problem can be defined in the following terms (see **Figure 1**). On each trial, an experimenter records both a temporal signal (e.g., an EEG waveform) and a behavioral response that can be graded (e.g., RT, perceived intensity on a continuous scale) or have only a few discrete labels (e.g., stimulus perceived vs. not perceived, two- or multiple-alternative discrimination forced-choice; see **Figure 1A**). In certain situations, the graded behavioral response can be turned into a discrete variable by binning neighboring values (for example, labeling each RT with the corresponding quintile value, from 1 to 5). For our purposes, the temporal signal will correspond to brain activity *prior* to the presentation of the stimulus that must be perceived or classified by the observer (of course, similar methods can also be used for analysis of stimulus-evoked brain activity, but these will not be discussed here). In addition, we will assume that the experimental paradigm is designed using randomized inter-trial intervals, such that the moment of stimulus onset is unpredictable, and therefore the distribution of oscillatory phase values at or before stimulus onset is uniform across all trials. In technical terms, our main question is whether this phase distribution will significantly depart from uniformity, once the behavioral outcome is taken into account.

In practice, for a discrete behavioral variable, trials are grouped according to the behavioral response, and the uniformity of the distribution of phases is evaluated for each trial group. This last step can be done explicitly, by averaging across trials in the complex domain – with phase being represented by the angle of the complex vector (see **Figure 1C**); this was the approach used to compute results in **Figures 2A,C**. It can also be performed implicitly, in the temporal domain (see **Figure 1B**); for example, Fries et al. (2001) compared pre-stimulus LFP averages for groups of trials separated by short vs. long firing latencies; similarly, Mathewson et al. (2009) compared band-passed pre-stimulus EEG averages for perceived vs. unperceived visual stimuli. Finally, for a continuous behavioral variable, specific methods exist (Berens, 2009) that estimate the correlation between the pre-stimulus EEG phase (a circular variable) and the behavioral response (generally given on a linear scale); this was the approach used for the analysis illustrated in **Figure 2B**. There are, of course, other alternatives to measure phase dependency, but the methods listed here already cover most of those used in the existing literature.

Obviously, each of the approaches listed above also needs to be accompanied by appropriate statistics. Simple parametric tests are sometimes sufficient: for example, time-domain signals can be directly compared between two groups of trials corresponding to two distinct behavioral outcomes, using a Student's *t*-test – the null hypothesis being that the pre-stimulus means for the two conditions are equivalent at each point in time. Care must be taken, however, to correct for the increased likelihood of false positives due to the number of multiple comparisons (in the above example, each time point yields a distinct, but not necessarily independent, statistical comparison). Such correction methods (Bonferroni's, among others) are beyond the scope of this article. In addition, circular variables (e.g., phase angles expressed in the complex domain) are highly non-linear (for example, the mean of two angles of 10° and 350° does not correspond to the arithmetic mean of 180; instead, the circular mean of these two vectors is 0° – or 360°). Therefore, it is often preferable to devise non-parametric statistical tests based on permutation or bootstrapping methods (Vinck et al., 2010a,b). For example, phase-locking values measured across trials grouped by behavioral outcome cannot be directly compared with a null hypothesis of zero phase locking, because the null hypothesis actually depends on the exact number of trials in each

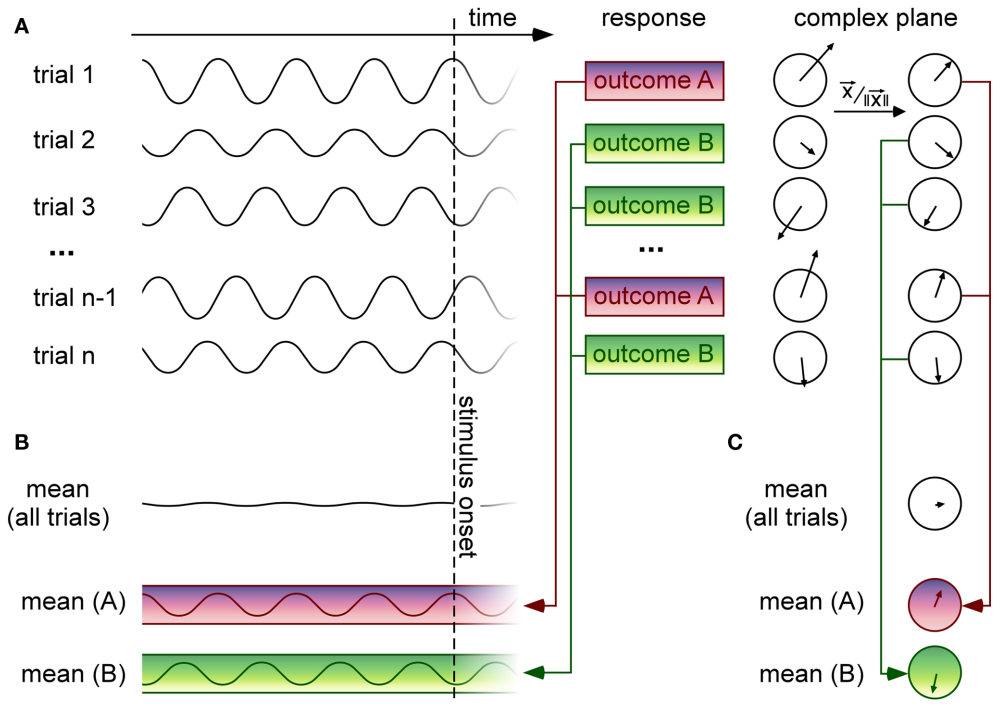


FIGURE 1 | Methods for linking pre-stimulus EEG phase to trial-by-trial variations of a discrete behavioral response. (A) On each trial, a time-varying signal (such as an EEG waveform) is recorded before stimulus onset. Here the signal is represented as oscillatory, but it could also correspond to a band-pass filtered version of a noisy, non-oscillatory signal. For each trial, the observer provides a behavioral response with discrete outcomes [here two possible outcomes **(A,B)**] corresponding, e.g., to distinct percepts (seen vs. unseen, category **(A vs. B)**], or to quantized reaction times (e.g., slow vs. fast), etc. On the right, the time-domain signal recorded on each trial is expressed as a vector in the complex plane (which can be calculated using Fourier or wavelet time–frequency decomposition methods). The direction of the vector and its length represent the phase and the amplitude (respectively) of the oscillatory signal recorded in a given time window of interest. **(B)** Since the moment of stimulus onset is unpredictable, the pre-stimulus mean over all trials is near-zero. However, if certain oscillatory phases systematically induce one or the other behavioral outcomes, then selectively averaging the signals for trials grouped according to the behavioral response should reveal oscillations with distinct phase angles. Here the magnitude of the resulting oscillation appears constant at all pre-stimulus times but in practice, due to external and measurement noise, the phase difference is more likely to be visible just before stimulus onset. This method essentially corresponds to an ERP computed before, rather than after the stimulus onset. **(C)** Another method consists in selectively averaging the vectors for each trial group in the complex domain; prior to averaging, each vector is normalized to a unit length, implying that its phase will always equally contribute to the average, regardless of its amplitude [without this normalizing step, the method would actually give equivalent results to the one described in **(B)**]. This step is important because amplitude modulations that would occur independently of phase effects would tend to obscure them. The length of the resulting vector after averaging is called "phase-locking value" (PLV) or "inter-trial coherence" (ITC; Lachaux et al., 1999). The phase locking is weak when the distribution of phase angles across trials is uniform (as should occur when all trials are pooled together regardless of behavioral outcome); if certain phase angles systematically induce one specific behavioral outcome, on the other hand, this phase locking should be significantly increased by considering only the trials with this outcome.

group (even for a uniform phase distribution, the expected phase locking in any *finite* group of trials is significantly above zero; see **Figure 1C**). A solution is to randomly reassign each trial to one of the behavior-defined groups, keeping the respective number of trials constant, and then re-calculate phase locking for this surrogate dataset; repeating this operation several times provides a distribution of phase-locking values under the null hypothesis, with which the real phase-locking value can be compared to estimate its statistical significance. The same approach can be applied to a situation with a continuous behavioral variable, by shuffling the assignment of behavioral values (e.g., RTs) to the corresponding EEG signals, each time re-calculating the circular test statistic under the null hypothesis (e.g., circular-to-linear correlation). In our experiments (**Figure 2**), we have favored such permutation methods because of their robustness and relative lack of assumptions about the data structure.

ONGOING EEG PHASE PREDICTS PERCEPTUAL VARIABILITY

Recent studies by our group and others have started exploring the impact of the phase of ongoing pre-stimulus EEG oscillations on the subsequent perception of a visual stimulus. In a first study (Busch et al., 2009), we presented brief (6 ms) and dim peripheral flashes of light to our observers ($n = 12$), with the luminance of the flash adjusted individually so that the exact same stimulus would be perceived on approximately half of the trials, but go completely unnoticed on the other half. We computed pre-stimulus phase-locking separately for the two trial groups corresponding to perceived and unperceived flashes, and found for each group a significant increase (compared to phase-locking computed on the same number of trials but drawn randomly, irrespective of perceptual outcome). This increase occurred just before stimulus onset, at a frequency of ~7 Hz (**Figure 2A**), and the effect was maximal over fronto-central electrodes. In fact, by considering the phase of the 7-Hz band-pass

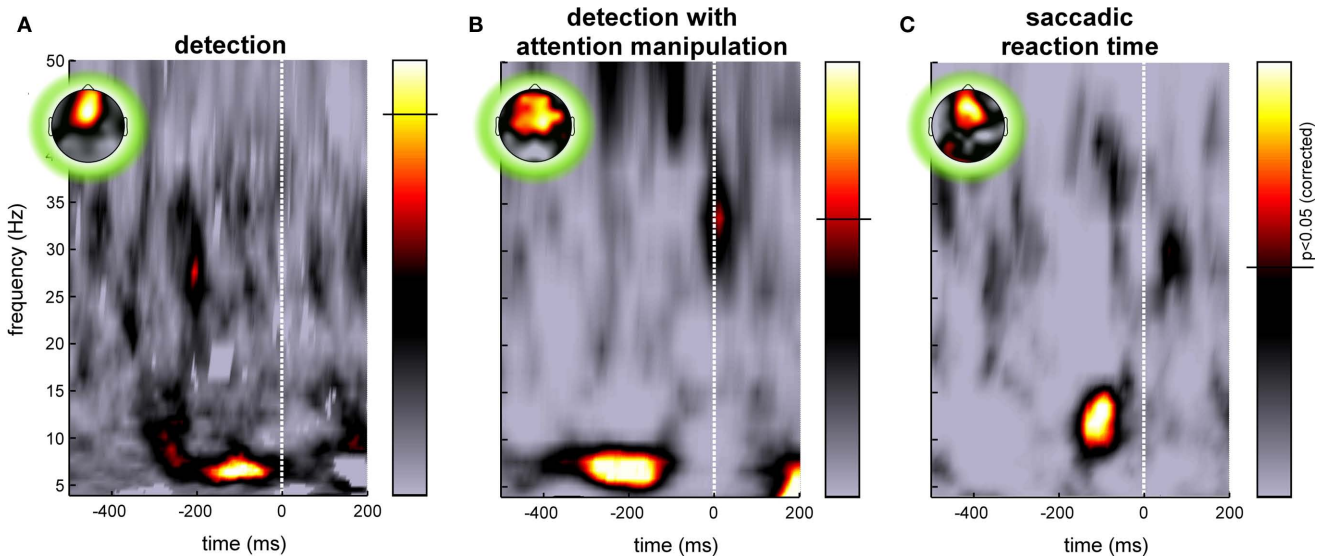


FIGURE 2 | Examples of pre-stimulus EEG phase influences on behavioral response variability. Each image illustrates the significance of a relation between the phase of ongoing EEG oscillations and a subsequent behavioral response recorded on the same trials, for various frequency bands (*y*-axis) and at different pre-stimulus times (*x*-axis). Time zero marks the (unpredictable) onset of the stimulus. The colorbar represents *p*-values, with the significance threshold marked by a horizontal line ($p < 0.05$, corrected for multiple comparisons across time and frequency points using the FDR procedure). The insets illustrate the topography of the effect at the optimal time–frequency point. **(A)** The response specified whether or not the observer had perceived a peripheral flash of light. The influence of phase was measured by comparing the phase-locking computed for two groups of trials corresponding to perceived and unperceived stimuli to surrogate phase-locking values obtained under the null hypothesis (random permutation of behavioral responses). At ~7 Hz and 100–200 ms before the stimulus appears, the phase of frontal EEG on each trial was strongly predictive of the perceptual outcome (Busch et al., 2009). **(B)** In a separate experiment, observers again reported their perception of a flash of light, but the focus of spatial attention was manipulated with a cueing procedure. A significant relation between ongoing EEG phase and trial-by-trial perception was recorded only when the target was flashed at the attended location. Here the image illustrates the significance of a circular-to-linear correlation between pre-stimulus phase (the circular variable) and post-stimulus global field power (GFP, a linear variable which we used as a marker of subjective perception; indeed, this GFP was virtually zero when the target was undetected). As in the previous case, the EEG phase at ~7 Hz, recorded 100–400 ms prior to stimulus onset on frontal electrodes, was maximally predictive of target perception (Busch and VanRullen, 2010). **(C)** In another study, pre-stimulus phase was linked to the subsequent saccadic reaction time (here for a choice discrimination task between two shapes presented left and right of fixation). Reaction times for each subject were binned in five quintiles, and phase locking for each quintile was compared statistically with surrogate phase-locking values obtained under the null hypothesis (i.e., using the same number of trials but randomly drawn, regardless of reaction time). A strongly significant phase-locking increase was again observed on frontal electrodes around 100 ms pre-stimulus, but this time at a frequency of ~13 Hz. Note that in this experiment, time zero corresponds to the beginning of the first stimulus-locked event, a "gap" or disappearance of the fixation point, the choice display itself being presented only after 200 ms (Drewes and VanRullen, 2011).

filtered EEG recorded at those electrodes just before stimulus onset in each trial, we could predict the subsequent response of the subject well above chance. Up to 16% of the trial-by-trial differences in perception were accounted for by comparing trials having the optimal phase angle with those at the opposite angle (Busch et al., 2009).

These findings were globally consistent with the conclusions of an independent study by Mathewson et al. (2009), who also reported that the phase of low-frequency oscillations (around 10 Hz) just before stimulus onset predicted trial-by-trial perception, in a situation where only half of the targets were consciously detected. However, important differences also exist between the two studies. The paradigm differed from ours, first, in that conscious visibility was regulated not by using dim stimuli, but by displaying a strong-contrast stimulus (a "mask") shortly after the target. In addition, the inter-trial interval duration was fixed in that experiment; the possibility that certain oscillatory rhythms could have been reset by the stimulus onset in the previous trial thus makes it difficult to draw strong conclusions in terms of ongoing or spontaneous oscillations. Finally, the analysis method relied on time-domain averaging (as illustrated in **Figure 1A**) which renders phase effects dependent on potential oscillatory amplitude differences between the perceptual conditions; such pre-stimulus amplitude differences between correctly and incorrectly perceived trials are known to exist, particularly over occipital regions at alpha-band frequencies around 10 Hz (Ergenoglu et al., 2004; Hanslmayr et al., 2005, 2007; Thut et al., 2006; van Dijk et al., 2008). This may also contribute to explain why the principal phase effect was observed by Mathewson et al. (2009) at 10 Hz on occipital electrodes, instead of 7 Hz on frontal electrodes in our study. Nonetheless, the fact that both studies point to a similar conclusion reinforces the general idea that pre-stimulus oscillatory phase at 7 and/or 10 Hz can determine to some extent the trial-by-trial changes in our conscious perception of a repetitive event.

ONGOING EEG PHASE REFLECTS PERIODIC ATTENTIONAL SAMPLING

In our next study, we asked whether the influence of ongoing phase on perception was mediated by top-down attentional factors (Busch and VanRullen, 2010). The previous results had been obtained

under conditions in which target location was always known in advance, and therefore subjects may have paid covert attention to that location in order to improve their detection performance. Would the same ongoing phase influence still occur for a target appearing at an unattended location?

Before each trial began, a central cue indicated to the observers ($n = 13$) the location on the screen where they should expect the target (left or right). When the flash of light did occur on this side, everything happened in fact exactly as in the previous experiment – and indeed we confirmed our previous results in this condition, with a strong impact of ~7 Hz pre-stimulus EEG phase recorded at fronto-central electrodes on the probability of target perception (**Figure 2B**). However, the target also sometimes occurred on the opposite side of the screen. In this case the subjects had more difficulty in perceiving the light (as indicated by the higher light intensity that proved necessary to achieve a 50% detection rate in this condition); this confirms that the observers were focusing on the cued side at the detriment of the rest of the screen. Critically, for those trials where the target appeared outside the focus of attention, no significant phase-locking effect was recorded for the perceived or unperceived trials. In other words, perception was related to ongoing EEG phase only via the action of attention. We thus hypothesized that attention samples visual information periodically, and that each ~7 Hz ongoing EEG cycle is the reflection of a new attentional sample (Busch and VanRullen, 2010). Stimuli occurring at around the optimal phase enjoy all the benefits of attention, while others are processed merely as if they were out of the attention focus. For some reason, likely related to its architecture and its neuronal substrates, the attention system could not apply the optimal strategy (optimal for such a detection task with unpredictable target onset) of steadily monitoring the expected location. These findings concur with conclusions from a previous psychophysical study in which we reported that attention samples information periodically at ~7 Hz, even when only a single item needs to be attended (VanRullen et al., 2007). In this context, the topographic localization of the phase effects over fronto-central electrodes may reveal the contribution of the frontal eye field (FEF), an area known, among other things, for its involvement in visual attention (Crowne, 1983; Kodaka et al., 1997; Corbetta and Shulman, 1998; Schall, 2004; Wardak et al., 2006). However, our EEG data would require independent corroboration using more accurate anatomical localization methods, before the implication of FEF can be definitely established.

ONGOING EEG PHASE PREDICTS REACTION TIME VARIABILITY

Attention and saccade programming are heavily intertwined brain functions (Rizzolatti et al., 1987; Deubel and Schneider, 1996; Smith et al., 2004). In particular, the FEF is involved in both visual attention orienting and saccadic motor outputs (Moore and Fallah, 2001; Murthy et al., 2001; Juan et al., 2004; Schall, 2004; Wardak et al., 2006). Our next experiment thus tested whether saccadic responses would also be affected by ongoing pre-stimulus phase (Drewes and VanRullen, 2011).

We used three different versions of a saccadic response task, performed by the same 13 observers. All three tasks required speeded choice responses using leftward or rightward eye movements, but the difficulty of the choice varied. In the easiest case, subjects simply made alternating left and right saccades toward a target whose position was fully predictable, trial after trial. In the second task the position was unpredictable, but since only one target appeared on each trial the task could be performed using mostly reflexive responses. Finally, the third task required a discrimination between two shapes displayed simultaneously; the saccade was made toward the shape that presented an opening at the top. In all three tasks, the fixation point disappeared 200 ms before the appearance of the saccade target display; this so-called "gap" procedure is normally used to maximize the occurrence of rapid "express" saccades (Fischer and Boch, 1983; Fischer and Ramsperger, 1984). The disappearance of the fixation point is itself a transient event that can modify or even reset ongoing oscillations; therefore, we considered the beginning of the gap as time zero, and concentrated only on spontaneous oscillations occurring before this time.

Behavioral results revealed that, as expected, the mean discrimination performance decreased while the average RT increased with task difficulty. This time, we focused on RT variability across trials, and how this variability would relate to ongoing EEG phase differences. In order to apply the analysis methods described previously (see **Figure 1C**), we binned the RTs of each observer into five groups, corresponding to the five quintiles of the RT distribution (this was done separately for each of the three tasks). For each trial group, we calculated the pre-stimulus phase locking (at each frequency and each pre-stimulus time point) and compared it to surrogate phase-locking values calculated using the same number of trials which were randomly drawn, irrespective of RT. Again, a significant increase in pre-stimulus phase locking was found over frontal electrodes for all five quintiles; this time, however, the effect was maximal at a frequency around 13 Hz. This phase effect was stronger in the easy and in the medium difficulty tasks; in fact, during the difficult task, the effect was only observed for subjects who provided very rapid (but often inaccurate) responses (**Figure 2C**). Thus the influence of pre-stimulus phase on RT variability seems to depend on cognitive load and subject strategy; for tasks involving a considerable cognitive effort, many other factors (discrimination and decision processes, motivation) may come into play and contribute their own variability, which would act to conceal the effects of pre-stimulus phase.

The eventuality of a trial-by-trial relationship between pre-stimulus oscillatory phase and subsequent RTs had already been investigated in the past, but these early studies provided only mixed results (Walsh, 1952; Lansing, 1957; Callaway and Yeager, 1960; Dustman and Beck, 1965). One specificity of our experiment that could explain its comparative success, is that our analysis did not assume a one-to-one relationship between phases and RTs – contrary to previous studies that all searched for specific phase angles systematically inducing the fastest, or the slowest RTs. As we found out, the relation between ongoing EEG rhythms and subsequent RTs can actually span more than a single oscillatory cycle. Each range of RT values will be associated with a specific and unique phase angle (as our analysis revealed), but the reverse is not true, that is, a given range of phases might be linked to two, or even more distinct RT values. This sort of relation would be missed by an analysis that would first group the trials by phase, and then compare the RTs in each group (a strategy employed in most of the studies cited above).

It would also be missed by directly calculating the circular-to-linear correlation between phase angles and RTs. A proper analysis for such a situation is the one described in **Figure 1**, in which trials are first grouped according to RTs and phase values are then considered in each group – in other words, the very analysis that we used (**Figure 2C**). The existence of a phase–RT relationship spanning multiple oscillatory cycles suggests that the underlying ~13 Hz ongoing oscillation creates multiple successive and regularly spaced "windows of opportunity" for saccade production.

WHAT ELSE DOES ONGOING EEG PHASE PREDICT?

The evidence reviewed so far unambiguously indicates that ongoing ~7 and ~13 Hz EEG oscillations dynamically modulate information processing in the visual system, and in particular the sampling of visual information by attention. Therefore, the single-trial phase recorded just before stimulus onset can be used as a predictor for subsequent behavioral and perceptual variables. Aside from conscious visual detection and RTs, ongoing EEG phase could also contribute to the trial-by-trial variability of several other cognitive functions. Some of the possible associations are listed below, together with a preview of the significance that their discovery could have for our understanding of brain function. Needless to say, our group is currently exploring several of these issues.

- *Does ongoing EEG phase predict transcranial magnetic stimulation (TMS) phosphene perception at threshold?* The perception of a flash of light is the result of a complex sequence of neuronal processing events, from the retina to the cortex via the thalamus. It is unclear (and heavily debated) which cortical region, or which network of cortical areas is critical for conscious perception to occur. Within this context, our results of a rhythmic influence of ongoing oscillations onto the conscious detection of a flash are difficult to interpret. However, the conscious experience of light can also be induced by direct stimulation of the occipital cortex (the seat of the visual system), for example using TMS pulses. Just as in our experiments, the intensity of the TMS pulse can be individually adjusted so that the perception of the induced "phosphene" only takes place in half of the trials, and it is possible to record EEG while applying TMS pulses (Thut et al., 2005; Taylor et al., 2008; Thut and Miniussi, 2009). Furthermore, focusing on pre-stimulation oscillations means that the study would be immune to the numerous artifacts generally evoked by the pulse in concurrent TMS/EEG studies (Ilmoniemi and Kicic, 2010; Thut and Pascual-Leone, 2010). Would the perceptual outcome in this experiment also depend on ongoing 7 Hz frontal EEG phase? If these oscillations are the reflection of attentional sampling, and assuming that visual attention enhances phosphene perception (Bestmann et al., 2007), the answer is likely to be positive. In addition, would the perceptual outcome on each trial also be affected by the phase of locally generated oscillations within occipital cortex itself – and if so, at what frequency? The fact that TMS-induced perception bypasses many of the early visual processing stages should maximize the chances of directly observing the local interplay between ongoing activity and visual responses, which so far has eluded our previous experimental efforts.
- *Does ongoing EEG phase predict threshold perception in other sensory modalities (e.g., audition)?* It is easy to adjust auditory, or even somato-sensory stimuli so that they are consciously registered only half of the time. Would a pre-stimulus phase difference be observed between perceived and unperceived stimuli? If yes, are the same or different mechanisms involved as in the visual modality, in terms of oscillation frequency, and cortical origin? The results of this experiment may indicate whether the source of the periodic modulation is central, or more sensory-specific. Existing evidence in animals already hints at a supramodal coordination of theta-band oscillatory activities (Lakatos et al., 2009).
- *Does ongoing EEG phase predict the latency of attentional shifts?* Just like saccadic RTs are variable when you move your eyes, it also takes variable amounts of time to shift your attention covertly from one location to another. This variability can be measured using specific paradigms, e.g., using a running analog "clock" at the target location and asking the observer to report the first "time" they can read on the clock after an attention orienting event. This time is generally delayed with respect to the onset of the orienting cue, and the delay is taken to reflect the latency of shifting attention to the target location (Carlson et al., 2006). Even for identical repetitions of an experimental trial, the latency is not fixed but varies by tens of milliseconds. Can the method employed to study the latency of saccadic responses (often taken to represent shifts of "overt" attention) also be used to reveal a relation between ongoing EEG phase and covert attentional shifts? In this case, would the relevant oscillation frequency be found at ~7 Hz (like in our study of attentional sampling; Busch and VanRullen, 2010) or at ~13 Hz (like in our study of saccadic latency; Drewes and VanRullen, 2011)? In any case, a positive outcome would lend credence to our proposal that ongoing ~7 and/or ~13 Hz oscillations mirror the rhythm at which attention samples visual information (VanRullen et al., 2005b, 2006, 2007; Busch and VanRullen, 2010).
- *Does ongoing EEG phase predict the capacity and/or ordering of items in visual short-term memory?* Visual attention and visual working memory share several traits, such as their limited capacity (Luck and Vogel, 1997; Cowan, 2001; Alvarez and Cavanagh, 2004), and it has even been proposed that the two cognitive functions overlap in part (LaBar et al., 1999; Awh et al., 2000, 2006; Downing, 2000; Awh and Jonides, 2001; de Fockert et al., 2001), although this conclusion is debated (Woodman et al., 2001, 2007). An influential model of working memory organization posits that remembered items are maintained in memory as an ordered sequence (Sternberg, 1966), each element being represented by one cycle of a high-frequency oscillation (e.g., 30–80 Hz gamma activity) nested within a lower-frequency cycle (e.g., 4–8 Hz theta activity) supposed to encompass the entire list (Lisman and Idiart, 1995). The limited capacity of working memory (about four to seven items) is explained in this model by the number of high-frequency cycles that can be slotted in one period of the low-frequency oscillation. A recent study of neural responses in monkey prefrontal cortex reported that spikes fired at distinct phases of an ongoing ~32 Hz oscillatory rhythm carried

information about distinct objects from a to-be-remembered list (Siegel et al., 2009). Although the study was presented as evidence for Lisman's model, its findings depart significantly from Lisman's predictions (indeed, the phase dependence should occur at the lower rhythm frequency, not at the higher one). Nonetheless, the findings clearly suggest that pre-stimulus oscillations could relate to short-term memory performance. Remaining questions include (i) whether the phase of ongoing lower-frequency oscillations (e.g., in the theta band) at which a test item is presented would predict the trial-by-trial variability in response time for deciding whether or not the item belongs to the remembered list – this prediction follows naturally from Lisman's model since the phase encodes the item's rank in the list and the rank determines the time needed for retrieval; (ii) whether inter-individual or inter-trial differences in peak oscillation frequency would correlate with capacity and performance measures; (iii) whether interfering with ongoing oscillations at specific phases (for example, using TMS applied on frontal areas) would disrupt working memory maintenance only for specific items within the list.

- *Does ongoing EEG phase predict long-term memory encoding and/or recall?* Memory encoding over longer time scales (minutes, days, or even more) depends both on frontal structures and on the medial temporal lobe system, including the hippocampus (Poldrack and Gabrieli, 1997; Desgranges et al., 1998; Kramer et al., 2005; Ramus et al., 2007). This latter area displays very large amplitude oscillations in the theta band (4–8 Hz) which are known to underlie spatial memory formation in rodents (O'Keefe, 1993; Buzsaki, 2002, 2006). More precisely, the firing of certain hippocampal neurons signals a remembered location in the rat's environment, and the specific phase of the ongoing theta rhythm at which this firing occurs reflects the relative position of the rat with respect to this location – a mechanism coined "theta phase precession" (O'Keefe and Recce, 1993; Skaggs et al., 1996). In fact, theta phase precession also coordinates the firing of prefrontal neurons to the hippocampal theta rhythm (Jones and Wilson, 2005; Siapas et al., 2005). A recent study using single-neuron recordings in humans revealed that trials in which long-term memory formation was successful were characterized by stronger phase locking of hippocampal neurons to the ongoing theta rhythm, even before the onset of the stimulus to be recalled (Rutishauser et al., 2010). This neuronal result does not directly imply the existence of a relationship between the phase of ongoing theta oscillations at the moment of presentation of a visual stimulus and the subsequent recall of this stimulus, but it makes such a relationship worth testing in future experiments.

DO ONGOING OSCILLATIONS PRODUCE PERCEPTUAL SNAPSHOTS?

One critical test of the relation between ongoing rhythms and conscious perception has eluded the community for so long that it deserves a dedicated section in this review. In 1981, inspired by earlier theories claiming that our perceptual experience was built upon discrete processing events, similar to the discrete frames of a movie sequence (Pitts and McCulloch, 1947; Stroud, 1956; Harter, 1967; Allport, 1968), Varela et al. (1981) reported apparently direct evidence for such perceptual "frames" or "snapshots." By presenting identical stimuli (two successive flashes separated by a short delay) at different phases of the ongoing alpha (10 Hz) EEG, they induced drastic changes in the observers' experience of temporal simultaneity: while at one alpha phase they judged the two flashes to have occurred at the same time, at the opposite phase they perceived the *same* two flashes as clearly separate events. The implication is that each alpha cycle slices the continuous temporal sequence of visual inputs into a new discrete chunk or snapshot – when the two flashes straddle the critical phase of the cycle, they are sliced into separate snapshots. Unfortunately, these promising first results did not prove as clear-cut in a follow-up study by the same author (Gho and Varela, 1988), and could simply not be replicated afterward, despite repeated efforts by our group and others (D. Eagleman, personal communication).

This failure is fateful: whereas the evidence reviewed so far of a relation between ongoing oscillatory phase and trial-by-trial variations in conscious detection, attention, or motor outputs implies the existence of periodic components in perception, it only indirectly alludes to the issue of discrete vs. continuous perception. A relation between ongoing phase and temporal framing, on the other hand, would directly, and unambiguously demonstrate the discrete nature of perception (VanRullen and Koch, 2003). Until such a demonstration is provided, the ongoing debate must continue to rely on indirect experimental signatures of the postulated discreteness, such as the motion reversals occurring in continuous light during the "wagon wheel illusion" (Purves et al., 1996; Kline et al., 2004; Andrews and Purves, 2005; Andrews et al., 2005; Holcombe et al., 2005; VanRullen et al., 2005b, 2006; Kline and Eagleman, 2008). In the end, even if it turns out that discrete temporal framing does not occur after all, or that it is restricted to specific sensory domains or experimental situations, the studies reviewed in this article should hopefully convince the reader that the outcome of many important brain functions depends in a periodic manner on the ongoing state of the brain, as reflected by the phase of certain pre-stimulus oscillations; and further, that it is possible to reveal this dependence using careful analysis of single-trial EEG activity.

REFERENCES

Allport, D. A. (1968). Phenomenal simultaneity and the perceptual moment hypothesis. *Br. J. Psychol.* 59, 395–406.

Alvarez, G. A., and Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychol. Sci.* 15, 106–111.

Andrews, T., and Purves, D. (2005). The wagon-wheel illusion in continuous light. *Trends Cogn. Sci. (Regul. Ed.)* 9, 261–263.

Andrews, T., Purves, D., Simpson, W. A., and VanRullen, R. (2005). The wheel keeps turning: reply to Holcombe et al. *Trends Cogn. Sci. (Regul. Ed.)* 9, 561.

Awh, E., Anllo-Vento, L., and Hillyard, S. A. (2000). The role of spatial selective attention in working memory for locations: evidence from event-related potentials. *J. Cogn. Neurosci.* 12, 840–847.

Awh, E., and Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends Cogn. Sci. (Regul. Ed.)* 5, 119–126.

Awh, E., Vogel, E. K., and Oh, S. H. (2006). Interactions between attention and working memory. *Neuroscience* 139, 201–208.

Berens, P. (2009). CircStat: a MATLAB Toolbox for circular statistics. *J. Stat. Softw.* 31, 1–21.

Bestmann, S., Ruff, C. C., Blakemore, C., Driver, J., and Thilo, K. V. (2007). Spatial attention changes excitability of human visual cortex to direct stimulation. *Curr. Biol.* 17, 134–139.

Brandt, M. E. (1997). Visual and auditory evoked phase resetting of the alpha EEG. *Int. J. Psychophysiol.* 26, 285–298.

Busch, N. A., Dubois, J., and VanRullen, R. (2009). The phase of ongoing EEG oscillations predicts visual perception. *J. Neurosci.* 29, 7869–7876.

Busch, N. A., and VanRullen, R. (2010). Spontaneous EEG oscillations reveal periodic sampling of visual attention. *Proc. Natl. Acad. Sci. U.S.A.* 107, 16048–16053.

Buzsaki, G. (2002). Theta oscillations in the hippocampus. *Neuron* 33, 325–340.

Buzsaki, G. (2006). *Rhythms of the Brain.* New York: Oxford University Press.

Callaway, E. I., and Yeager, C. L. (1960). Relationship between reaction time and electroencephalographic alpha phase. *Science* 132, 1765–1766.

Calvin, W. H., and Stevens, C. F. (1967). Synaptic noise as a source of variability in the interval between action potentials. *Science* 155, 842–844.

Carlson, T. A., Hogendoorn, H., and Verstraten, F. A. (2006). The speed of visual attention: what time is it? *J. Vis.* 6, 1406–1411.

Corbetta, M., and Shulman, G. L. (1998). Human cortical mechanisms of visual attention during orienting and search. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 353, 1353–1362.

Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav. Brain Sci.* 24, 87–114.

Crowne, D. P. (1983). The frontal eye field and attention. *Psychol. Bull.* 93, 232–260.

de Fockert, J. W., Rees, G., Frith, C. D., and Lavie, N. (2001). The role of working memory in visual selective attention. *Science* 291, 1803–1806.

Desgranges, B., Baron, J. C., and Eustache, F. (1998). The functional neuroanatomy of episodic memory: the role of the frontal lobes, the hippocampal formation, and other areas. *Neuroimage* 8, 198–213.

Deubel, H., and Schneider, W. X. (1996). Saccade target selection and object recognition: evidence for a common attentional mechanism. *Vision Res.* 36, 1827–1837.

Downing, P. E. (2000). Interactions between visual working memory and selective attention. *Psychol. Sci.* 11, 467–473.

Drewes, J., and VanRullen, R. (2011). This is the rhythm of your eyes: the phase of ongoing EEG oscillations modulates saccadic reaction time. *J. Neurosci.* 31, 4698–4708.

Dustman, R. E., and Beck, E. C. (1965). Phase of alpha brain waves, reaction time and visually evoked potentials. *Electroencephalogr. Clin. Neurophysiol.* 18, 433–440.

Engel, A. K., and Singer, W. (2001). Temporal binding and the neural correlates of sensory awareness. *Trends Cogn. Sci. (Regul. Ed.)* 5, 16–25.

Ergenoglu, T., Demiralp, T., Bayraktaroglu, Z., Ergen, M., Beydagi, H., and Uresin, Y. (2004). Alpha rhythm of the EEG modulates visual detection performance in humans. *Brain Res. Cogn. Brain Res.* 20, 376–383.

Fischer, B., and Boch, R. (1983). Saccadic eye movements after extremely short reaction times in the monkey. *Brain Res.* 260, 21–26.

Fischer, B., and Ramsperger, E. (1984). Human express saccades: extremely short reaction times of goal directed eye movements. *Exp. Brain Res.* 57, 191–195.

Fries, P., Neuenschwander, S., Engel, A. K., Goebel, R., and Singer, W. (2001). Rapid feature selective neuronal synchronization through correlated latency shifting. *Nat. Neurosci.* 4, 194–200.

Fries, P., Nikolic, D., and Singer, W. (2007). The gamma cycle. *Trends Neurosci.* 30, 309–316.

Gho, M., and Varela, F. J. (1988). A quantitative assessment of the dependency of the visual temporal frame upon the cortical rhythm. *J. Physiol.* 83, 95–101.

Hanslmayr, S., Aslan, A., Staudigl, T., Klimesch, W., Herrmann, C. S., and Bauml, K. H. (2007). Prestimulus oscillations predict visual perception performance between and within subjects. *Neuroimage* 37, 1465–1473.

Hanslmayr, S., Klimesch, W., Sauseng, P., Gruber, W., Doppelmayr, M., Freunberger, R., and Pecherstorfer, T. (2005). Visual discrimination performance is related to decreased alpha amplitude but increased phase locking. *Neurosci. Lett.* 375, 64–68.

Harter, M. R. (1967). Excitability cycles and cortical scanning: a review of two hypotheses of central intermittency in perception. *Psychol. Bull.* 68, 47–58.

Holcombe, A. O., Clifford, C. W., Eagleman, D. M., and Pakarian, P. (2005). Illusory motion reversal in tune with motion detectors. *Trends Cogn. Sci. (Regul. Ed.)* 9, 559–560.

Ilmoniemi, R. J., and Kicic, D. (2010). Methodology for combined TMS and EEG. *Brain Topogr.* 22, 233–248.

Jansen, B. H., and Brandt, M. E. (1991). The effect of the phase of prestimulus alpha activity on the averaged visual evoked response. *Electroencephalogr. Clin. Neurophysiol.* 80, 241–250.

Jones, M. W., and Wilson, M. A. (2005). Phase precession of medial prefrontal cortical activity relative to the hippocampal theta rhythm. *Hippocampus* 15, 867–873.

Juan, C. H., Shorter-Jacobi, S. M., and Schall, J. D. (2004). Dissociation of spatial attention and saccade preparation. *Proc. Natl. Acad. Sci. U.S.A.* 101, 15541–15544.

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Res. Brain Res. Rev.* 29, 169–195.

Kline, K., and Eagleman, D. M. (2008). Evidence against the temporal subsampling account of illusory motion reversal. *J. Vis.* 8, 13.1–13.5.

Kline, K., Holcombe, A. O., and Eagleman, D. M. (2004). Illusory motion reversal is caused by rivalry, not by perceptual snapshots of the visual field. *Vision Res.* 44, 2653–2658.

Kodaka, Y., Mikami, A., and Kubota, K. (1997). Neuronal activity in the frontal eye field of the monkey is modulated while attention is focused on to a stimulus in the peripheral visual field, irrespective of eye movement. *Neurosci. Res.* 28, 291–298.

Kramer, J. H., Rosen, H. J., Du, A. T., Schuff, N., Hollnagel, C., Weiner, M. W., Miller, B. L., and Delis, D. C. (2005). Dissociations in hippocampal and frontal contributions to episodic memory performance. *Neuropsychology* 19, 799–805.

Kruglikov, S. Y., and Schiff, S. J. (2003). Interplay of electroencephalogram phase and auditory-evoked neural activity. *J. Neurosci.* 23, 10122–10127.

LaBar, K. S., Gitelman, D. R., Parrish, T. B., and Mesulam, M. (1999). Neuroanatomic overlap of working memory and spatial attention networks: a functional MRI comparison within subjects. *Neuroimage* 10, 695–704.

Lachaux, J. P., Rodriguez, E., Martinerie, J., and Varela, F. J. (1999). Measuring phase synchrony in brain signals. *Hum. Brain Mapp.* 8, 194–208.

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., and Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science* 320, 110–113.

Lakatos, P., O'Connell, M. N., Barczak, A., Mills, A., Javitt, D. C., and Schroeder, C. E. (2009). The leading sense: supramodal control of neurophysiological context by attention. *Neuron* 64, 419–430.

Lansing, R. W. (1957). Relation of brain and tremor rhythms to visual reaction time. *Electroencephalogr. Clin. Neurophysiol.* 9, 497–504.

Large, E. W., and Jones, M. R. (1999). The dynamics of attending: how people track time-varying events. *Psychol. Rev.* 106, 119–159.

Levitan, H., Segundo, J. P., Moore, G. P., and Perkel, D. H. (1968). Statistical analysis of membrane potential fluctuations. Relation with presynaptic spike train. *Biophys. J.* 8, 1256–1274.

Lisman, J. E., and Idiart, M. A. (1995). Storage of 7±2 short-term memories in oscillatory subcycles. *Science* 267, 1512–1515.

Luck, S. J., and Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature* 390, 279–281.

Mathewson, K. E., Fabiani, M., Gratton, G., Beck, D. M., and Lleras, A. (2010). Rescuing stimuli from invisibility: inducing a momentary release from visual masking with pre-target entrainment. *Cognition* 115, 186–191.

Mathewson, K. E., Gratton, G., Fabiani, M., Beck, D. M., and Ro, T. (2009). To see or not to see: prestimulus alpha phase predicts visual awareness. *J. Neurosci.* 29, 2725–2732.

Montemurro, M. A., Rasch, M. J., Murayama, Y., Logothetis, N. K., and Panzeri, S. (2008). Phase-of-firing coding of natural visual stimuli in primary visual cortex. *Curr. Biol.* 18, 375–380.

Moore, T., and Fallah, M. (2001). Control of eye movements and spatial attention. *Proc. Natl. Acad. Sci. U.S.A.* 98, 1273–1276.

Murthy, A., Thompson, K. G., and Schall, J. D. (2001). Dynamic dissociation of visual selection from saccade programming in frontal eye field. *J. Neurophysiol.* 86, 2634–2637.

O'Keefe, J. (1993). Hippocampus, theta, and spatial memory. *Curr. Opin. Neurobiol.* 3, 917–924.

O'Keefe, J., and Recce, M. L. (1993). Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus* 3, 317–330.

Pitts, W., and McCulloch, W. S. (1947). How we know universals: the perception of auditory and visual forms. *Bull. Math. Biophys.* 9, 127–147.

Poldrack, R. A., and Gabrieli, J. D. (1997). Functional anatomy of long-term memory. *J. Clin. Neurophysiol.* 14, 294–310.

Purves, D., Paydarfar, J. A., and Andrews, T. J. (1996). The wagon wheel illusion in movies and reality. *Proc. Natl. Acad. Sci. U.S.A.* 93, 3693–3697.

Ramus, S. J., Davis, J. B., Donahue, R. J., Discenza, C. B., and Waite, A. A. (2007). Interactions between the orbitofrontal cortex and the hippocampal memory

system during the storage of long-term memory. *Ann. N. Y. Acad. Sci.* 1121, 216–231.

Rizzolatti, G., Riggio, L., Dascola, I., and Umilta, C. (1987). Reorienting attention across the horizontal and vertical meridians: evidence in favor of a premotor theory of attention. *Neuropsychologia* 25, 31–40.

Rutishauser, U., Ross, I. B., Mamelak, A. N., and Schuman, E. M. (2010). Human memory strength is predicted by theta-frequency phase-locking of single neurons. *Nature* 464, 903–907.

Schall, J. D. (2004). On the role of frontal eye field in guiding attention and saccades. *Vision Res.* 44, 1453–1467.

Schroeder, C. E., and Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends Neurosci.* 32, 9–18.

Siapas, A. G., Lubenov, E. V., and Wilson, M. A. (2005). Prefrontal phase locking to hippocampal theta oscillations. *Neuron* 46, 141–151.

Siegel, M., Warden, M. R., and Miller, E. K. (2009). Phase-dependent neuronal coding of objects in short-term memory. *Proc. Natl. Acad. Sci. U.S.A.* 106, 21341–21346.

Skaggs, W. E., McNaughton, B. L., Wilson, M. A., and Barnes, C. A. (1996). Theta phase precession in hippocampal neuronal populations and the compression of temporal sequences. *Hippocampus* 6, 149–172.

Smith, D. T., Rorden, C., and Jackson, S. R. (2004). Exogenous orienting of attention depends upon the ability to execute eye movements. *Curr. Biol.* 14, 792–795.

Stern, E. A., Kincaid, A. E., and Wilson, C. J. (1997). Spontaneous subthreshold membrane potential fluctuations and action potential variability of rat corticostriatal and striatal neurons in vivo. *J. Neurophysiol.* 77, 1697–1715.

Sternberg, S. (1966). High-speed scanning in human memory. *Science* 153, 652–654.

Stroud, J. M. (1956). "The fine structure of psychological time," in *Information Theory in Psychology*, ed. H. Quastler (Chicago, IL: Free Press) 174–205.

Tallon-Baudry, C., and Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends Cogn. Sci. (Regul. Ed.)* 3, 151–162.

Taylor, P. C., Walsh, V., and Eimer, M. (2008). Combining TMS and EEG to study cognitive function and cortico-cortico interactions. *Behav. Brain Res.* 191, 141–147.

Thut, G., Ives, J. R., Kampmann, F., Pastor, M. A., and Pascual-Leone, A. (2005). A new device and protocol for combining TMS and online recordings of EEG and evoked potentials. *J. Neurosci. Methods* 141, 207–217.

Thut, G., and Miniussi, C. (2009). New insights into rhythmic brain activity from TMS-EEG studies. *Trends Cogn. Sci. (Regul. Ed.)* 13, 182–189.

Thut, G., Nietzel, A., Brandt, S. A., and Pascual-Leone, A. (2006). Alpha-band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. *J. Neurosci.* 26, 9494–9502.

Thut, G., and Pascual-Leone, A. (2010). Integrating TMS with EEG: how and what for? *Brain Topogr.* 22, 215–218.

van Dijk, H., Schoffelen, J. M., Oostenveld, R., and Jensen, O. (2008). Prestimulus oscillatory activity in the alpha band predicts visual discrimination ability. *J. Neurosci.* 28, 1816–1823.

VanRullen, R., Carlson, T., and Cavanagh, P. (2007). The blinking spotlight of attention. *Proc. Natl. Acad. Sci. U.S.A.* 104, 19204–19209.

VanRullen, R., Guyonneau, R., and Thorpe, S. J. (2005a). Spike times make sense. *Trends Neurosci.* 28, 1–4.

VanRullen, R., Reddy, L., and Koch, C. (2005b). Attention-driven discrete sampling of motion perception. *Proc. Natl. Acad. Sci. U.S.A.* 102, 5291–5296.

VanRullen, R., and Koch, C. (2003). Is perception discrete or continuous? *Trends Cogn. Sci. (Regul. Ed.)* 7, 207–213.

VanRullen, R., Reddy, L., and Koch, C. (2006). The continuous wagon wheel illusion is associated with changes in electroencephalogram power at approximately 13 Hz. *J. Neurosci.* 26, 502–507.

Varela, F., Lachaux, J. P., Rodriguez, E., and Martinerie, J. (2001). The brainweb: phase synchronization and large-scale integration. *Nat. Rev. Neurosci.* 2, 229–239.

Varela, F. J., Toro, A., John, E. R., and Schwartz, E. L. (1981). Perceptual framing and cortical alpha rhythm. *Neuropsychologia* 19, 675–686.

Vinck, M., Lima, B., Womelsdorf, T., Oostenveld, R., Singer, W., Neuenschwander, S., and Fries, P. (2010a). Gamma-phase shifting in awake monkey visual cortex. *J. Neurosci.* 30, 1250–1257.

Vinck, M., van Wingerden, M., Womelsdorf, T., Fries, P., and Pennartz, C. M. (2010b). The pairwise phase consistency: a bias-free measure of rhythmic neuronal synchronization. *Neuroimage* 51, 112–122.

Walsh, E. G. (1952). Visual reaction time and the alpha-rhythm, an investigation of a scanning hypothesis. *J. Physiol.* 118, 500–508.

Wardak, C., Ibos, G., Duhamel, J. R., and Olivier, E. (2006). Contribution of the monkey frontal eye field to covert visual attention. *J. Neurosci.* 26, 4228–4235.

Woodman, G. F., Luck, S. J., and Schall, J. D. (2007). The role of working memory representations in the control of attention. *Cereb. Cortex* 17(Suppl. 1), i118–i124.

Woodman, G. F., Vogel, E. K., and Luck, S. J. (2001). Visual search remains efficient when visual working memory is full. *Psychol. Sci.* 12, 219–224.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 12 January 2011; paper pending published: 26 January 2011; accepted: 24 March 2011; published online: 09 April 2011.

Citation: VanRullen R, Busch NA, Drewes J and Dubois J (2011) Ongoing EEG phase as a trial-by-trial predictor of perceptual and attentional variability. Front. Psychology ***2****:60. doi: 10.3389/fpsyg.2011.00060*

This article was submitted to Frontiers in Perception Science, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

METHODS ARTICLE
published: 09 March 2011
doi: 10.3389/fpsyg.2011.00036

Probing of brain states in real-time: introducing the ConSole environment

Thomas Hartmann[1]*, Hannah Schulz[1] and Nathan Weisz[1,2]

[1] *Department of Psychology, Universität Konstanz, Konstanz, Germany*
[2] *Zukunftskolleg, Universität Konstanz, Konstanz, Germany*

Edited by:
Guillaume A. Rousselet, University of Glasgow, UK

Reviewed by:
Rufin Vanrullen, Centre de Recherche Cerveau et Cognition, France
Cyril R. Pernet, University of Edinburgh, UK

***Correspondence:**
Thomas Hartmann, Department of Psychology, Universität Konstanz, P.O. Box D25, 78457 Konstanz, Germany.
e-mail: thomas.hartmann@uni-konstanz.de

Recent years have seen huge advancements in the methods available and used in neuroscience employing EEG or MEG. However, the standard approach is to average a large number of trials for experimentally defined conditions in order to reduce intertrial-variability, i.e., treating it as a source of "noise." Yet it is now more and more accepted that trial-to-trial fluctuations bear functional significance, reflecting fluctuations of "brain states" that predispose perception and action. Such effects are often revealed in a pre-stimulus period, when comparing response variability to an invariant stimulus. However such offline analyses are disadvantageous as they are correlational by drawing conclusions in a *post hoc*-manner and stimulus presentation is random with respect to the feature of interest. A more direct test is to trigger stimulus presentation when the relevant feature is present. The current paper introduces *Con*stance *S*ystem for *On*line *E*EG (ConSole), a software package capable of analyzing ongoing EEG/MEG in real-time and presenting auditory and visual stimuli via internal routines. Stimulation via external devices (e.g., transcranial magnetic stimulation) or third-party software (e.g., PsyScope X) is possible by sending TTL-triggers. With ConSole it is thus possible to target the stimulation at specific brain states. In contrast to many available applications, ConSole is open-source. Its modular design enhances the power of the software as it can be easily adapted to new challenges and writing new experiments is an easy task. ConSole is already pre-equipped with modules performing standard signal processing steps. The software is also independent from the EEG/MEG system, as long as a driver can be written (currently two EEG systems are supported). Besides a general introduction, we present benchmark data regarding performance and validity of the calculations used, as well as three example applications of ConSole in different settings. ConSole can be downloaded at: http://console-kn.sf.net.

Keywords: EEG, single-trial analysis, real-time analysis, oscillation, brain states, neurofeedback

INTRODUCTION

RATIONALE

An increasing amount of electrophysiological (EEG/MEG) studies have recently shed new light on our understanding of how the brain processes and represents internally and externally generated input. The still-dominant approach of stimulus averaging across several trials implicitly or explicitly assumes an invariance of a neuronal response toward a certain stimulus and treats trial-by-trial fluctuations as noise. This notion is particularly pronounced in ERP studies in which even temporally fluctuating neuronal responses ("induced activity"; Tallon-Baudry and Bertrand, 1999) are removed as a consequence of averaging. However, it is becoming increasingly obvious that trial-by-trial fluctuations bear functionally meaningful information and explain a significant amount of the trial-to-trial variability observed in overt behavior. Evidence for this view originates in a variety of different approaches that encompass different levels of neuronal activity. For example, Romei et al. (2008b) showed that the level of ongoing alpha power in visual areas *interindividually* predicts the intensity needed to elicit phosphenes at chance level. Trial-by-trial fluctuations of pre-stimulus alpha power also predicted whether or not a phosphene would be perceived *intraindividually*, when subjects were stimulated at threshold intensity (Romei et al., 2008a). Other studies have shown better performance in a visual detection task in trials preceded by low power in the alpha band prior to stimulus onset (Ergenoglu et al., 2004; Hanslmayr et al., 2007; van Dijk et al., 2008). This is also the case for visual discrimination tasks (Hanslmayr et al., 2005a) and the performance can even be modulated by neurofeedback (Hanslmayr et al., 2005b). Although they are the most frequently reported phenomenon, pre-stimulus effects are not restricted to the alpha band. Similar results have been found for beta (Schubert et al., 2009) and gamma (Wyart and Tallon-Baudry, 2009). Taken together, these results suggest that the fluctuations of ongoing cortical oscillations represent certain brain states that determine the "fate" of how an incoming stimulus will be further processed. Of interest is that the patterns of these results found on a neurophysiological as well as behavioral level strongly resemble the patterns found in studies in which ongoing oscillatory activity is modulated in a top-down manner by differential experimental conditions. Prominent examples include attention (Klimesch et al., 1998; Worden et al., 2000; Bastiaansen et al., 2001; Romei et al., 2008a) as well as working

memory (Jensen and Jokisch, 2007; Tuladhar et al., 2007). Evidence for the significance of these fluctuations is also provided by studies that show that ongoing brain activity produces highly structured patterns – similar to those evoked by an actual stimulus – on the level of single units (Kenet et al., 2003) as well as on the system level (Fox and Raichle, 2007). In functional terms, these findings support the aforementioned notion that fluctuations of ongoing brain activity represent fluctuations of brain states that are associated with differential predispositions for a certain cognitive performance. Instead of averaging variability away, a growing community of neuroscientists is realizing that understanding trial-to-trial variability may hold one key to a deeper understanding of brain functions. However, this research path requires the challenging analysis of data on a *single-trial level*.

A common approach to analyzing the contributions of trial-by-trial variability in MEG/EEG data involves first transforming the data from the time-domain to the time–frequency-domain. The resulting amplitude and phase values can then be averaged over trials (e.g., "seen" versus "not seen" comparison; Romei et al., 2008a) to increase the signal-to-noise ratio or analyzed on a single-trial level [e.g., correlation of EEG with motor-evoked potential (MEP) following transcranial magnetic stimulation (TMS); Sauseng et al., 2009]. It is crucial to keep in mind that this approach assumes that oscillations are the lingua franca for establishing communication within and between neuronal assemblies. The simplest approach is to analyze the change in power of certain frequency bands while reacting to a stimulus. While the power of an oscillation represents the amount of local synchronization (e.g., a brain region or a fraction thereof), other methods can be employed to explore connectivity between distant brain regions (Friston, 2002). These methods either assess the statistical dependency between two time series of two sensors or sources on a single-trial level (functional connectivity; e.g., coherence or phase synchrony; for an overview see e.g., Varela et al., 2001) or measure the "causal" association between two signals that yield information about the directionality of the information flow (effective connectivity; e.g., Partial Directed Coherence; Baccalá and Sameshima, 2001).

All these studies nonetheless face a major limitation: although they aim to show a direct relationship between certain features of cortical oscillations and an assumed functional state on a trial-by-trial basis, their conclusions are drawn *ex post facto*. Since the direct (real-time) control of cortical oscillations is difficult *in vivo* (an interesting avenue may be the recently reported "entrainment" paradigms; e.g., Mathewson et al., 2010; Romei et al., 2010), even though it is not resolved how "entrained" oscillations are really associated to genuinely spontaneously produced oscillations), the experimental setup that comes closest to allowing causal inferences is to temporally trigger events as close as possible to the hypothetically relevant brain activity feature, which fluctuates over the course of an experiment. A real-time framework even holds advantages for the more conventional offline analysis approach, including a clearer distinction between high and low alpha trials that enables a better contrast between the hypothesized brain state and its assumed behavioral impact (e.g., on reaction time). Currently, the presentation of stimuli is entirely random with respect to features of ongoing brain oscillations and the conclusions that can be drawn from it are correlational (see **Figure 1**).

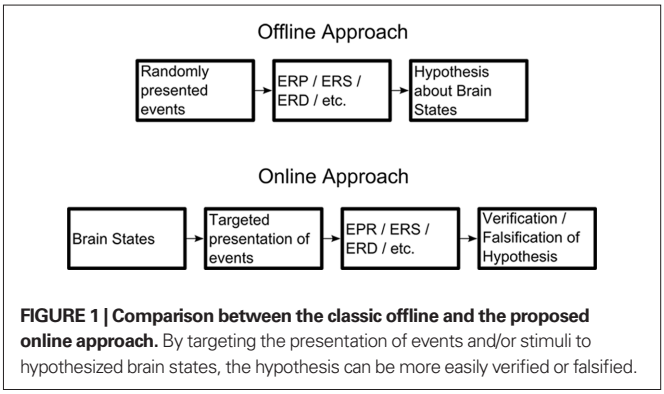


FIGURE 1 | Comparison between the classic offline and the proposed online approach. By targeting the presentation of events and/or stimuli to hypothesized brain states, the hypothesis can be more easily verified or falsified.

REQUIREMENTS OF A REAL-TIME APPLICATION

In order to pursue the hypothesis-driven research approach outlined above, a system is needed that monitors in real-time the relevant feature of ongoing oscillatory activity and controls the course of the experiment dependent on certain criteria (e.g., level of power in a particular frequency band, level of synchrony between distinct sources) *defined in advance* by the experimenter. Such an approach would ideally complement explorative *ex post facto* studies in order to allow for stronger inference regarding the functional relevance of certain brain states. In this paper, we present a user-friendly and open-source software environment called *Con*stance *S*ystem for *On*line *E*EG (ConSole), which allows EEG/MEG researchers to pursue such a hypothesis-driven approach. Although some commercial and non-commercial (not necessarily open-source) programs exist that enable the implementation of some aspects of the described research approach, they are either targeted at clinical neurofeedback or brain–computer-interfaces (BCI). Even though this kind of research is also feasible within the ConSole environment, the main intention is its use as cognitive neuroscientific tool – that is, to test hypotheses regarding the functional relevance of brain oscillations in humans.

A few applications currently exist that make experimental setups like these possible. These, however, display shortcomings that cannot be neglected. Many EEG system vendors ship real-time software along with their products; yet these applications are closely tied to a specific brand or even type of amplifier and most are designed for neurofeedback. Thus, they do not provide the features essential for controlling the course of the experiment (e.g., a TMS or experimental computer). Switching to a different hardware system is virtually impossible – a serious limitation in cases when laboratories with different hardware have decided to collaborate. Moreover, these systems are mostly proprietary, meaning that one has to rely on the features provided by the manufacturer with no possibility to alter or add functionality, let alone distribute these changes so that other scientists can profit from them. Another option would be to use one of the two available systems targeted at BCI development – BCI2000 (Schalk et al., 2004) and OpenVIBE (Renard et al., 2010). While only the first system suffers from the licensing issues described above, both options are primarily designed for BCI and neurofeedback research and not for conducting real-time cognitive neuroscience experiments. The FieldTrip community has developed another promising approach by integrating their extensive Matlab routines with a real-time acquisition system (Oostenveld et al., 2011). As

the whole toolbox is released under an open-source license, it also offers great flexibility. It lacks however the modularity of ConSole (see below) and suffers from lower processing speed since it requires the Matlab environment. The modules of ConSole are compiled binaries written in C++, a programming language that can be greatly optimized for speed with modern compilers and also provides an interface to Matlab, thus offering the best of both worlds.

Another important aspect sometimes neglected in software designs in this area is the distinction between developer, scientist and investigator, who all have different requirements for such an application. The available FieldTrip/EEGLAB toolboxes, for example, require a certain degree of proficiency in Matlab programming. However, the scientist who plans and sets up the experiment often has little or no programming experience. It is thus vital to provide an interface flexible enough to exploit all features of the program but still easy and intuitive enough to achieve this in a short time. Whenever the scientist needs a feature that is not yet available, the developer needs a clear and well documented framework for its rapid development. It is essential that this can be carried out without in-depth knowledge of the whole system as this makes the creation of new features easier, faster, and less error-prone. A modular system consisting of independent components that use a standardized means of communication is the ideal solution to this problem. This approach leaves a consistent interface for the scientist, thus further facilitating the setup of experiments. The investigator's (not necessarily the scientist; frequently a Masters student, student intern, etc.) needs in such a program are different, as he or she is the one using the application in the actual experiments. The constrained availability of laboratory time and the need for a professional setting for the participant place further demands on the software. First and most importantly, the software must be simple, fast, and intuitive to use. All unnecessary aspects of the experiment should be hidden, as this reduces the demand on the investigator and thus minimizes errors. Secondly, the application must be mature enough to run stably and must provide a means of recovering a started session after a crash without intervention from the investigator.

*Con*stance *S*ystem for *On*line *E*EG is an innovative system for real-time experiments in cognitive neuroscience, designed to meet the aforementioned demands of the software architecture. Importantly, it can be downloaded for free[1] and is released under the GPL (Free Software Foundation, 1991), making it possible to share modifications or amendments made to the code. Furthermore, ConSole clearly separates between the different tasks mentioned above by using independent modules that can be combined to build the actual experiment. These modules are written in C++, which we consider to be the best compromise between high-execution speed and the demand for high-level programming. C++ is also a standard programming language and is thereby familiar to most developers. ConSole provides a set of built-in modules for virtually any task related to conducting real-time EEG, including different amplifiers, filters, FFT, source projection, and stimulation (see Methods section for further details). A Matlab interface allows for faster proof-of-principle checks in case a method is not available yet in C++ but can be found in one of the MEG/EEG Matlab toolboxes. Moreover, ConSole is able to control external devices via TTL pulses (e.g., TMS device, PsyScope X; Macwhinney et al., 1997). To set up an experiment, an XML file is written, specifying which modules to use, and how these are connected. The XML-dialect used for this purpose is easy to understand, well documented, and does not require prior knowledge of any programming language in order to prepare even moderately complex experiments. In order to make the actual runs of the experiment as easy as possible, ConSole provides the user with a simple and intuitive graphical user interface (GUI). No knowledge about the internal details of ConSole or the specific experiment is needed to run this.

In the current article, we present details on the techniques used with ConSole, including benchmark results. To further demonstrate the capabilities of ConSole, we provide three examples of possible applications. The first application was inspired by a recent paper by Romei et al. (2008a), which showed that phosphene detection not only depends on the correct site of stimulation and intensity but also on the current brain state represented by occipital alpha oscillations. Using the possibilities opened up by ConSole, we directly tested online whether a phosphene was more or less likely to be perceived when alpha was respectively low or high. This example serves as a proof-of-principle that the idea of online EEG experiments (see above) is feasible in practice.

The second example is based on findings of our workgroup (Dohrmann et al., 2007b) and illustrates how neurofeedback experiments can be implemented in ConSole. Patients suffering from chronic tinnitus (a sound lacking a physical source, usually described as a tone or a hissing) learned to normalize their alpha oscillations and thereby putatively increased inhibitory drive in auditory regions. While our first training approach was conducted using proprietary software supplied by the manufacturer, the training outlined in this article benefited to a great extent from the flexibility of ConSole.

To further emphasize the flexibility of ConSole and to demonstrate the Matlab interface, we also present an application for measuring single-trial MEP elicited by a TMS pulse.

SOFTWARE DESCRIPTION

*Con*stance *S*ystem for *On*line *E*EG was designed with two main goals: ease of use and flexibility. As a monolithic architecture is unable to provide the flexibility needed for such an application, we decided to adopt a modular approach, dividing the application into three distinctive parts: the GUI, the framework, providing internal management functions, and the actual signal processing routines. While the first two parts are included in the main application, the signal processing (as well as drivers for amplifiers, data visualization, etc.) is implemented using totally independent modules – so-called plug-ins. These plug-ins are developed independent of ConSole, which only provides the framework and basic algorithms. This approach eases the development of new modules, as it is only necessary to understand the plug-in interface and not the whole structure of ConSole. The actual experiment is written in an XML-dialect, describing the modules used and the connections between these. This further abstraction from the internal structure of ConSole hides irrelevant details from the designer of the experiment, thus making it possible to design a paradigm without programming skills. For an overview of the structure of an experiment designed using ConSole, see **Figure 2**.

[1]http://console-kn.sf.net

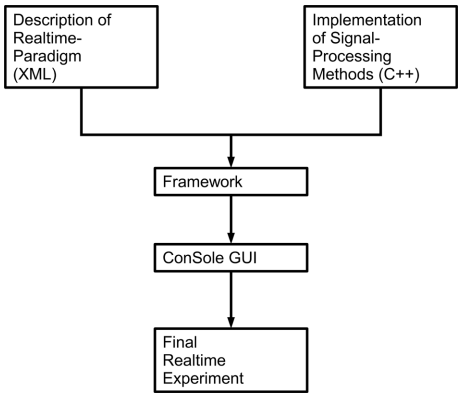


FIGURE 2 | Coarse overview over the structure of ConSole. The general purpose modules implemented in C++ are used in conjunction with the actual real-time paradigm. Both are joined by the framework that ConSole provides. ConSole then displays all necessary information on a computer screen via its GUI.

The description of an experiment for ConSole follows a hierarchical approach: the highest element is the paradigm, which comprises the whole experiment. The definition of the paradigm includes general information about the experiment, such as the patient-specific information that needs to be acquired. A paradigm includes one or more Setups which define the signal processing and thus the logic of the experiment.

ABOUT MODULES

To achieve our goal of writing an easy-to-use and flexible application for real-time EEG experiments, we decided to restrict each module to one specific task independent from other modules. In this regard, modules are the basic entity of ConSole. This approach is beneficial to both module developers and experiment designers. The experiment designer is provided with modules, each specialized in one specific task that mimics steps from offline analysis. These modules and their connections have to be specified in order to design an experiment (see **Figure A1** in Appendix). As all modules function independent of one another and only serve one specific purpose, the module developer only needs to focus on the specific function that the module is supposed to provide. Furthermore, all modules run in separate threads, thus making the application highly scalable on the multi-core PCs that have become a standard in recent years.

The module is used as a simple concept in ConSole. Each module can have one or more input and/or output ports. Incoming data (EEG data or any kind of other information-like events) enter the module via one of the input ports (or through an external source, e.g., modules receiving data from amplifiers). The internal logic processes the data and sends the results to the next module via one of the output ports. It must be emphasized that the individual module is completely agnostic of the module it receives data from or sends data to. Moreover, each module can provide settings such as cutoff frequency and filter order for a module that implements frequency filters. To use a module in an experiment, the designer first declares which module will be used, then adjusts the relevant settings and finally connects the ports.

The following example further clarifies the use of modules in a real-time ConSole experiment. For simplicity, only two modules are considered: the source of the EEG data (i.e., the module acquiring the data from the amplifier and feeding it into the system) and a highpass filter. As the source is mandatory for an experimental paradigm, it must not be declared (unlike, e.g., the filter module). The filter module also takes settings defining the cutoff frequency and the order of the filter. In this example, we use a recursive Butterworth filter with a cutoff frequency of 2 Hz and an order of 3.

```
<module>
        <plugin>IIR BW Highpass</plugin>
        <name>Highpass Filter </name>
        <set cutoff = ”2”/>
        <set order = ”3”/>
</module>
```

Afterward, the source module is connected to the filter module. The source module provides an output port called “output,” while the filter module provides both an input port called “input” and an output port called “output.” In this fashion, data flows from the output port of the source to the input port of the filter. After the data is processed in the filter it is sent to the next module via the filter’s output port.

```
<connect>
        <input module = ”Source” port = ”output”/>
        <output module = ”Highpass Filter” port = ”input”/>
</connect>
```

On top of providing a plug-in framework that enables the implementation of any module, ConSole already ships with well-tested standard modules that use methods commonly employed in offline analysis tools. These include:

- Signal acquisition for a variety of devices
- Various signal processing modules. See **Table 1** for details.
- Acquisition, processing, and output of triggers
- Graphic and sound output
- Raw data displays
- Neurofeedback display

ADDITIONAL FEATURES OF THE CONSOLE FRAMEWORK

Apart from providing modules for the most important signal processing tasks as well as an easy plug-in framework for extending the functionality of ConSole, the application further provides the experiment designer with important features that enable simple as well as complex paradigms.

Simulating a data source

As with non-real-time experiments, developing an experimental design with ConSole includes testing the code. ConSole provides a special input source that reads data from a file. It is thus possible to test an experiment as well as individual modules using real and simulated data. Two data formats are presently supported: Simple Binary Matrix, a format used for example by BESA and BDF, as used by BioSemi. Future releases will feature a variety of supported data formats for simulating as well as for saving acquired data to disk.

Table 1 | Signal processing modules implemented in ConSole including details and references where applicable.

Module	Details	References
Average reference	Re-reference data to average reference. Subtract the mean over all channels at each sample.	
Check peak	Calculates the FFT on the block of data and rejects blocks that do not show a peak in a specified frequency range.	
Combine orientations	Combines the orientations resulting from source projection by either rotating the components using a PCA and choosing the one with the highest eigen value or by calculating the total energy.	
Complex demodulation	Complex demodulation of the incoming signal.	
Distribution	Calculates the percentile of the data based on a distribution acquired in a calibration run.	
FFT	Fast Fourier transform using the fftw library.	Frigo and Johnson (2005)
FIR filter (lowpass and highpass)	Finite impulse response filter calculation using Windowed-Sinc algorithm with Blackman-Window.	Octave-Forge (2010)
Hilbert	Calculates the Hilbert transform.	
ICA artifact correction	Corrects the data for artifacts using filters calculated by PCA or ICA (currently only JADE is implemented)	Cardoso and Souloumiac (2002)
IIR filter (lowpass, highpass, and bandpass)	Infinite impulse response filter calculation using the Butterworth algorithm.	Octave-Forge (2010)
Interpolator	Interpolates the signal of all channels in a block of data that are identified of including artifacts based on variance and maximum amplitude using spline interpolation.	Perrin et al. (1989, 1990)
Matlab	Sends the data to Matlab and runs a script on the data. The result is fed back to ConSole.	
Normalizer	Compute *z*-values of the data based on a calibration run.	
RejectVarMax	Rejects blocks of data that show high variance or amplitude specified in the paradigm.	
Source projection	Dipole source projection using a four-shell concentric sphere model. Adapted from Fieldtrip (Oostenveld et al., 2011).	Cuffin and Cohen (1979), Scherg (1990)
Spatial filter	Applies an externally calculated spatial filter to the data by matrix multiplication.	

Subject and session management

A very important feature of an integrated experimental framework like ConSole is a proper and easy-to-use subject and session management. The purpose of such a system is to provide standardization in terms of where to store data acquired within the experiment (including events and externally generated information such as, for instance, key presses), general subject data (e.g., subject-id, age, and other information related to the experiment), and log files.

Crash management

Most, if not all, real-time experiments feature several runs that must be completed in one experimental session. It is thus vital for an efficient and unobstructed work flow that potential application and operating system crashes or technical issues like power failure cause the least possible impact on the current session. ConSole is therefore equipped with an efficient crash management system. The system automatically analyzes the chosen subject's dataset for incomplete sessions and is able to resume the program at the point of the crash, thus minimizing the loss of data and time.

Matlab interface

To facilitate the rapid implementation of innovative ideas, ConSole is equipped with an interface to Matlab. Module development can therefore be carried out using a widely accepted programming language familiar to many neuroscientists and innovative ideas can be implemented and tested very quickly. This approach, however, also faces downsides. Transferring data to Matlab and back produces a considerable overhead. Moreover, calculations in Matlab tend to be much slower compared to calculations using C++, although this greatly depends on the amount of optimization in the Matlab code. On the other hand, modules written in Matlab have access to a vast amount of functions and toolboxes not available to C++, such as EEGLab, FieldTrip, and NutMEG. Although it is of course possible to port those functions

to C++, this is not always feasible due to time constraints, thus further increasing the value of the Matlab Interface for initial testing purposes.

Artifact correction

Depending on the regions of interest and the setup of a particular experiment, a powerful artifact correction or artifact rejection algorithm is essential to a real-time experiment. The impact of artifact-contaminated data on the outcome of an experiment is much greater for real-time, brain state-driven studies, as an artifact is not only a period of unusable data but might also lead to a false calculation of the current brain state and could therefore, for example, trigger a stimulus that should not have been triggered. The impact of certain artifacts of course depends to a great extent on the region of interest for the calculation of the brain states and the reference used. So the choice of whether to employ artifact correction, rejection of contaminated trials or no correction at all in an experiment should be carefully considered. Artifact correction algorithms alter the data and reduce independence between channels. Rejecting contaminated trials is possible as long as the experiment does not depend on ongoing, uninterrupted output. Identifying contaminated trials is not a trivial task in a real-time experiment. In contrast to offline analysis, visual inspection of the data is not possible. Only unsupervised, automatic algorithms, which commonly employ simple thresholds for the absolute maximum or the variance in the data, can be used. This method is, of course, far from perfect and results in false positives and false negatives. Proper tuning of the relevant parameters is therefore crucial.

Presently, the best method for artifact correction suitable for online approaches is using spatial filters calculated that use one of the many ICA algorithms available. In general, ICA algorithms calculate components of the data that are as independent as possible from one another. This property renders the approach very suitable for online EEG analysis, as (1) common artifacts like eye movements, blinks, and movements of the head or body are completely independent from any cortical signal, and (2) as opposed to PCA approaches, the contamination of artifact components with cortical data is kept to a minimum as the ICA algorithms ensure maximum independence between the components.

*Con*stance *S*ystem for *On*line *E*EG implements ICA-based artifact correction using a modular framework, essentially allowing any spatial filter-based algorithm to be included in the application. To calculate the weights, a calibration run is performed before the actual experiments on which the weights are calculated. ConSole provides the user with a very convenient way of choosing artifact components by offering a three-split window (see **Figure A2** in Appendix). The upper panel shows the original data, the middle panel the calculated components, and the lower panel the resulting data after correction. Artifact components can be marked in the middle panel with an immediate effect on the lower panel. Together with the possibility of displaying the topography of the component on a 3D head, choosing the correct components to reject is a fast and easy process. To use the calculated weights and chosen components in the experiment, the "ICA Artifact Correction" module has to be placed between the "Average Reference" module and the rest of the processing queue.

Two ConSole modules provide rejection of contaminated trials. The "Reject VarMax" module implements thresholding of the incoming data. It first identifies channels exceeding a certain variance or absolute maximum. These channels are then set to zero. The remaining channels are then tested for variance and absolute maximum. If one of the defined thresholds is exceeded, the trial is not forwarded to the next module.

Another approach for separating good and bad trials in an experiment investigating oscillatory activity is checking for peaks in the frequency spectrum of the trial. For instance, in an experiment studying the impact of high/low alpha on some cognitive measure, each trial can be checked for a peak in the alpha band using the "Check Peak" module. Trials containing only noise or some other artifact interfering with the normal spectrum are thus rejected.

MEASURING TIMING ACCURACY

A crucial property of a real-time experiment system is the precise timing of the triggered events. Events should be detected and propagated with as little delay and jitter as possible. Both parameters have many influencing factors. It is thus important to separate these factors in order to optimize the system.

Factors influencing delay and jitter

As pure electrical transmission occurs almost instantly, the first factor to be taken into account is the amplifier, including the transmission to the PC (in most cases today via USB). The delay and jitter introduced until this point cannot be reliably estimated, and for the amplifier used in our setup (ANT, 128 channels), no specifications exist. We thus used a custom-built button box capable of generating TTL pulses to trigger the amplifier and a function generator (WAVETEK 10 MHz DDS model 29). The function generator produced a negative pulse at one of the amplifier's channels. When this negative pulse was detected by ConSole, it sent a second TTL pulse to the amplifier that could be differentiated from the pulse generated by the button box (see **Figure A3** in Appendix). It is thus possible to calculate the delay between the button press and the TTL pulse generated by ConSole. We used 398 pulses to measure the minimum delay and jitter of the system.

These are, of course, ideal conditions. In a normal experiment, delay and jitter are necessarily greater since the higher amount of running analysis modules will add to both. Furthermore, certain signal processing methods such as digital filters inherently add delay and jitter. In addition, many methods require blocks of data to operate on. In those cases, a decision between better and more valid signal processing against faster signal processing has to be made. These parameters largely depend on data quality and the features to be extracted and must be balanced with timing constraints. To measure a more realistic setup, we adapted the design of the second experiment to measure timing accuracy. Incoming data was filtered (IIR Highpass, order 2, cutoff 2 Hz; IIR bandpass, order 3, passband 3–40 Hz) and then grouped into 1-s blocks (512 samples) every 125 ms. These blocks were then prepared for Hilbert transformation by applying an IIR bandpass filter (order 3, passband 8–12 Hz). To avoid filter artifacts at the beginning of the blocks, the first 256 samples (500 ms) of the blocks were discarded before the Hilbert transform. The instantaneous amplitude of the remaining data was then computed by calculating the square root of the sum of squares

of the real and imaginary part of the analytic signal. A preparatory run was used to calibrate the system to only react on high levels of the 10-Hz oscillation. As soon as the calibrated level was exceeded, a TTL pulse was sent to the amplifier.

We used the same function generator as in the first timing test to generate one hundred ten 10-Hz oscillations. The oscillation was triggered by pressing a button on the custom device and stopped after the button was released. The button press also sent a TTL pulse to the amplifier. As soon as ConSole detected the oscillation, it sent a TTL pulse to the amplifier. We used the difference between the two TTL pulses to estimate the delay and jitter of the setup.

As we had expected, the results differed greatly between the simple setup that only included the detection of a rectangular pulse and the more elaborate one in which a 10-Hz oscillation had to be detected. The first setup showed a mean delay of 17.5 ms with a SD of 0.5 ms. The delay was in the range between 0.8 and 27.3 ms and the distribution was approximately uniform (see **Figure 3A**). In the oscillation detection task, the delay increased to 477.4 ms on average with a SD of 46.3 ms and the distribution was approximately normal (see **Figure 3B**). This increase can be very well explained by the use of blocks of 500 ms and the inter-block interval of 125 ms.

Consequences and possible optimizations

The aforementioned benchmark results show that the delay and the jitter introduced by ConSole are reasonably low to conduct real-time experiments that analyze the power of oscillations. To further optimize for the delay and jitter between the input and the system's reaction the amount of data averaged in one block could be decreased as well as the inter-block interval. While the latter would only increase the load on the computer system used to run the software on, decreasing the amount of data analyzed in each block would also increase the susceptibility to random fluctuations (noise) of the analysis. For the benchmarking and the first example in this paper, we chose a window of 5 cycles of the center frequency. This corresponds to window-lengths commonly used in offline analysis settings that range from about 1.5–7 cycles (Tallon-Baudry et al., 1996; Rizzuto et al., 2003; Fujioka and Ross, 2008; Romei et al., 2008b; Schubert et al., 2009). As we did not average over trials, we decided to use a window-size from the upper range. Therefore with regards to online power analysis, there is an inherent trade-off that the researcher is faced with in setting the optimal parameters. A good strategy in general would be to run some offline experiments first and to use ConSole's possibility to simulate an online experiment using previously recorded data (see Simulating a Data Source). By this means the researcher has the possibility to optimize the code in order to be temporally as close to the events of interest as possible. In general "brain states" marked by an extended period (on the range of hundreds of milliseconds) of an increase/decrease of power are optimal for investigation using ConSole. Triggering on short-lived "bursts" of an oscillation is not recommended, at least with the standard analysis methods at hand.

The mentioned temporal limitations – i.e., to initiate events (e.g., TMS) as closely to the neuronal event of interest as possible – applies also to the phase of an oscillation which has received great interest recently (e.g., Hanslmayr et al., 2005a; Busch et al., 2009; Mathewson et al., 2009). Theoretically it would be interesting to directly stimulate at peaks and troughs of an oscillation, however the current delay using standard methods is not sufficient for this endeavor. An exception to this may be very slow oscillations below 1 Hz which have also gained increased interest recently.

EXAMPLE 1: PHOSPHENES AND REAL-TIME ALPHA

INTRODUCTION

In recent years, an increasing number of articles on the functional role of spontaneous alpha oscillations have been published. Most of these studies use a task involving covert attention to a region left or right of a fixation cross, which detects a target in these areas. The great majority of these studies come to the conclusion that ongoing alpha oscillations are modulated by drawing attention to one area, either by an increase of power in ipsilateral cortical regions of the visual cortex or by a decrease contralateral to the attended visual hemifield (e.g., Thut et al., 2006). There is even evidence for a retinotopic organization of the modulation of ongoing alpha oscillations (Rihs et al., 2007). Overall, these studies give strong evidence that alpha represents a mechanism of active inhibition,

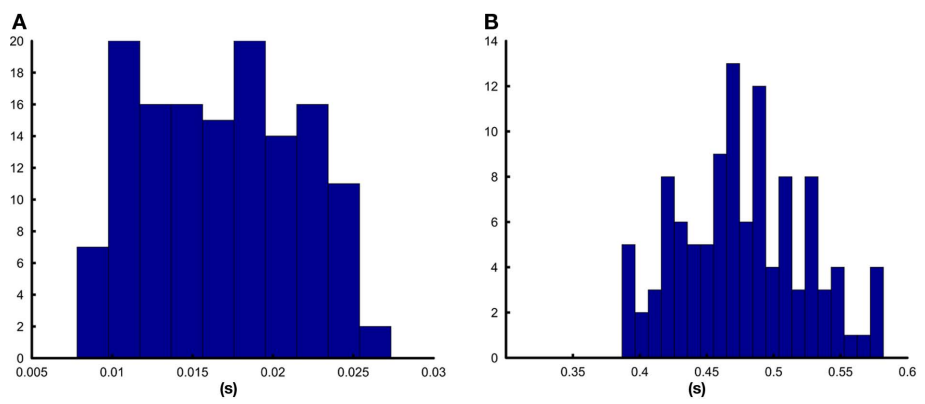


FIGURE 3 | Distribution of the delays between signal generation and reaction of ConSole. (A) Shows the histogram for the detection of a pulse, **(B)** shows the histogram for the detection of a 10-Hz oscillation.

pointing to a top-down mechanism modulating alpha oscillations. Although this assumption is most likely correct in regards to the aforementioned evidence, these studies do not answer the question of whether *spontaneous* fluctuations of these oscillations or even the different phases of the oscillation themselves have an impact on the processing of stimuli. The answer to the second question can be found in studies that account for the phase of the ongoing alpha oscillation immediately before stimulus presentation (Mathewson et al., 2009) or even earlier in the pre-stimulus period (Busch et al., 2009; Busch and VanRullen, 2010). Romei and Thut took an interesting approach to the first question and showed not only that the tonic level of occipital alpha predicts the interindividual phosphene threshold (Romei et al., 2008b), but also observed a dependency between fluctuations of ongoing alpha and the probability of phosphene perception when subjects where stimulated at phosphene threshold.

Although the dependency between ongoing alpha oscillations and phosphene perception could be causal, as mentioned above, offline studies only provide correlational evidence. In order to test the hypothesis that the likelihood of phosphene perception causally depends on the energy of spontaneous alpha oscillations, a real-time experiment is needed that can specifically stimulate the primary visual cortex at those points in time when alpha is either low or high.

METHODS

Six subjects (four female, mean age ± SD: 24 ± 3.9 years) were screened and trained according to procedures previously used in similar studies (Romei et al., 2008a,b) after giving written informed consent. The procedure was approved by the local ethics committee. Screening and training was necessary in order to select only those participants who reported seeing phosphenes and to train the reliability of their answers. We blindfolded the participants and used single-pulse TMS to elicit the perception. Neuronavigated TMS (Magstim Rapid2) was used to reproduce stimulation sites between sessions. In the last session, TMS pulses were triggered depending on the current alpha power in primary and secondary visual areas, which was computed using a minimum-norm estimation (Hämäläinen and Ilmoniemi, 1994) relative to a 5-min baseline measurement. The participants performed four runs of 5 min each. A TMS pulse was triggered when the current alpha power fell into the upper or lower quartile of the distribution acquired in the baseline run. To achieve a balance between low and high alpha power trials, the system did not react on high alpha if the total amount of high alpha trials exceeded the total amount of low alpha trials by two. The same strategy was used to limit the amount of low alpha trials with respect to high alpha trials. The participants indicated with a mouse click whether or not they had perceived a phosphene.

In order to verify the effectiveness of ConSole and the applied algorithms as well as to test the hypothesis that phosphenes are more likely to be seen when occipital alpha is low, we cut the data into trials 3000 ms before to 500 ms after the TMS pulse. The epochs were then subjected to a detrending algorithm in order to remove linear trends from the data. Afterward, the data were bandstop filtered (fourth order forward Butterworth filter; frequency range 49–51 Hz) in order to eliminate possible power line noise from the data. As the experiment relied on stimulating the participants at their individual threshold intensity, we discarded all runs that, according to a binomial test, did not show a balance between seen and unseen trials. Data from the remaining runs were then grouped into four categories (Seen/Not Seen High/Low Alpha) and visually inspected for artifacts in the time range between 1000- and 50-ms pre-stimulus. As we were interested in effects relative to the stimulation side, we mirrored data from the left hemisphere for those participants who were stimulated at the right hemisphere and *vice versa*. We then applied a time–frequency analysis using variable window-sizes of four times the wavelength of the respective frequency tapered with a Hanning window. We used a non-parametric, cluster-based statistic with 1000 randomizations to control for type 1 errors (Maris and Oostenveld, 2007). In this paper we show the preliminary results of six participants.

RESULTS

On average, ConSole triggered 71 ± 20 TMS pulses for each subject. Fifty-two percent of those trials were identified as low alpha.

To test the validity of the separation between low and high alpha trials, we applied a cluster-based statistic on the time–frequency representation of the data from trials identified as high versus low alpha by ConSole. The algorithm found one significant cluster ($p < 0.001$) that confirmed higher alpha from 700 to 125 ms before the stimulation (see **Figure 4**). Alpha was increased rather globally, however the maximum increase was found at the site of stimulation as targeted by ConSole. The data also show a good timing accuracy. The window for the Hilbert transform of the paradigm was 500 ms. According to the aforementioned results, we would have expected delays of about 400–500 ms. The data, however, show that the maximum difference is between 350- and 175-ms pre-stimulation. Interestingly, higher synchronization in trials identified as containing high alpha is not restricted to the alpha band. Higher synchronization also extends into the beta range.

Although our results are similar to those reported by Romei et al. (2008a) in that, on average, trials in which no phosphene was perceived were preceded by higher alpha, we were not able to show that pre-stimulus alpha power predicted the probability of the participant seeing a phosphene by comparing the responses to high and low alpha trials. At first glance, these results do not seem to fit together. This could be due to the small number of participants, not optimally defined thresholds or a more complex relationship between alpha and phosphene perception that a "simple" offline contrast of "seen" versus "unseen" would suggest. Interestingly, however, when comparing only the high alpha trials in which the phosphene was seen versus those in which it was not seen, a strong trend in frontal areas is revealed. If alpha was high under the stimulated site, alpha had to also be high in frontal areas in order to produce a phosphene. This could serve as a starting point for a more in-depth analysis, as this result could imply that the connectivity between these regions is relevant, but only if alpha is high in occipital areas.

DISCUSSION

One purpose of this study was to test whether ConSole was capable of separating low and high alpha trials correctly and with reasonable delay in a real online setting. Our results show that this

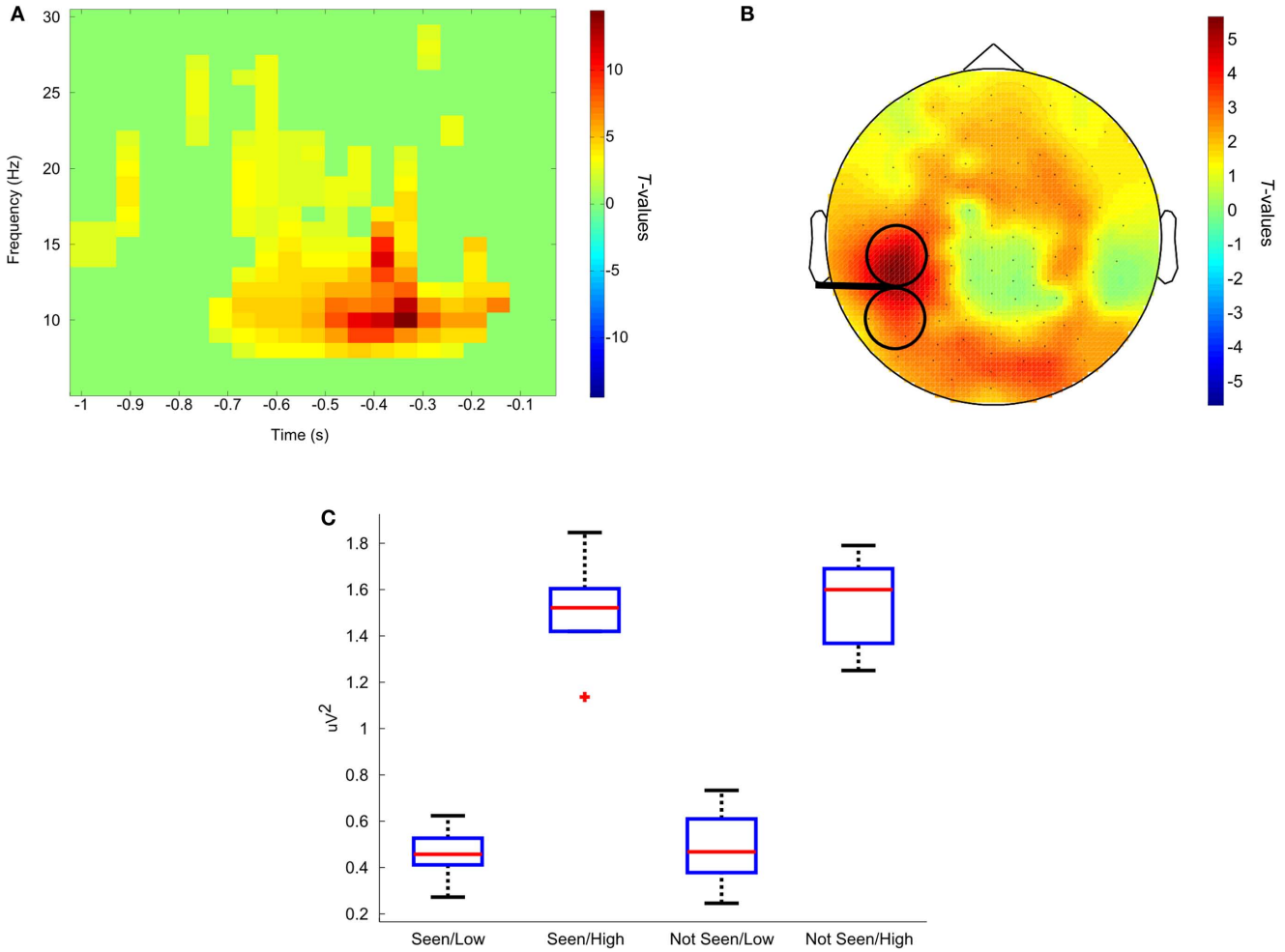


FIGURE 4 | Results of example 1. (A,B) Show the comparison between trials identified as high alpha versus low alpha. **(A)** Time course at one representative occipital electrode. The maximum difference is between 400 and 300 ms before ConSole sent the trigger. **(B)** Topography at the point of maximum difference. Higher alpha is not restricted to the area analyzed but extends to the other hemifield as well as to frontal areas. The stimulated area is depicted by the TMS coil sketch. **(C)** Box and whiskers plot showing the distribution of alpha power for all four conditions. The red lines represent the medians, the edges of the boxes mark the first and third quartile. Whiskers extend to the most extreme value not considered an outlier. Outliers are values that exceed the first or third quartile by 1.5 times the total range between the first and third quartile.

separation was perfect but could be optimized to minimize the delay between classification (high versus low alpha) and stimulus onset. One concern for real-time experiments is that the amount of trials cannot be predicted. The example however shows that on average 71 trials out of the maximum number of 120 were triggered. The balance between high and low identified trials was guaranteed by limiting the amount of imbalance by definition in the paradigm file.

Another purpose of the study was to test the hypothesis of Romei et al. (2008a) that the perception of phosphenes depends on the current level of alpha synchronization over the occipital cortex at the stimulated hemisphere. Although the offline analysis of the data points in the same direction, the online analysis seems at odds with a simple "high" versus "low" alpha functional distinction since we did not find any difference in the probability of seeing a phosphene in trials with high or low alpha as we had originally hypothesized. This might be due to the small number of participants, as the results shown are only preliminary. More interestingly, however, we observed that if alpha was high under the stimulated area, frontal areas had to be synchronized in the alpha band as well in order to produce the perception of a phosphene, thereby suggesting the influence of a connection between these two regions. Overall, the data suggest that the "high" alpha category is functionally more diverse than simply reflecting an inhibitory state and that under certain conditions it may even favor a perception near-threshold. In future approaches we will scrutinize the actual distribution of the underlying electrophysiological marker in more depth, as one further possibility for not seeing the hypothesized results might also be that "high alpha" does not only cover the uppermost 25% of the distribution but extends into the low alpha range far deeper than we had expected. This might also indicate that the relationship between alpha power and behavior is not linear, as assumed in the study, but, for instance, logarithmic.

Furthermore, we plan to extend this approach to other modalities such as the visual and motor system, as alpha is believed to have a common functional meaning (Weisz et al., submitted) – at least in primary and secondary sensory and motor cortical areas.

EXAMPLE 2: TREATING TINNITUS WITH NEUROFEEDBACK

INTRODUCTION

Neurofeedback is generally defined as the operant conditioning of signals acquired from the brain via various methods such as EEG, MEG, fMRI. For almost 40 years, this approach has been used to teach patients how to normalize abnormal brain states (Sterman and Friar, 1972; Lubar and Shouse, 1976; Rockstroh et al., 1993; Masterpasqua and Healey, 2003; Dohrmann et al., 2007a,b). Recent studies by our group have shown significant improvements of tinnitus loudness and distress in those patients who were able to learn to normalize temporal alpha and/or delta activity but not in those who failed to learn the task (Dohrmann et al., 2007a,b). To increase the number of patients who are able to learn the task and thus benefit from the training, we used the fact that sensory alpha rhythms desynchronize on sensory input – the extent to which albeit depending strongly on top-down factors such as attention (Bastiaansen et al., 2001; Müller et al., submitted; Hartmann et al., submitted).

METHODS

Nine otherwise healthy patients (one female, mean age ± SD: 57 ± 8.8 years) suffering from chronic tinnitus were recruited via advertisements in the local newspaper. All patients were informed about the training and gave written consent. The procedure was approved by the local ethics committee. The patients took part in 10 session within 3 weeks. Each session consisted of one baseline measurement to calibrate the neurofeedback system, four training runs, and another baseline measurement after the training to assess changes in cortical activity within each session. In the training runs, patients were shown a feedback on a screen for 5 s without hearing a tone (see **Figure 5A**). They were instructed to consider this period as a baseline that showed how auditory areas of their brain behaved without any input. Afterward, patients were stimulated with a sound that was filtered to match their tinnitus sound as closely as possible. Because of the aforementioned effect, patients saw a decrease in alpha activity via the feedback. They were instructed that one possible strategy for increasing alpha activity was to ignore the sound. Baseline measurements before and after the training runs differed from those only by not providing feedback to the patients. The patients were instructed to passively listen to the sounds with eyes open.

For the training, data acquired from 29 electrodes on the scalp and 2 electrodes beside and above the right eye to facilitate artifact correction were sent to ConSole, highpass filtered, average referenced, and artifact corrected (via ICA). The data were then projected onto eight regional sources. The data of the two temporal sources were subjected to a frequency analysis and the relative energy at the individual's alpha frequency was fed back onto a computer screen. We analyzed the power at the individual's alpha frequency before and after each training.

RESULTS

To first test whether the participants in the neurofeedback study were able to learn how to modulate alpha power, we used a linear mixed-effects model with normalized alpha power at both temporal sources as the dependent variable. The individual session and the time of measurement (pre or post-training) were the independent variables. The advantage of a linear mixed-effects model is the possibility to add so-called random variables that control for variance between individuals. "Participant" was thereby chosen as the random variable. Our results show that, on average, alpha increased within each session ($p < 0.001$) by 19% (SD: 26%) as well as between the sessions ($p < 0.05$) as shown by an increase from the first to the last session by 38% (SD: 68%; see **Figure 5B**). A similar result was drawn by the behavioral results. We used a custom-made questionnaire with six Likert scales about different features of the tinnitus, like loudness, annoyance. On average, the total score decreased by 13% (SD: 17%) within the sessions ($p < 0.001$) as well as between them ($p < 0.05$) as shown by a decrease from the first to the last session by 11% (SD: 7%; see **Figure 5C**).

DISCUSSION

The increasing amount of studies into the effect of neurofeedback on tinnitus show that, though patients benefit from the treatment, these effects are moderate and subject to strong interindividual variability. This might be owing to a number of reasons that remain unclear. In general, three factors are important for a successful neurofeedback training: contingent feedback about the correct features acquired from the brain, good adjustment to the demands of the task, and keeping the patient motivated. The approach we have presented in this example shows another way of neurofeedback training. The results are comparable to those of recent studies in terms of how much the patients improved. What is however more important is that we have demonstrated that ConSole provides scientists with a tool for rapidly developing and testing new approaches to the treatment of conditions including (but not limited to) tinnitus using neurofeedback. New trainings might include more tailor-made designs like QEEG, different frequency bands, interareal functional connectivity or other ways of giving feedback to the patient, just to name a few possibilities. ConSole not only facilitates the development of these new approaches but also makes testing at multiple sites with different equipment easier due to its independence from the acquisition device used.

EXAMPLE 3: MONITORING TMS-INDUCED MEPs

INTRODUCTION

If a strong and brief magnetic pulse is applied to brain tissue associated with motor functioning via TMS, changes in corticospinal excitability can be observed (e.g., Van Der Werf and Paus, 2006). In suprathreshold stimulations, these changes result in a typical pattern of periphysiological muscle activity. This summation of electromyographic, response is commonly called MEP (for further details refer to Barker et al., 1985; Rothwell et al., 1999; Di Lazzaro et al., 2004).

Motor-evoked potential amplitude and shape depend on several factors, such as TMS pulse duration (Rothkegel et al., 2010), coil orientation (e.g., Mills et al., 1992), distance from the coil to

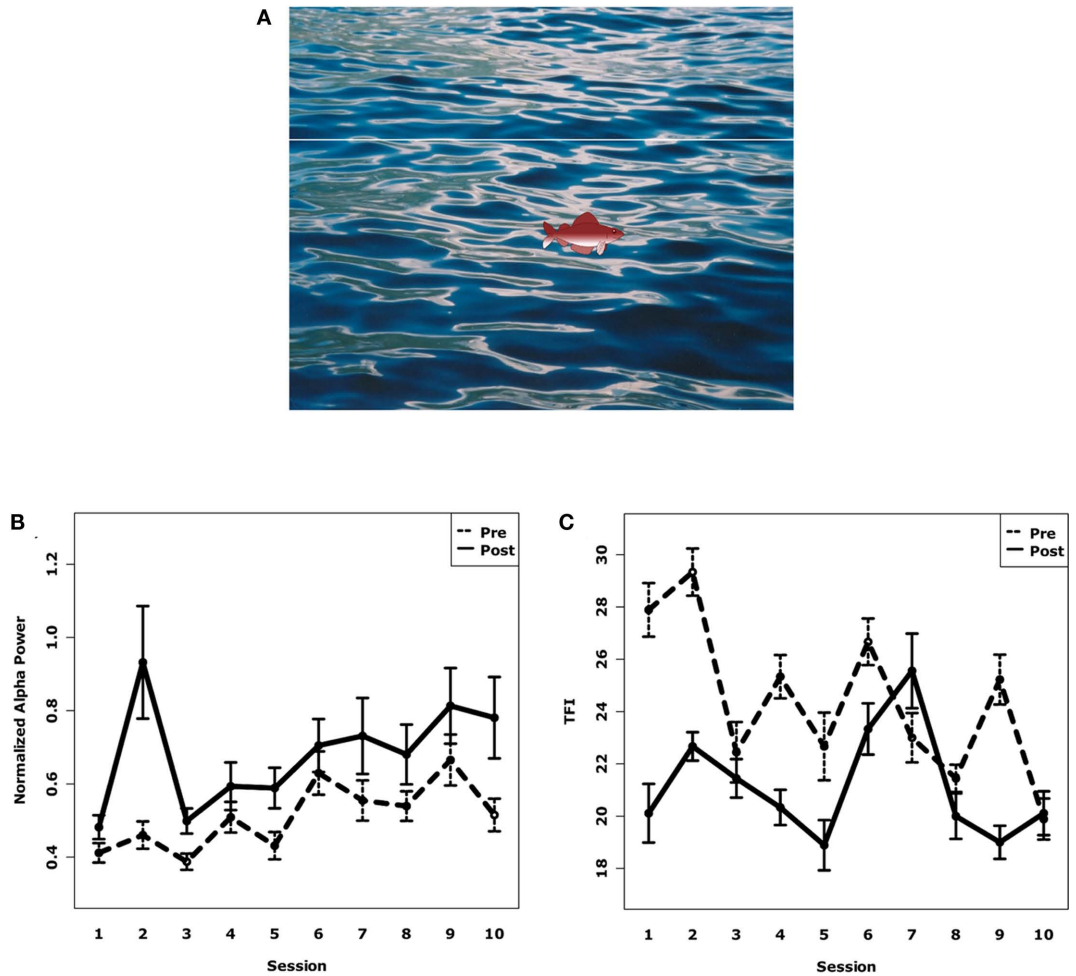


FIGURE 5 | (A) Screenshot of the patient's training screen. The fish takes 10 s to move from the left to the right of the screen. The first 5 s are the "baseline" period for the patient without any stimulation. In the second half, the patients were stimulated with a sound that resulted in an alpha desynchronization. The patient's task was then to increase temporal alpha power which was indicated by the height of the displayed fish. The patient was rewarded after the trial if the fish stayed above the target line for a sufficient amount of time. **(B)** Normalized alpha power of all subjects over all 10 sessions before and after neurofeedback training. Alpha power increased significantly within and between sessions. Error bars denote SE. **(C)** Distress rating of all subjects over all 10 sessions before and after neurofeedback training. Distress was reduced significantly within and between sessions. Error bars denote SE.

the underlying cortex (Kozel et al., 2000), or individual biological differences (Wassermann, 2002). We ran a number of combined EEG–TMS studies with the aim of investigating oscillatory brain activity and its influence on MEPs. In order to achieve a comparable reference value between subjects, we determined resting motor thresholds using a Matlab-based decision-making algorithm within the ConSole environment. Data were acquired and preprocessed using the ConSole framework and the built-in modules. Although this could have also been carried out entirely in Matlab (apart from the acquisition), using the modules provided by ConSole is more convenient and much faster. The part of the paradigm that involved the special analysis of the MEP was confined to a block of data specific to the time range when a MEP was expected to occur. The analysis did therefore not have to run constantly and was not time-critical. This fact and the greater flexibility in having a Matlab script qualified this part of the online analysis for execution in Matlab.

METHODS

Surface electromyography was recorded from the right first interosseous muscle (FDI) and was sampled at 2048 Hz. ConSole acquired the voltage difference between two surface electrodes attached in a belly tendon montage with a ground electrode taped to the right ear contralateral to TMS Stimulation (Magstim Rapid2). The data were highpass filtered (third order, cutoff: 4 Hz). Upon receiving the trigger representing the TMS pulse, ConSole sent a corresponding block of data via its Matlab interface to a routine that first rectified the epoch and then searched for a peak of >50 mV. The result was fed back to ConSole and displayed on the screen as a text message.

Neuronavigated high intensity single-pulse TMS (on average 60% stimulator output) was applied to the left hand knob area with the handle of the TMS coil pointing backward approximately 45° to the midsagittal line. Coil position was adjusted until absolute FDI MEP amplitude was maximal in three consecutive trials and

was marked as so-called "hot spot." Afterward the resting motor threshold was determined by a maximum likelihood paradigm using the software MTAT 2.0 (Awiszus, 2003).

RESULTS AND DISCUSSION

By using the aforementioned setup demonstrating the capabilities of interactions between ConSole and Matlab, we were able to receive reliable feedback about whether a TMS pulse had triggered a MEP (see **Figure 6** for a schematic example of the feedback in case a MEP was detected and in case, the TMS failed to elicit an MEP). This example demonstrates the use of ConSole to analyze peripheral data in response to a stimulus. Future applications are not limited to TMS and MEPs but may include, for instance, reactions of peripheral signals like skin conductance on electrical or tactile stimulation.

DISCUSSION

The analysis of brain activity acquired by EEG or MEG on a single trial level is becoming an increasingly important topic. Treating the variability between trials as a valuable source of information and not as noise that is to be discarded will lead to great advances in our understanding of how brain activity relates to behavior. However, the analysis of data on a trial-by-trial basis raises the importance of good data quality during acquisition as well as advanced analysis strategies. Nevertheless, the offline analysis of the data does not exploit the full potential of the methods available as stimulation is done at random points in time and thus completely independent of the current brain state. Such a correlation approach makes sense as long as prior knowledge about the effect is small and the underlying hypothesis cannot be strictly defined. However, as soon as a clear hypothesis about the functional significance of distinct brain states exists, a more causal test of the hypothesis is favorable. Analyzing EEG or MEG in real-time is an approach that takes advantage of prior knowledge by controlling the experiment depending on those points in time that putatively mark a brain state of interest. This approach not only leads to a stronger test of the hypothesis but can also provide more in-depth insights in phenomena related to the primary effect (e.g., whether alpha is a correlate of active inhibition). In contrast to the offline correlational approach, differences between physiologically clearer defined brain states can be analyzed. For example, in the study comparing the effect of high and low alpha power on the perception of a phosphene, our online approach, despite superb separation of high and low posterior alpha, indicates that high levels of alpha may be functionally not a unitary phenomenon (e.g., inhibition). The contrast within the high alpha trials points to a critical role of frontal regions and suggests that, in some cases, high alpha activity may reflect a more efficient communication between sensory and frontal regions – potentially implemented via specific phase relationships (Busch et al., 2009).

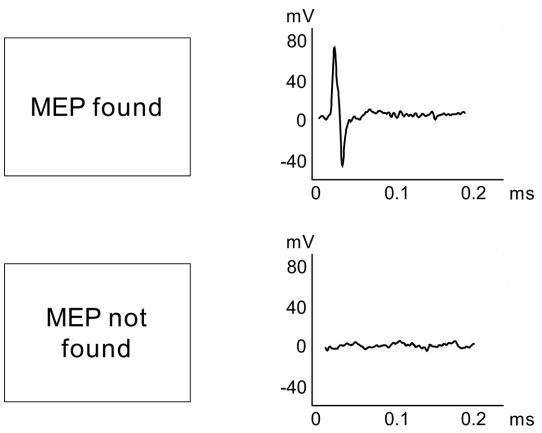


FIGURE 6 | Schematic drawing of the setup to automatically identify TMS-induced MEPs.

In the current paper, we present a flexible software environment that helps scientists in conducting many kinds of real-time experiments. Its modular approach and clear file format makes setting up experiments easy, even without any knowledge of computer programming.

One of the most important requirements of a real-time EEG/MEG application is the exact and fast timing of the triggers that are used to stimulate the participant. Current computer systems with multi-core processors and gigabytes of RAM are generally fast enough to deal with the acquisition and processing of 128+ channels at high sampling rates. Our system with an Intel Core I7-970 and 4 GB of RAM is easily capable of handling the above presented experimental setups at very low processor load and with acceptable delay and jitter.

Three current examples from our laboratory were given to illustrate that ConSole is a flexible tool that can be universally applied in numerous real-time paradigms and that extends the possibilities in these areas. Apart from the envisaged real-time cognitive neuroscience experiments (Example 1), ConSole can be used, for instance, in clinical settings for neurofeedback (Example 2) or to obtain online feedback about the presence or absence of a peripheral physiological response (Example 3). The latter example was mainly intended to demonstrate the ConSole/Matlab interface, further facilitating the process of creating a real-time analysis-flow, even for scientists not experienced in programming languages like C++.

In conclusion we wish to promote the idea that, apart from conducting offline data analysis, cognitive neuroscience-specific hypotheses should in the future be tested in real-time experiments. This puts high demands on data quality and online signal processing. Furthermore, in order to enable as many neuroscientists as possible to pursue this research strategy, user-friendly, and open-source software frameworks are needed. With ConSole, we attempt to make such a framework publically available and have given first proof-of-concept evidence with regards to its functionality.

ACKNOWLEDGMENTS

This work was supported by the Deutsche Forschungsgemeinschaft (WE 4156/2-1), the TRI (TE 06 02), and the Ministerium für Wissenschaft, Forschung und Kunst Baden Württemberg. The authors wish to thank Gabriela Salagean, Sylvie Roth, Hadas Gorodetzky, Teresa Übelacker, and Christiane Wolf for their help in acquiring the data for the examples.

REFERENCES

Awiszus, F. (2003). TMS and threshold hunting. *Suppl. Clin. Neurophysiol.* 56, 13–23.

Baccalá, L. A., and Sameshima, K. (2001). Partial directed coherence: a new concept in neural structure determination. *Biol. Cybern.* 84, 463–474.

Barker, A. T., Jalinous, R., and Freeston, I. L. (1985). Non-invasive magnetic stimulation of human motor cortex. *Lancet* 1, 1106.

Bastiaansen, M. C., Böcker, K. B., and Brunia, C. H. (2001). Event-related desynchronization during anticipatory attention for an upcoming stimulus: a comparative EEG/MEG study. *Clin. Neurophysiol.* 112, 393–403.

Busch, N. A., Julien, D., and VanRullen, R. (2009). The phase of ongoing EEG oscillations predicts visual perception. *J. Neurosci.* 29, 7869–7876.

Busch, N. A., and VanRullen, R. (2010). Spontaneous EEG oscillations reveal periodic sampling of visual attention. *Proc. Natl. Acad. Sci. U.S.A.* 107, 16048–16053.

Cardoso, J. F., and Souloumiac, A. (2002). Blind beamforming for non-Gaussian signals. *Radar Signal Process. IEE Proc. F* 140, 362–370.

Cuffin, B. N., and Cohen, D. (1979). Comparison of the magnetoencephalogram and electroencephalogram. *Electroencephalogr. Clin. Neurophysiol.* 47, 132–146.

Di Lazzaro, V., Oliviero, A., Pilato, F., Saturno, E., Dileone, M., Mazzone, P., Insola, A., Tonali, P. A., and Rothwell, J. (2004). The physiological basis of transcranial motor cortex stimulation in conscious humans. *Clin. Neurophysiol.* 115, 255–266.

Dohrmann, K., Elbert, T., Schlee, W., and Weisz, N. (2007a). Tuning the tinnitus percept by modification of synchronous brain activity. *Restor. Neurol. Neurosci.* 25, 371–378.

Dohrmann, K., Weisz, N., Schlee, W., Hartmann, T., and Elbert, T. (2007b). Neurofeedback for treating tinnitus. *Prog. Brain Res.* 166, 473.

Ergenoglu, T., Demiralp, T., Bayraktaroglu, Z., Ergen, M., Beydagi, H., and Uresin, Y. (2004). Alpha rhythm of the EEG modulates visual detection performance in humans. *Brain Res. Cogn. Brain Res.* 20, 376–383.

Fox, M. D., and Raichle, M. E. (2007). Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat. Rev. Neurosci.* 8, 700–711.

Free Software Foundation. (1991). *General Public License Version 2*. Available at: http://www.gnu.org/licenses/gpl-2.0.html

Frigo, M., and Johnson, S. G. (2005). The design and implementation of FFTW3. *Proc. IEEE* 93, 216–231.

Friston, K. (2002). Functional integration and inference in the brain. *Prog. Neurobiol.* 68, 113–143.

Fujioka, T., and Ross, B. (2008). Auditory processing indexed by stimulus-induced alpha desynchronization in children. *Int. J. Psychophysiol.* 68, 130–140.

Hämäläinen, M. S., and Ilmoniemi, R. J. (1994). Interpreting magnetic fields of the brain: minimum norm estimates. *Med. Biol. Eng. Comput.* 32, 35–42.

Hanslmayr, S., Aslan, A., Staudigl, T., Klimesch, W., Herrmann, H. S., and Bäuml, K. H. (2007). Prestimulus oscillations predict visual perception performance between and within subjects. *Neuroimage* 37, 1465–1473.

Hanslmayr, S., Klimesch, W., Sauseng, P., Gruber, W., Doppelmayr, M., Freunberger, R., and Pecherstorfer, T. (2005a). Visual discrimination performance is related to decreased alpha amplitude but increased phase locking. *Neurosci. Lett.* 375, 64–68.

Hanslmayr, S., Sauseng, P., Doppelmayr, M., Schabus, M., and Klimesch, W. (2005b). Increasing individual upper alpha power by neurofeedback improves cognitive performance in human subjects. *Appl. Psychophysiol. Biofeedback* 30, 1–10.

Jensen, O., and Jokisch, D. (2007). Modulation of gamma and alpha activity during a working memory task engaging the dorsal or ventral stream. *J. Neurosci.* 27, 3244–3251.

Kenet, T., Bibitchkov, D., Tsodyks, M., Grinvald, A., and Arieli, A. (2003). Spontaneously emerging cortical representations of visual attributes. *Nature* 425, 954–956.

Klimesch, W., Doppelmayr, M., Russegger, H., Pachinger, T., and Schwaiger, J. (1998). Induced alpha band power changes in the human EEG and attention. *Neurosci. Lett.* 244, 73–76.

Kozel, F. A., Nahas, Z., DeBrux, C., Molloy, M., Lorberbaum, J. P., Bohning, D., Risch, S. C., and George, M. S. (2000). How coil-cortex distance relates to age, motor threshold, and antidepressant response to repetitive transcranial magnetic stimulation. *J. Neuropsychiatry Clin. Neurosci.* 12, 376.

Lubar, J., and Shouse, M. (1976). EEG and behavioral changes in a hyperkinetic child concurrent with training of the sensorimotor rhythm (SMR). *Appl. Psychophysiol. Biofeedback* 1, 293–306.

Macwhinney, B., Provost, J., and Cohen, J. (1997). The PsyScope experiment-building system. *Spat. Vis.* 11, 99–101.

Masterpasqua, F., and Healey, K. N. (2003). Neurofeedback in psychological practice. *Prof. Psychol. Res. Pr.* 34, 652–656.

Maris, E., and Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* 164, 177–190.

Mathewson, K. E., Fabiani, M., Gratton, G., Beck, D. M., and Lleras, A. (2010). Rescuing stimuli from invisibility: inducing a momentary release from visual masking with pre-target entrainment. *Cognition* 115, 186–191.

Mathewson, K. E., Gratton, G., Fabiani, M., Beck, D. M., and Ro, T. (2009). To see or not to see: prestimulus alpha phase predicts visual awareness. *J. Neurosci.* 29, 2725.

Mills, K. R., Boniface, S. J., and Schubert, M. (1992). Magnetic brain stimulation with a double coil: the importance of coil orientation. *Electroencephalogr. Clin. Neurophysiol.* 85, 17–21.

Octave-Forge. (2010). *Octave-Forge*. Available at: http://octave.sourceforge.net/index.html

Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J.-M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* 2011, 1–9.

Perrin, F., Pernier, J., Bertrand, O., and Echallier, J. F. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalogr. Clin. Neurophysiol.* 72, 184–187.

Perrin, F., Pernier, J., Bertrand, O., and Echallier, J. F. (1990). Corrigenda. *Electroencephalogr. Clin. Neurophysiol.* 76, 565–566.

Renard, Y., Lotte, F., Gibert, G., Congedo, M., Maby, E., Delannoy, V., Bertrand, O., and Lécuyer, A. (2010). OpenViBE: an open-source software platform to design, test, and use brain–computer interfaces in real and virtual environments. *Presence* 19, 35–53.

Rihs, T. A., Michel, C. M., and Thut, G. (2007). Mechanisms of selective inhibition in visual spatial attention are indexed by alpha-band EEG synchronization. *Eur. J. Neurosci.* 25, 603–610.

Rizzuto, D. S., Madsen, J. R., Bromfield, E. B., Schulze-Bonhage, A., Seelig, D., Aschenbrenner-Scheibe, R., and Kahana, M. J. (2003). Reset of human neocortical oscillations during a working memory task. *Proc. Natl. Acad. Sci. U.S.A.* 100, 7931–7936.

Rockstroh, B., Elbert, T. R., Birbaumer, N., Wolf, P., Düchting-Röth, A., Reker, M., Daum, I., Lutzenberger, W., and Dichgans, J. (1993). Cortical self-regulation in patients with epilepsies. *Epilepsy Res.* 14, 63–72.

Romei, V., Brodbeck, V., Michel, C., Amedi, A., Pascual-Leone, A., and Thut, G. (2008a). Spontaneous fluctuations in posterior alpha-band EEG activity reflect variability in excitability of human visual areas. *Cereb. Cortex* 18, 2010–2018.

Romei, V., Rihs, T., Brodbeck, V., and Thut, G. (2008b). Resting electroencephalogram alpha-power over posterior sites indexes baseline visual cortex excitability. *Neuroreport* 19, 203–208.

Romei, V., Gross, J., and Thut, G. (2010). On the role of prestimulus alpha rhythms over occipito-parietal areas in visual input regulation: correlation or causation? *J. Neurosci.* 30, 8692–8697.

Rothkegel, H., Sommer, M., Paulus, W., and Lang, N. (2010). Impact of pulse duration in single pulse TMS. *Clin. Neurophysiol.* 121, 1915–1921.

Rothwell, J. C., Hallett, M., Berardelli, A., Eisen, A., Rossini, P., and Paulus, W. (1999). Magnetic stimulation: motor evoked potentials. The international federation of clinical neurophysiology. *Electroencephalogr. Clin. Neurophysiol. Suppl.* 52, 97–103.

Sauseng, P., Klimesch, W., Gerloff, C., and Hummel, F. C. (2009). Spontaneous locally restricted EEG alpha activity determines cortical excitability in the motor cortex. *Neuropsychologia* 47, 284–288.

Schalk, G., McFarland, D. J., Hinterberger, T., Birbaumer, N., and Wolpaw, J. R. (2004). BCI2000: a general-purpose brain-computer interface (BCI) system. *IEEE Trans. Biomed. Eng.* 51, 1034–1043.

Scherg, M. (1990). Fundamentals of dipole source potential analysis. Auditory evoked magnetic fields and electric potentials. *Adv. Audiol.* 5, 40–69.

Schubert, R., Haufe, S., Blankenburg, F., Villringer, A., and Curio, G. (2009). Now you'll feel it, now you won't: EEG rhythms predict the effectiveness of perceptual masking. *J. Cogn. Neurosci.* 21, 2407–2419.

Sterman, M. B., and Friar, L. (1972). Suppression of seizures in an epileptic following sensorimotor EEG feedback training. *Electroencephalogr. Clin. Neurophysiol.* 33, 89–95.

Tallon-Baudry, C., and Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends Cogn. Sci.* 3, 151–162.

Tallon-Baudry, C., Bertrand, O., Delpuech, C., and Pernier, J. (1996). Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. *J. Neurosci.* 16, 4240–4249.

Thut, G., Nietzel, A., Brandt, S. A., and Pascual-Leone, A. (2006). Alpha-band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. *J. Neurosci.* 26, 9494–9502.

Tuladhar, A. M., Huurne, N., Schoffelen, J. M., Maris, E., Oostenveld, R., and Jensen, O. (2007). Parieto-occipital sources account for the increase in

alpha activity with working memory load. *Hum. Brain Mapp.* 28, 785–792.
Van Der Werf, Y. D., and Paus, T. (2006). The neural response to transcranial magnetic stimulation of the human motor cortex. I. Intracortical and cortico-cortical contributions. *Exp. Brain Res.* 175, 231–245.
van Dijk, H., Schoffelen, J. M., Oostenveld, R., and Jensen, O. (2008). Prestimulus oscillatory activity in the alpha band predicts visual discrimination ability. *J. Neurosci.* 28, 1816.
Varela, F., Lachaux, J. P., Rodriguez, E., and Martinerie, J. (2001). The brainweb: phase synchronization and large-scale integration. *Nat. Rev. Neurosci.* 2, 229–239.
Wassermann, E. M. (2002). Variation in the response to transcranial magnetic brain stimulation in the general population. *Clin. Neurophysiol.* 113, 1165–1171.
Worden, M. S., Foxe, J. J., Wang, N., and Simpson, G. V. (2000). Anticipatory biasing of visuospatial attention indexed by retinotopically specific alpha-band electroencephalography increases over occipital cortex. *J. Neurosci.* 20, RC63.
Wyart, V., and Tallon-Baudry, C. (2009). How ongoing fluctuations in human visual cortex predict perceptual awareness: baseline shift versus decision bias. *J. Neurosci.* 29, 8715–8725.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 19 January 2011; paper pending published: 06 February 2011; accepted: 22 February 2011; published online: 09 March 2011.

Citation: Hartmann T, Schulz H and Weisz N (2011) Probing of brain states in real-time: introducing the ConSole environment. Front. Psychology ***2****:36. doi: 10.3389/fpsyg.2011.00036*
This article was submitted to Frontiers in Perception Science, a specialty of Frontiers in Psychology.

APPENDIX

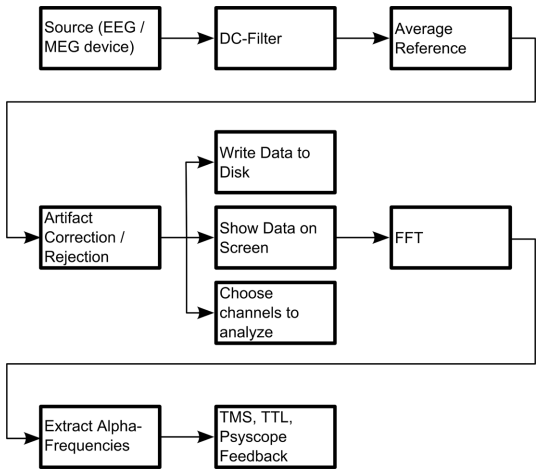


FIGURE A1 | Schematic of a typical ConSole paradigm. Data are acquired from the source and propagated from module to module, each further processing the data. The data can also be sent to multiple modules. Eventually, ConSole uses the final results to trigger an event like a TMS pulse or the display of a stimulus.

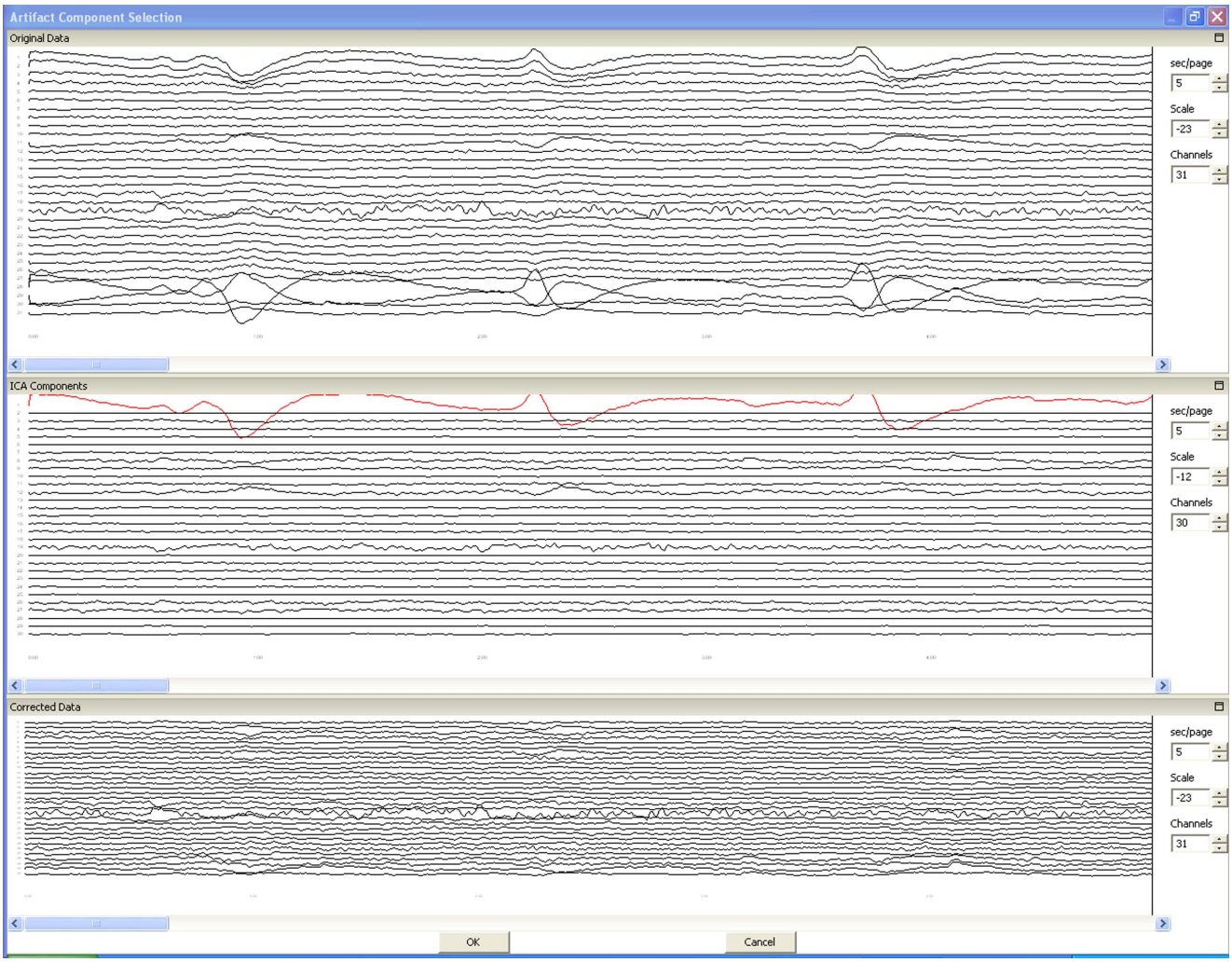


FIGURE A2 | This figure shows the artifact component selection process of ConSole's built-in artifact correction. An ICA algorithm was applied to calibration data. The upper panel shows the original data to identify the time points when artifacts were present. The middle panel shows the components calculated by the ICA algorithm. The user selects components representing artifacts by clicking on the number to the left. The lower panel shows the resulting data after correction to verify the outcome.

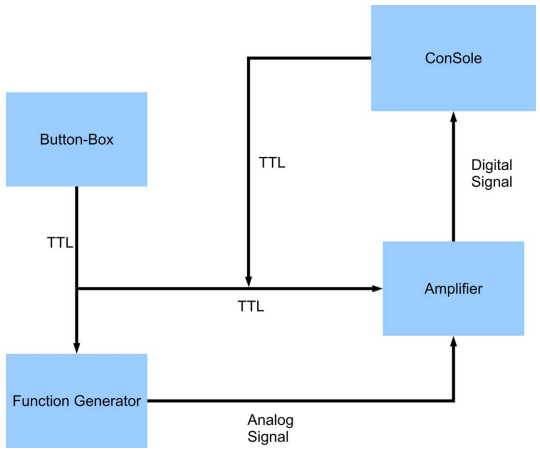


FIGURE A3 | Schematic overview of the setup to measure the delay and jitter of ConSole. The Button-Box sends a TTL pulse to the Function Generator and the amplifier. The Function Generator immediately sends the analog signal (negative pulse or 10 Hz sine wave) to the amplifier. Both the TTL pulse and the analog signal are recorded by ConSole. ConSole then sends a TTL pulse to the amplifier as soon as it reacts on the incoming analog signal generated by the Function Generator. This second TTL pulse is also acquired by the amplifier and recorded by ConSole. The difference between the two TTL pulses is the delay between signal generation and ConSole's reaction.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 20 May 2011
doi: 10.3389/fpsyg.2011.00091

Adaptive thresholding for improving sensitivity in single-trial simultaneous EEG/fMRI

Megan deBettencourt, Robin Goldman, Truman Brown and Paul Sajda*

Department of Biomedical Engineering, Columbia University, New York, NY, USA

Edited by:
Guillaume A. Rousselet, University of Glasgow, UK

Reviewed by:
Cyril R. Pernet, University of Edinburgh, UK
Stefan Debener, University of Oldenburg, Germany

****Correspondence:***
Paul Sajda, Department of Biomedical Engineering, Columbia University, 404 CEPSR Building, 530 West 120th Street, New York, NY, USA.
e-mail: ps629@columbia.edu

A common approach used to fuse simultaneously recorded EEG and fMRI is to correlate trial-by-trial variability in the EEG, or variability of components derived therefrom, with the blood oxygenation level dependent response. When this correlation is done using the conventional univariate approach, for example with the general linear model, there is the usual problem of correcting the statistics for multiple comparisons. Cluster thresholding is often used as the correction of choice, though in many cases it is utilized in an *ad hoc* way, for example by employing the same cluster thresholds for both traditional regressors (stimulus or behaviorally derived) and EEG-derived regressors. In this paper we describe a resampling procedure that takes into account the *a priori* statistics of the trial-to-trial variability of the EEG-derived regressors in a way that trades off cluster size and maximum voxel *z*-score to properly correct for multiple comparisons. We show that this data adaptive procedure improves sensitivity for smaller clusters of activation, without sacrificing the specificity of the results. Our results suggest that extra care is needed in correcting statistics when the regressor model is derived from noisy and/or uncertain measurements, as is the case for regressors constructed from single-trial variations in the EEG.

Keywords: single-trial, simultaneous EEG/fMRI, multiple comparisons, resampling, auditory oddball

INTRODUCTION

It is becoming increasingly common to use single-trial EEG-derived values to model simultaneously acquired fMRI data. The combination of these two complementary neuroimaging modalities enables the variability of neural activity to be related to the blood oxygenation level dependent (BOLD) response (e.g., Debener et al., 2005; Benar et al., 2007; Goldman et al., 2009). Particularly exciting is that this type of fusion of modalities exploits variability that is not observable via behavioral responses, and thus provides a window into latent states of the human brain.

Many efforts combining simultaneous EEG and fMRI in this way employ the general linear model (GLM; Worsley and Friston, 1995). The GLM is a univariate approach which, when applied to massive datasets such as fMRI, requires correcting the statistics for multiple comparisons in order to properly perform hypothesis testing. The issue of multiple comparison correction has been addressed extensively in the literature and there are a number of procedures that have been adopted, both for fMRI analysis alone (Genovese et al., 2002; Nichols and Hayasaka, 2003) and for simultaneous EEG/fMRI (Debener et al., 2005; Eichele et al., 2005; Benar et al., 2007; Esposito et al., 2009; Goldman et al., 2009; Mayhew et al., 2010; Mulert et al., 2010; Scheibe et al., 2010; Novitskiy et al., 2011).

One concern that is specific to using the variability of EEG-derived regressors is the added uncertainty in the "interpretation" of the resulting statistics within the context of statistical parametric maps. This can be easily seen if we consider the sources of uncertainty for different types of regressors. In conventional fMRI, regressors might be constructed based on the timing of the stimulus presented to the subject (see **Figure 1A**). We know the times exactly and we have reason to believe that these times are relevant to the activity in the brain. The amplitudes of the regressor events are constant so there is no uncertainty in terms of their magnitude – i.e., a constant scale would not change the GLM analysis. We could also construct regressors derived from behavioral measures, such as response time (RT), and though this may vary across trial and subject, we can measure RTs very accurately and thus measurement noise is low. However there is additional uncertainty in the interpretation of the amplitude variations. For example a long RT (e.g., high amplitude on a trial) might be caused by attentional lapses, natural alertness fluctuations, or additional neural processing for increasing decision confidence. Thus, compared to the case in **Figure 1A**, the behavioral regressor in **Figure 1B** could be seen as having greater uncertainty.

When constructing regressors derived from EEG variability (**Figure 1C**), we must consider a number of factors affecting our uncertainty. One is related to the noise in the measurements – i.e., EEG has roughly a −20 dB signal to noise ratio (SNR; Parra et al., 2008). Thus we have measurement noise that is greater than for the cases seen in **Figures 1A,B**. Secondly, there is the interpretation of the meaning of the EEG-derived variability. We might hypothesize that the variability relates to attentional modulation, workload, perceived error strength, etc., but this too is a hypothesis with its own noise/uncertainty – similar to the issue for the RT derived regressor of **Figure 1B**.

In order to take the uncertainty in the prior into account, we could resample from the prior distribution to construct our null hypothesis (H0). There are many ways to resample or bootstrap (e.g., for a review of techniques in signal processing see (Zoubir and Iskander, 2007). One of the simplest is to draw samples

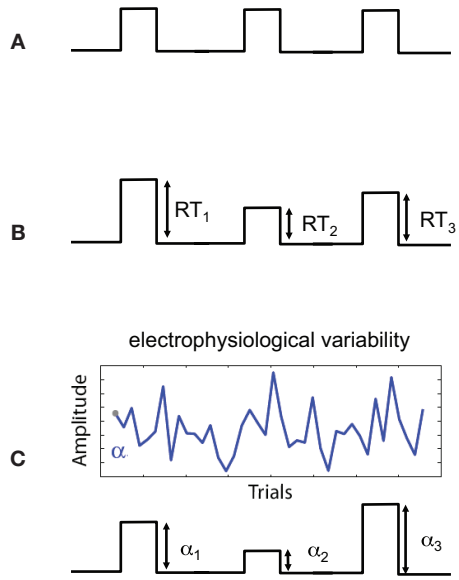


FIGURE 1 | Sources of uncertainty in different regressor models. Shown are three regressor types (before convolution with the hemodynamic response function). **(A)** Regressor constructed from the onset times of the stimulus. The high certainty in this type of regressor is a result of stimulus times being set by the experiment and measured exactly. There is no uncertainty in the interpretation of amplitudes since they are constant. **(B)** Regressor constructed by modulating the stimulus onset times by response times (RTs). Measurement noise of RTs is low (i.e., RTs are high SNR measurements) however there is additional uncertainty (relative to case **A**) in terms of the interpretation of the trial-by-trial variations in response time. **(C)** Regressor constructed by modulating stimulus by activity derived from electrophysiological variability. Sources of uncertainty include both measurement noise, due to the low SNR of the signal, as well as the interpretation of the trial-to-trial electrophysiological variability.

from a distribution of the noise by permuting or randomizing the EEG-derived variability. In previous work we used a simple resampling procedure to compute cluster thresholds for correcting for multiple comparisons (Goldman et al., 2009). However this method for estimating a cluster size threshold was limited because it sacrificed sensitivity for overly conservative specificity. Here we expand this work to a more exhaustive search of the parameter spaces by developing a better representation of the relationship between characteristics of the data. This enables us to maintain both sensitivity and specificity over a range of cluster sizes.

We present these results for two such characteristics: cluster size and maximum *z*-score within a cluster. We use the relationship between them to correct for multiple comparisons by trading-off *z*-score for cluster size. Specifically, we propose a resampling technique, adapted to the noise distribution of the single-trial EEG-derived regressors, that enables increased sensitivity by identifying significance values that are a joint function of cluster size and maximum *z*-score – i.e., enables us to construct the null hypothesis H0 and set a joint threshold to test significance. This approach increases sensitivity by allowing smaller clusters having higher *z*-score, while maintaining specificity.

MATERIALS AND METHODS

The data used in this study are taken from Goldman et al. (2009). Detailed methods of subjects, paradigm, data acquisition, and single-trial analysis are given there (Goldman et al., 2009). In brief, simultaneous EEG and fMRI data were acquired for 11 healthy normal subjects (six female, mean age 31, range 25–38) during an auditory oddball paradigm. Informed consent was obtained from all participants in accordance with the guidelines and approval of the Columbia University Institutional Review Board.

Subjects listened to standard (frequency 350 Hz, probability of occurrence 0.8) and oddball tones (frequency 500 Hz, probability of occurrence 0.2). Subjects were instructed to respond with a button press to the target oddball tone. There were a total of 50 target and 200 standard trials for each subject.

EEG was acquired simultaneously with fMRI using a custom-built MR-compatible system using 36 bipolar twisted pair electrodes sampled at 1 kHz (Goldman et al., 2005; Sajda et al., 2007) in a 1.5-T scanner (Philips Medical Systems, Bothell, WA, USA). Whole brain functional EPI data were acquired with 15 axial slices (TE = 50 ms; TR = 3000 ms; matrix = 64 × 64 voxels, 3.125 mm in-plane resolution, 8 mm thickness). EEG pre-processing included a 0.5-Hz high-pass filter, 60 and 120 Hz notch filters, and gradient artifact removal (mean subtraction as well as 10 ms median filter). BCG artifacts were removed by principal component analysis (PCA) by first estimating the principal components on data after high-pass filtering at 4 Hz and then applying these estimates to the original EEG. The EEG data was epoched into trials in two ways: stimulus-locked (SL, aligned to the onset of the tone) and response-locked (RL, for the target tones, they were aligned to the subject's button press and for the standard tones they were randomly chosen from the RT distribution of the target tones). Individual subject single-trial analysis of EEG was then performed via logistic regression to discriminate between the EEG responses to two classes of stimuli (targets and standards) within consecutive 50 ms training windows. The output of this process is a discriminating component, *y*, specific to that trial and discriminating window, where *y* represents the distance to the discriminating hyperplane (Parra et al., 2002, 2005; Goldman et al., 2009). The amplitudes of the resulting discriminating components within the training window for each SL and RL window were then used to model the BOLD response on a single-trial basis. **Figure 2** illustrates our approach, mapping from trial-to-trial variability in EEG components to fMRI regressors (in this case using time windows of 200 and 350 ms post-stimulus).

RESAMPLING

Our goal in the resampling was to maintain the overall distributions of the EEG discriminating components (*y*-values) for target and standard trials while removing the specific trial-to-trial correlations in the individual experimental runs. To this end, we constructed two empirical distributions (one for "target *y*-values" and one for "standard *y*-values") by pooling the *y*-values for each condition across all subjects and runs. To best demonstrate the contribution of the trial-to-trial variability, we constructed our empirical distributions using *y*-values from only behaviorally correct trials and computed from a time window (450 ms SL) which yielded both substantial discrimination in the EEG (across subject, average

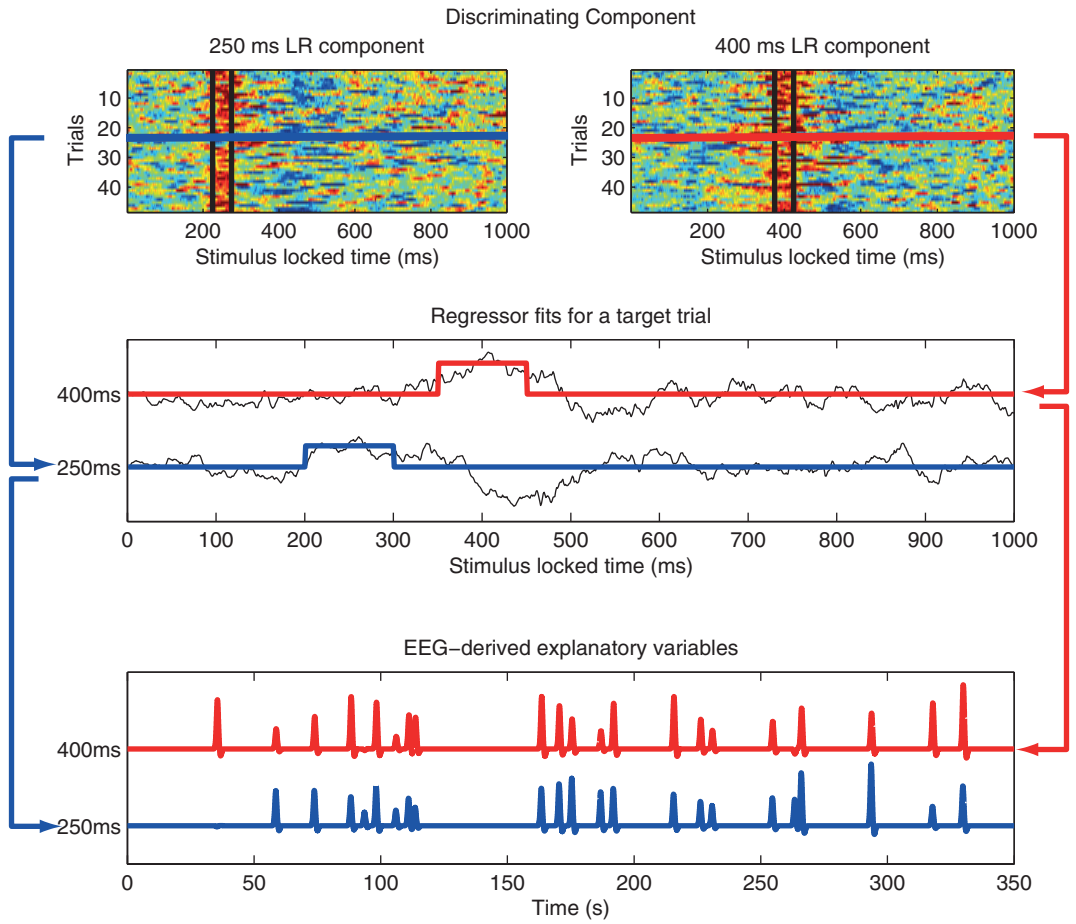


FIGURE 2 | Method used to construct fMRI regressors from EEG component trial-to-trial variability. *Top*: *y*-values for all target trials of the single-trial EEG discriminator for two stimulus-locked windows. Data between black vertical bars indicates those *y*-values used in the analysis. In this example the window width is 50 ms, with one window centered at 200 ms and the other at 350 ms post stimulus-onset. Hot to cold color scale indicates high likelihood to low likelihood for a target. *Middle*: *y*-value for a single target trial for each of the two components (black curves), showing the fMRI event model amplitude as the average of the discriminator output within each 50 ms window, with one modulated event shown for 200 ms (blue) and 350 ms (red). *Bottom*: Single-trial fMRI regressor for target trials across the entire session for the 200 and 350 ms windows, shown after convolution with the hemodynamic response function. Note that the event timing for each of the two windows is the same, but the event amplitudes are different. A separate fMRI analysis is run for each window, using that window's single-trial output to model single-trial variability. Figure from Goldman et al. (2009).

Az > 0.75; Figure 2 in Goldman et al., 2009) as well as significant correlation in the fMRI ($p < 0.005$ and cluster corrected for comparison of <73 voxels; Figure 2 in Goldman et al., 2009).

Given these empirical distributions, we constructed a "resampled run" by taking the ordered vector of targets and standard trials (250 trials per subject, with 50 targets and 200 standards intermixed) and drew randomly, without replacement, from the distribution corresponding to the label of that trial. Thus for a resampled run, all trials were drawn from the distribution with the correct "label" (target or standard) however the specific *y*-value was mixed between trials, runs and subjects. All subjects had the same resampled run *y*-values for a given iteration, though the resulting regressors for each subject were different given that the each had a random sequence of target and standard trials. This entire procedure was repeated 100 times, yielding 100 resamplings of the entire data set (run, subject, group). These procedures allowed us to construct the null hypothesis H0, and establish a joint threshold on cluster size and max *z*-score (see below).

fMRI ANALYSIS

For each of the 100 resampled iterations, a full three-level analysis (run, subject, group) was performed. Standard fMRI pre-processing was performed: slice-timing correction, motion correction, spatial smoothing at 8 mm full-width at half-maximum, and high-pass filtering at 0.01 Hz. Our design matrix included 11 regressors, six of which related to the degrees of motion correction. Of the remaining five regressors (which were convolved with a double-gamma hemodynamic response function), two were a traditional event-related model of target and standard tones (onset at the time of stimulus presentation, duration 100 ms, amplitude of 1), one represented the RTs (with onset at the time of stimulus presentation, duration

100 ms, amplitude corresponding to RT), and the final two regressors modeled the amplitude variability of the single-trial discriminating component (onset time at window time, duration 100 ms, amplitude given by the single-trial EEG discriminator y-value) and were orthogonalized to the traditional target and standard regressors. These single-trial regressors allowed us to examine the BOLD signal related to the variation in the resampled single-trial EEG discriminating component values. Specifically, we looked at the cluster outputs from the resampled single-trial target and standard regressors.

ONE-DIMENSIONAL THRESHOLDS

All clusters that exceeded minimal thresholds [cluster size of 2 and z-score of 2.57 (per voxel $p = 0.005$)] from all 100 iterations were assembled for both positive and negative correlations for the resampled single-trial targets. Two characteristics of the resampled data were obtained: cluster size (number of voxels) and z-score (maximum value in cluster). The resampled data were sorted and thresholds were established based on the top 5% and top 1%. Canonical thresholds were also used: 10 voxel cluster size and 2.57 z-score. The results from the analysis of single-trial simultaneous EEG/fMRI from Goldman et al. (2009) were overlaid to examine those clusters that exceeded each of these thresholds.

CONSTRUCTING JOINT THRESHOLDS

We developed an approach for increasing the sensitivity for smaller cluster sizes by constructing a significance threshold based on the joint distribution of cluster size and maximum z-score. Our approach was to identify a joint distribution, based on these two measures, that showed a strong positive linear correlation which could be used to construct a one-dimensional projection for thresholding which was a function of both dimensions. We found that the log cluster size versus 1 − (maximum p-value in the cluster) resulted in a strong linear fit (Pearson's correlation coefficient $r = 0.88$) for all resampled clusters that exceeded the minimal thresholds [cluster size of 2 and z-score of 2.57 (per voxel $p = 0.005$)]. The projections of the resampled activations along this linear regression were sorted to obtain the maximal 5 and 1%. Significance thresholds ($p = 0.05$ and 0.01) were defined by orthogonal projections from this linear fit of the resampled activations. The results from the analysis of single-trial simultaneous EEG/fMRI from Goldman et al. (2009) were overlaid to examine those clusters that exceeded both of these thresholds.

RESULTS

STATISTICS

There were 655 clusters with a per-voxel $p < 0.005$ that survived the minimum cluster threshold of two voxels in the resampled analysis. The mean cluster size was 14 voxels with a standard deviation of 18. The mean z-score within these clusters was 2.8 and the standard deviation was 0.14. The clusters from the resampled analysis were overlaid onto a standard MNI template brain volume to ensure that clusters of activation from resampling came from regions throughout the brain and did not represent the same cluster (data not shown).

CLUSTER SIZE THRESHOLD

We examined the distribution of cluster sizes (number of voxels) for the resampled data and used these data to establish adaptive thresholds. The largest 1% ($n = 6$ of 655) of the resampled data had a cluster size that exceeded 94 voxels. The largest 5% ($n = 32$ of 655) of the resampled data clusters had a cluster size that exceeded 50 voxels. These thresholds were then applied to the clusters observed in the original data, which represent regions where BOLD signal correlated with either SL or RL single-trial EEG variability. The relationship between the resampled and original data is depicted in **Figure 3**. This figure indicates that there were only three clusters from the original data that exceeded the $p < 0.01$ resampled significance threshold of 94 voxels. Further, these three clusters were the only ones that exceeded the $p < 0.05$ resampled significance threshold of 50 voxels. Two of those activations were seen in the right lateral occipital cortex (**Table 1**, clusters 1 and 2). One cluster was found in the left postcentral gyrus (**Table 1**, cluster 3). Twenty-six percent (22 of 85) of the original clusters exceeded a canonical 10 voxel threshold. Thirty-seven percent (244 of 655) of the resampled clusters exceeded that canonical 10 voxel threshold.

MAX Z-SCORE THRESHOLD

We also plotted the distribution of maximum z-scores (the peak value within the clusters) for the resampled data and used these to establish adaptive thresholds. The upper 1% of the resampled data clusters had a z-score that exceeded 3.17, and the upper 5% had a z-score that exceeded 3.01. These thresholds were then applied to the z-scores from the original data. The relationship between the resampled and original data is depicted in **Figure 4**. This figure indicates one cluster that exceeds the $p < 0.01$ resampled significance threshold of $z = 3.17$. This cluster was located in the right lateral occipital cortex (**Table 1**, cluster 1). The figure also indicates three additional clusters that exceed the $p < 0.05$ resampled significance threshold $z = 3.01$. These clusters were located in the right lateral occipital cortex (**Table 1**, cluster 2), and bilaterally in the amygdala (**Table 1**, clusters 4 and 5).

JOINT THRESHOLDS

The relationship between cluster size (log scale) and 1 − (p-value) for the resampled and original data is depicted in **Figure 5**. The linear regression visualizes the first principal component of the resampled data ($n = 655$). Significance values ($p < 0.05$ and <0.01) were determined by the projections to the regression line of the 32nd and 6th largest resampled data clusters along this first principal component. Significance thresholds, the orthogonal projections from regression line, intersected for $p < 0.05$ at a cluster size of 52 and a 1 − p-value of 0.998 [$\mathbf{x}_{05} = (52,0.998)$] and for $p < 0.01$ at a cluster size of 88 and a 1 − p-value of 0.999 [$\mathbf{x}_{01} = (88,0.999)$]. Two clusters from the original data exceeded $p < 0.01$ significance (clusters 1 and 2) and three additional clusters exceeded $p < 0.05$ significance (clusters 3, 4, and 5).

DISCUSSION

Simultaneously acquired EEG and fMRI data offers the potential to investigate neural states with temporal and spatial precision that is not afforded by either method alone. However, while in

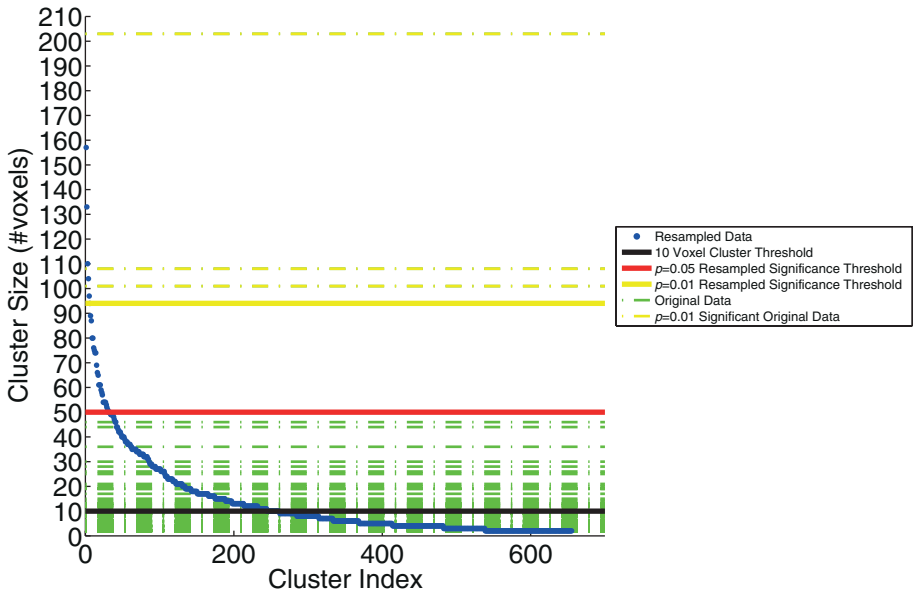


FIGURE 3 | The resampling method, used to generate the threshold based on cluster size. The resampled data (blue) are sorted and plotted in increasing cluster size (number of voxels). Three significance thresholds are presented: 10 voxel canonical cluster threshold (black solid line), $p = 0.05$ resampled significance threshold (red solid line), and $p = 0.01$ resampled significance threshold (yellow solid line). Horizontal dashed lines show the cluster sizes of the original data from Goldman et al. (2009), both those below (green) and the three clusters that exceed (yellow) the resampled cluster size significance threshold ($p < 0.01$).

Table 1 | Significant clusters of activation in the original data of Goldman et al. (2009).

Index	Cluster size	*z*-score	1 – (*p*-value) (joint threshold)	Hemisphere	Location (MNI)			Brain region	EEG single-trial logistic regression results			
					x	*y*	*z*		Correlation	Az	Locked to	Window
1	203	3.42	0.9997	R	40	–74	–8	Lateral occipital cortex	–	0.92	Response	50
2	108	3.04	0.9988	R	42	–68	–4	Lateral occipital cortex	–	0.86	Response	150
3	101	3.01	0.9987	L	–28	–36	60	Postcentral gyrus	–	0.76	Stimulus	450
4	46	3.17	0.9992	R	12	–4	–18	Amygdala	+	0.83	Response	200
5	44	3.14	0.9992	L	–32	2	–18	Amygdala	–	0.79	Response	–100

Index is a number to uniquely identify the cluster for later reference. Cluster size is given in number of voxels. For the one-dimensional thresholds, the background color (yellow, p < 0.01; red, p < 0.05) indicates if the cluster was significant in the corresponding column. For the joint threshold, the background color in the 1 – p-value column indicates the significance. Cluster location is given by hemisphere (R, right; L, left), location of the peak z-score (x, y, z in MNI space), and brain region. Also shown for each cluster is the direction of correlation between single-trial regressor and BOLD signal (–, +), Az value for the single-trial window, whether the window was locked to stimulus onset or response time, and the window onset time in milliseconds.

traditional event-related fMRI the model is known with high certainty, explanatory variables defined by a measured quantity such as EEG introduce more uncertainty into the model. While it is becoming increasingly common to acquire multi-modal data, as yet there is no generally agreed upon method for analysis or interpretation of this data that takes the uncertainty of the model into account.

In this study, we used the data from Goldman et al. (2009), in which the single-trial variability derived from the EEG was used to construct BOLD fMRI regressors, to illustrate a resampling method for determining significance in single-trial EEG/fMRI data. With this method, we can correct for multiple comparisons by adaptively resampling the noise distribution of the EEG-derived regressors. This resampling method demonstrated an interaction between activated cluster size and local maximum *z*-score and showed that the joint distribution of the size of activation clusters and the maximum *z*-score found in the cluster can be used to establish significance thresholds that provide a proper trade-off of sensitivity and specificity. Principled methods for trading-off spatial extent and individual voxel statistical significance were originally proposed for use in neuroimaging by Poline et al. (1997).

Figure 6 graphically summarizes the results from **Table 1** and also illustrates the improved sensitivity we obtained using the joint threshold method. Compared to the results in Goldman et al. (2009), these results show additional significant correlations in areas that overlap with the amygdala. Though this current paper is meant as a statistical methods paper and is not aimed at re-evaluating the specific findings

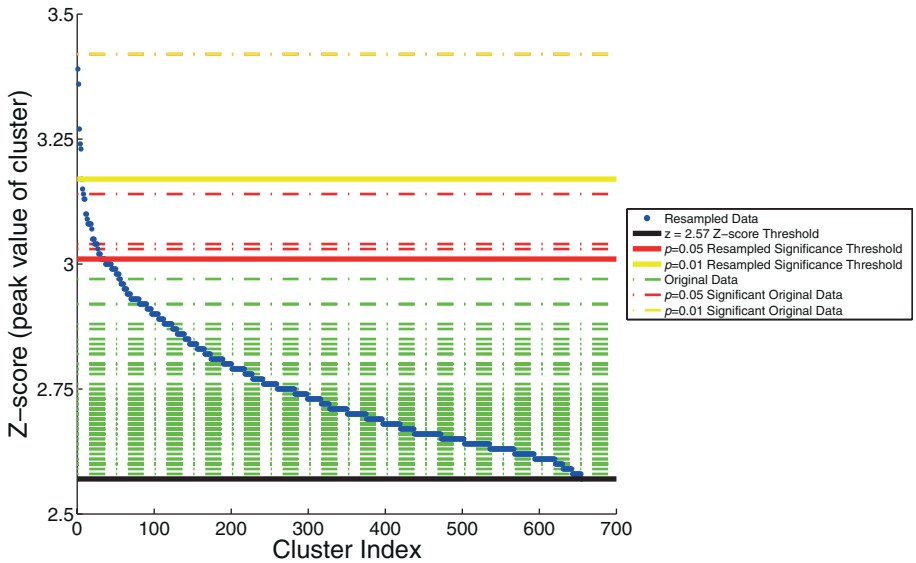


FIGURE 4 |The resampling method, used to generate thresholds based on the maximum *z*-score within the cluster. The resampled data (in blue) are sorted and plotted in increasing *z*-score. Horizontal dashed lines of the *z*-scores of the original data are overlaid (in green). Three significance thresholds are shown: $z = 2.57$ threshold (black), $p = 0.05$ resampled significance threshold (red), and $p = 0.01$ resampled significance threshold (yellow). The dashed lines of the four clusters from the original data of Goldman et al. (2009), that exceed significance thresholds are colored accordingly: three for which $p < 0.05$ (red) and one for which $p < 0.01$ (yellow).

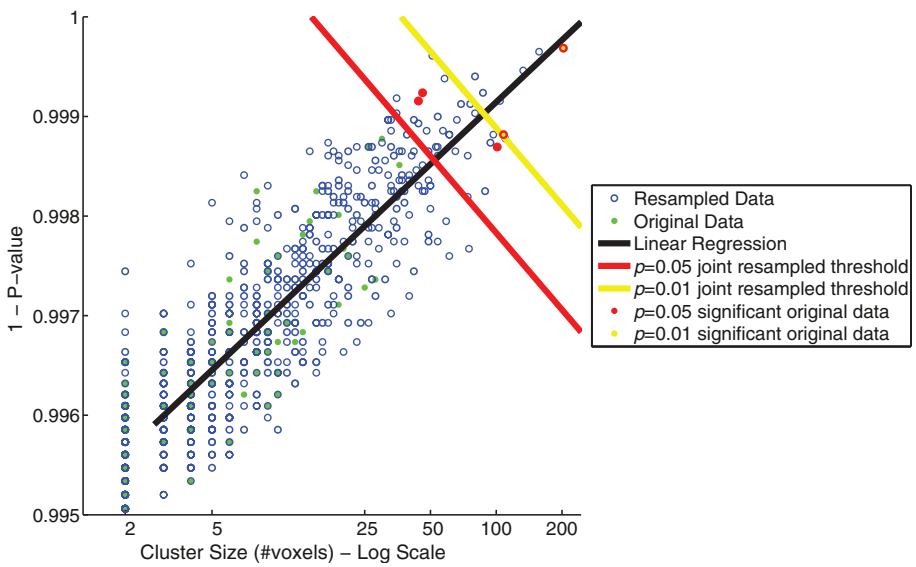


FIGURE 5 | Joint thresholds of cluster size (log scale) versus *p*-value (maximum of cluster) from resampling methods. Shown are scatter plots of cluster size versus *p*-value of the resampled data (blue open circles) and the original data of Goldman et al. (2009; green filled circles). Two joint significance thresholds are shown: $p = 0.05$ resampled significance threshold (red), and $p = 0.01$ resampled significance threshold (yellow). Those clusters from the original data that exceed the joint significance thresholds are colored accordingly.

of Goldman et al. (2009), it is worth noting that the responses in the amygdala have been observed, intracranially, for activity associated with P300 timing and polarity (Halgren et al., 1980).

Projection of the data onto the first principal component of this distribution allows us to simply compute a joint threshold for a one-sided significance level. One might ask "why not use the full distribution instead of the projection on the first principle component?" For example, one might imagine trying to empirically construct a one-sided multivariate test of significance by constructing contours of fixed probability mass about the mean and then utilize the contour cutoff so that an increase in single-voxel significance always yields a decrease in voxel size, and vice versa. Though such a procedure seems to be even more powerful than simply looking at the one-dimensional projection, it comes with the following disadvantages/costs. First it

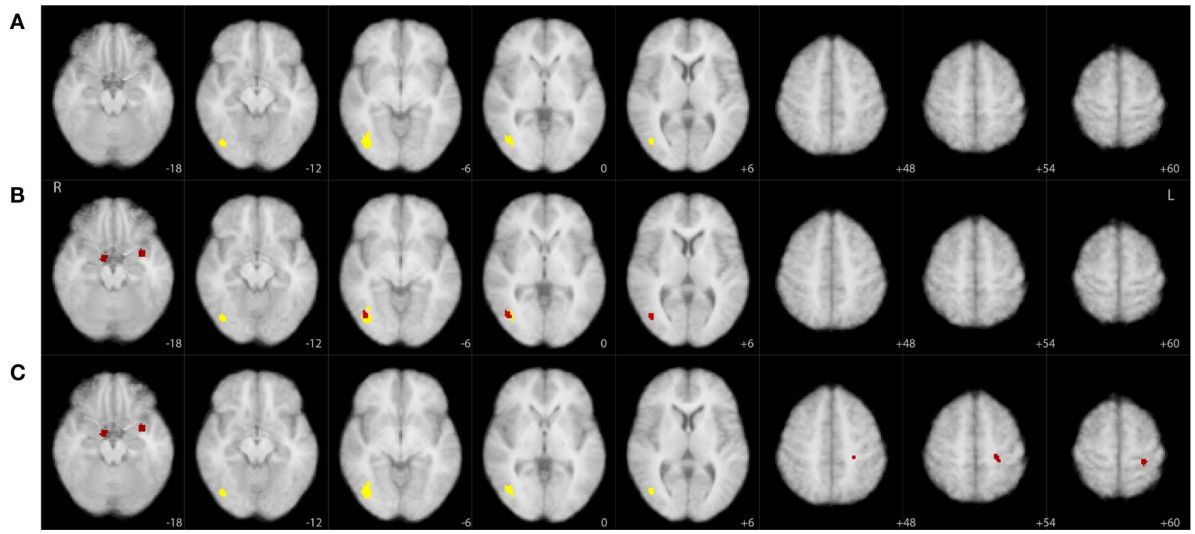


FIGURE 6 | Significant clusters (yellow, $p < 0.01$; red, $p < 0.05$) overlaid on axial slices of an MNI template brain image. Each row corresponds to a different thresholding technique derived from our resampling tests: **(A)** cluster size only based threshold, **(B)** maximum *z*-score based threshold, **(C)** joint threshold. Sampled slices not shown (slices sampled every 6 mm) had no significant clusters for any of the thresholding techniques. See **Table 1** for additional information on these clusters.

requires a parametric representation of the joint distribution so as to determine the shape of the contours and secondly it requires substantially more resampling to identify the contour/boundary and/or fit the parametric distributions if modeling the data. Since the distribution in **Figure 5** does not appear to be well-modeled by a multivariate normal, we chose to illustrate our procedure with the simplest method which maintains the key points and findings of the approach.

Our focus has been on a data adaptive approach to correct for multiple comparisons in the joint space of single-voxel *p*-value and cluster size. Cluster size thresholding has been a popular method for correcting for multiple comparisons in EEG/fMRI, but often these thresholds are selected *ad hoc* and then propagate in the literature, used by others with little statistical justification (e.g., see Scheibe et al., 2010). There are of course other data adaptive approaches that can be used to correct for multiple comparisons and are not based on using cluster-size. One of the more recently adopted techniques controls the false-discovery rate (FDR; Genovese et al., 2002; Nichols and Hayasaka, 2003), namely the expected ratio of false positives to true positives. The FDR is data adaptive in that the thresholds are inherently linked to the signal to noise in the data. FDR correction methods are easily interpretable and are computationally efficient, compared to the resampling procedure we describe here. However, FDR usually does not take into account any prior model or certainty about the regressor model and requires some assumptions about the structure of the noise distribution. FDR correction methods are none-the-less an attractive approach for multiple comparison correction and more work is needed to understand how they can best be employed in EEG/fMRI data analysis.

CONCLUSION

In summary, our results suggest that additional care is required when using cluster-size to correct for multiple comparisons and determine significance in EEG/fMRI data. Some groups use conservative thresholds (Goldman et al., 2009) while others use cutoffs that are more lenient (Debener et al., 2005; Benar et al., 2007; Mayhew et al., 2010; Scheibe et al., 2010). Conservative cluster thresholds reduce false positives, but they also limit the sensitivity of single-trial analysis. The resampling method proposed here suggests that standard *ad hoc* cluster size thresholds, for example of 10 voxels, are too lenient (at least in this case) since more than a third of the resampled data clusters were found to have a larger size. For this data, a more conservative threshold of 50 voxels, as determined by the variability in the data itself, is a better estimate of the true data significance. This is less conservative than the threshold (of 73) presented in Goldman et al. (2009) that was derived from a less expansive bootstrapping method. In any case, our results suggest that establishing statistical significance using a threshold criteria that has been used in another EEG/fMRI study may not be adequate. Instead these corrections are better off left to be data adaptive, and though this can be computationally expensive, it improves interpretation of the results and properly establishes statistical significance.

ACKNOWLEDGMENT

This work was supported by National Institutes of Health grants R01-MH085092 and R33-EB004730.

REFERENCES

Benar, C. G., Schon, D., Grimault, S., Nazarian, B., Burle, B., Roth, M., Badier, J. M., Marquis, P., Liegeois-Chauvel, C., and Anton, J. L. (2007). Single-trial analysis of oddball event-related potentials in simultaneous EEG–fMRI. *Hum. Brain Mapp.* 28, 602–613.

Debener, S., Ullsperger, M., Siegel, M., Fiehler, K., Von Cramon, D. Y., and Engel, A. K. (2005). Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. *J. Neurosci.* 25, 11730–11737.

Eichele, T., Specht, K., Moosmann, M., Jongsma, M. L., Quiroga, R. Q., Nordby, H., and Hugdahl, K. (2005). Assessing the spatiotemporal evolution of neuronal activation with single-trial event-related potentials

and functional MRI. *Proc. Natl. Acad. Sci. U.S.A.* 102, 17798–17803.

Esposito, F., Mulert, C., and Goebel, R. (2009). Combined distributed source and single-trial EEG–fMRI modeling: application to effortful decision making processes. *Neuroimage* 47, 112–121.

Genovese, C. R., Lazar, N. A., and Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage* 15, 870–878.

Goldman, R. I., Gerson, A. D., Cohen, M. S., Brown, T. R., and Sajda, P. (2005). "Simultaneous EEG and fMRI for event-related studies," in *Eleventh Annual Meeting of the Organization for Human Brain Mapping* (Toronto, ON: Elsevier).

Goldman, R. I., Wei, C. Y., Philiastides, M. G., Gerson, A. D., Friedman, D., Brown, T. R., and Sajda, P. (2009). Single-trial discrimination for integrating simultaneous EEG and fMRI: identifying cortical areas contributing to trial-to-trial variability in the auditory oddball task. *Neuroimage* 47, 136–147.

Halgren, E., Squires, N. K., Wilson, C. L., Rohrbaugh, J. W., Babb, T. L., and Crandall, P. H. (1980). Endogenous potentials generated in the human hippocampal formation and amygdala by infrequent events. *Science* 210, 803–805.

Mayhew, S. D., Dirckx, S. G., Niazy, R. K., Iannetti, G. D., and Wise, R. G. (2010). EEG signatures of auditory activity correlate with simultaneously recorded fMRI responses in humans. *Neuroimage* 49, 849–864.

Mulert, C., Leicht, G., Hepp, P., Kirsch, V., Karch, S., Pogarell, O., Reiser, M., Hegerl, U., Jager, L., Moller, H. J., and Mccarley, R. W. (2010). Single-trial coupling of the gamma-band response and the corresponding BOLD signal. *Neuroimage* 49, 2238–2247.

Nichols, T., and Hayasaka, S. (2003). Controlling the familywise error rate in functional neuroimaging: a comparative review. *Stat. Methods Med. Res.* 12, 419–446.

Novitskiy, N., Ramautar, J. R., Vanderperren, K., De Vos, M., Mennes, M., Mijovic, B., Vanrumste, B., Stiers, P., Van Den Bergh, B., Lagae, L., Sunaert, S., Van Huffel, S., and Wagemans, J. (2011). The BOLD correlates of the visual P1 and N1 in single-trial analysis of simultaneous EEG–fMRI recordings during a spatial detection task. *Neuroimage* 54, 824–835.

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

Parra, L. C., Christoforou, C., Gerson, A. D., Dyrholm, M., An, L., Wagner, M., Philiastides, M. G., and Sajda, P. (2008). Spatiotemporal linear decoding of brain state. *IEEE Signal Process. Mag.* 25, 107–115.

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

Poline, J. B., Worsley, K. J., Evans, A. C., and Friston, K. J. (1997). Combining spatial extent and peak intensity to test for activations in functional imaging. *Neuroimage* 5, 83–96.

Sajda, P., Goldman, R. I., Philiastides, M. G., Gerson, A. D., and Brown, T. R. (2007). "A system for single-trial analysis of simultaneously acquired EEG and fMRI," in *Third International IEEE EMBS Conference on Neural Engineering*, ed. M. Akay (Kohala Coast, HI: IEEE Press), 287–290.

Scheibe, C., Ullsperger, M., Sommer, W., and Heekeren, H. R. (2010). Effects of parametrical and trial-to-trial variation in prior probability processing revealed by simultaneous electroencephalogram/functional magnetic resonance imaging. *J. Neurosci.* 30, 16709–16717.

Worsley, K. J., and Friston, K. J. (1995). Analysis of fMRI time-series revisited – again. *Neuroimage* 2, 173–181.

Zoubir, A. M., and Iskander, D. R. (2007). Bootstrap methods in signal processing [from the guest editors]. *IEEE Signal Process. Mag.* 24, 7–8.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 11 March 2011; accepted: 27 April 2011; published online: 20 May 2011.
Citation: deBettencourt M, Goldman R, Brown T and Sajda P (2011) Adaptive thresholding for improving sensitivity in single-trial simultaneous EEG/fMRI. Front. Psychology ***2****:91. doi: 10.3389/fpsyg.2011.00091*
This article was submitted to Frontiers in Perception Science, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 11 March 2011
doi: 10.3389/fpsyg.2011.00042

Real-time measurement of face recognition in rapid serial visual presentation

Jon Touryan*, Laurie Gibson, James H. Horne and Paul Weber

Boulder Center for Neurotechnology, Science Applications International Corporation, Louisville, CO, USA

Edited by:
Paul Sajda, Columbia University, USA

Reviewed by:
Fang Fang, Peking University, China
Jesse Husk, McGill University, Canada

****Correspondence:***
Jon Touryan, Boulder Center for Neurotechnology, Science Applications International Corporation, 801 Main Street, Suite 300, Louisville, CO 80027, USA.
e-mail: jonathan.o.touryan@saic.com

Event-related potentials (ERPs) have been used extensively to study the processes involved in recognition memory. In particular, the early familiarity component of recognition has been linked to the FN400 (mid-frontal negative deflection between 300 and 500 ms), whereas the recollection component has been linked to a later positive deflection over the parietal cortex (500–800 ms). In this study, we measured the ERPs elicited by faces with varying degrees of familiarity. Participants viewed a continuous sequence of faces with either low (novel faces), medium (celebrity faces), or high (faces of friends and family) familiarity while performing a separate face-identification task. We found that the level of familiarity was significantly correlated with the magnitude of both the early and late recognition components. Additionally, by using a single-trial classification technique, applied to the entire evoked response, we were able to distinguish between familiar and unfamiliar faces with a high degree of accuracy. The classification of high versus low familiarly resulted in areas under the curve of up to 0.99 for some participants. Interestingly, our classifier model (a linear discriminant function) was developed using a completely separate object categorization task on a different population of participants.

Keywords: event-related potential, recognition memory, object categorization, classifier

INTRODUCTION

The neural substrates of recognition, an essential aspect of declarative memory, have been extensively studied with the use of event-related potentials (ERPs). Many of these studies attempt to dissect the two phenomenologically distinct processes involved in recognition: familiarity and recollection (Yonelinas, 2002; Yovel and Paller, 2004; Guo et al., 2005; Curran and Hancock, 2007; MacKenzie and Donaldson, 2007). A commonly accepted definition of familiarity is the sense of having previous experience with the probe stimulus (e.g., person, object, word) without any accompanying contextual information as to the nature of the previous encounter. Recollection, on the other hand, is when the memory of the probe stimulus is accompanied by contextual or associative detail. The neural correlates of familiarity have been linked to the early mid-frontal negativity (FN400) in the ERP. Specifically, the amplitude of the negative deflection between 300 and 500 ms is less for familiar as compared to novel stimuli (Rugg et al., 1998; Curran, 2000). Recollection, by contrast, has been linked to a later positive component over the central–parietal cortex. This positive deflection, between 500 and 800 ms, is greater for stimuli that have been consciously recollected (Smith, 1993). However, there is still an ongoing debate as to whether or not these ERP components are a reflection of two distinct neural processes (Yonelinas, 2002; Paller et al., 2007). In addition, the influences of conceptual priming (Voss and Paller, 2006; Voss et al., 2008, 2010) make some prior research on this question difficult to interpret.

For the purpose of this study we used the early (300–500 ms) and late (500–800 ms) windows, associated with familiarity and recollection respectively, as a means to quantity the gradations in the recognition response. Specifically, how does the magnitude of the recognition response change with the level of experience? And, does this change affect both early and late components in a similar fashion? Here, to manipulate level of experience with the stimuli, we used color photographs of faces from three distinct categories: novel individuals, famous persons or celebrities selected by the participant, and personal friends and family provided by the participant. In this way, we were able to quantify the effect of experience or knowledge of the depicted individual on the recognition response. While this study does not explicitly dissociate the neural processes involved in familiarity and recollection, the magnitude of the evoked responses (within the early and late integration windows) do have implications for the competing models of recognition.

In addition to this conventional analysis of the ERP, we also wanted to determine how accurately the recognition response could be classified on a single-trial basis. To accomplish this, we integrated the face stimuli into a real-time system that classifies the evoked response elicited by each stimulus, based on a linear model of the neural response pattern. The motivation for this approach stems from a potential application of the recognition response toward a novel brain–machine interface (BMI). BMI technologies often utilize the visual categorization response for binary output or classification (Parra et al., 2002; Sellers et al., 2006). An example of this is a rapid presentation of images or letters for which each P300 is classified as a "yes" or "no" response. This response can then be used to identify relevant objects in a stream of imagery or to select letters for building a word or phrase. Here, we sought to use a similar experimental paradigm to further quantify the recognition response. Specifically, can the recognition response be identified in a continuous presentation of faces where explicit recognition of each face is not required?

This unorthodox approach to measuring the recognition response utilized the rapid serial visual presentation (RSVP) paradigm (Chun and Potter, 1995). In this paradigm the participant viewed a continuous stream of rapidly presented faces for several minutes. The task was simply to identify, via button press, a small number of target faces from amongst the ensemble of images. There was no explicit study phase beyond the imagery that was provided by the participant. Likewise, there was no explicit test phase where the participants indicated which faces they recognized. Despite the less controlled nature of this study (relative to the prior research) we were able to clearly identify the recognition response both at an aggregate and trial-by-trial level.

MATERIALS AND METHODS

PARTICIPANTS

Twenty-two individuals participated in the experiment for payment of $20 per hour. The participants (17 female and 5 male) ranged in age from 22 to 53, with a mean age of 28. Participants were both right-handed and left-handed (19 right-handed, 3 left-handed). Five of the 22 individuals participated in a second, experimentally identical session roughly a week after the initial session. None of the participants were excluded from the analysis due to noise (bad channels), movement artifacts, or low behavioral accuracy. However, the real-time classification system did exclude individual trials based on pre-defined noise and movement artifact thresholds.

STIMULI

Stimuli consisted of 256 × 320 pixel color photographs of single faces that were manually centered, scaled, and cropped. The eyes were centered just above the midline and the entire face was contained within the cropped region. Photographs were excluded from the experiment if the face was obscured by sunglasses, hats, or costumes. Novel faces, both male and female, were obtained from the Flickr Creative Commons database[1]. All 447 novel faces were used in every experiment. Famous faces of movie stars, singers, or celebrities were obtained from Getty Images[2]. There were 79 famous females and 80 famous males in total. Participants were allowed to select the 20–30 most familiar individuals, roughly balancing for gender. The average number of faces selected by the participant was 24.6 ± 4.5 (std). Personal faces were obtained from each participant prior to the experimental session. On average, 15.7 ± 4.1 faces provided by the participant met the above criteria and were included in the experiment. The target faces were the three most recent presidents (Obama, Bush, and Clinton), which were shown to the participant prior to the experiment.

PROCEDURE

The participants were seated in front of a computer monitor at a distance of approximately 100 cm. All images were displayed at the center of the monitor and subtended a visual angle of about 7° horizontally and 9° vertically. The experiment consisted of 10–14 blocks, each containing roughly 200 faces presented in rapid sequence (i.e., the RSVP paradigm). Each face was presented in random order for 500 ms and there were no breaks between faces. The variation in block number was due to the dynamics of the real-time system. If the response corresponding to a particular face presentation contained EEG artifacts (blinks, eye movement, etc.) that face was then re-queued for subsequent presentation. The experiment was complete when all faces were shown to the participant, artifact free, at least four times. Participants were instructed to fixate at the center of the monitor and respond, via button press, when they saw a target face (Presidents Obama, Clinton, or Bush). There was a pause at the end of each block and the participant started the next block at their discretion. Thus, the participants experienced approximately 2 min of RSVP, followed by a self-paced rest period.

EEG RECORDING

Scalp EEG was collected with a 128-channel HydroCel Geodesic Sensor Net™ (Electrical Geodesics, Inc., Eugene, OR, USA) connected to an AC-coupled 128-channel, high-input impedance amplifier (200 MΩ, Net Amps™, Electrical Geodesics, Inc.). Individual sensors were adjusted until impedances were less than 50 kΩ. Amplified analog voltages (0.1–100 Hz bandpass) were digitized at 250 Hz and then low-pass filtered at 40 Hz. Recorded voltages were initially referenced to a vertex channel. Trials were excluded from analysis if they contained eye movements (vertical electro-oculogram channel differences greater than 70 μV) or more than five bad channels (changing more than 100 μV between samples, or reaching amplitudes over 200 μV). Data from individual bad channels were replaced using a spherical-spline interpolation algorithm. An average-referenced transform was then used for subsequent ERP analysis. All ERPs were baseline-corrected to a 100 ms pre-stimulus recoding interval.

REAL-TIME CLASSIFICATION

The real-time classification of the EEG signal was accomplished through analysis of the evoked response following the presentation of each face. The goal was to find a linear combination of the components of the signal that most reliably discriminated between the responses to familiar or recognized versus novel faces. Classification was initially performed using a standard or general model. This linear model was developed from the responses of multiple participants in a prior study (Curran et al., 2009; Touryan et al., 2010) examining the P300 object classification signal (Thorpe et al., 1996). Individualized or custom models were also developed for each participant using only that participants data. Below we describe the feature selection process (Perkins et al., 2003) we implemented to develop the linear models (both general and custom) used in the real-time EEG classification.

For each trial, N features are generated, giving a feature vector $\mathbf{x} \in \mathbf{R}^N$. The basic machine learning approach is to find a discriminant function $f: \mathbf{R}^N \mapsto \mathbf{R}$ that maps the features $\mathbf{x}$ into the probability that the trial was caused by one of two stimulus classes. For recognition, we define the classes as familiar and unfamiliar. Here, the probability of recognition, given the data vector $\mathbf{x}$ and a discriminant function $f(\mathbf{x})$, is

$$p(\text{familiar}\,|\,\mathbf{x}) = \frac{1}{1+e^{-f(\mathbf{x})}}. \quad (1)$$

The probability of non-recognition is

[1]http://www.flickr.com/creativecommons

[2]http://www.gettyimages.com

$$p(\text{unfamiliar}|\mathbf{x}) = 1 - p(\text{familiar}|\mathbf{x}) = \frac{1}{1+e^{+f(\mathbf{x})}}. \quad (2)$$

Thus if $f(\mathbf{x})$ is large and positive, the probability of recognition is near one, and if $f(\mathbf{x})$ is large and negative, the probability of recognition is near zero. The midpoint $f(\mathbf{x}) = 0$ corresponds to a probability of recognition of 1/2 and is used as the dividing line between the classes.

We use the following linear discriminant function to distinguishing between recognized and unrecognized trials:

$$f(\mathbf{x}) = w_0 + \sum_{i=1}^{N} w_i x_i, \quad (3)$$

where N is the number of features and w_i is the weight of the ith feature. This collection of weights represents the linear model that separates the two classes by means of a hyperplane. The goal then is to find the collection of weights that maximally separates the two classes. To accomplish this, a dataset is built from the initial experimental session. These sessions include repeated presentations of all faces (novel, famous, and personal) in random order. The session data is decomposed into a set of feature vectors: $\mathbf{x}_a^r, a = 1,\ldots,N_r$ for the familiar or recognized faces, and $\mathbf{x}_a^u, a = 1,\ldots,N_u$ for the unfamiliar faces. The session is then randomly and iteratively split into separate training and validation portions. The training data is used to find a set of weights which are then applied to the feature vectors in the validation data. The linear model is then applied to any subsequent experimental session for that participant.

An important factor to consider with dense array EEG (128 channels at 250 Hz) is the large number of features in each evoked response. Since every channel contains some noise, it is easy to find a set of weights that perfectly fits the training data, but gives poor results on the validation data or subsequent sessions. Thus, a model is more robust if the majority of weights w_i are zero, so that only the most important (and stable) features are used. The approach described by Perkins et al. (2003) is to penalize any non-zero w_i. Specifically, the weights are found by minimizing the regularized empirical risk function

$$L(w_k, \mathbf{x}_a^u, \mathbf{x}_a^r) = \frac{1}{N_u + N_r}\left(-\sum_{a=1}^{N_u} \log \frac{1}{1+e^{f(\mathbf{x}_a^u)}} - \sum_{a=1}^{N_r} \log \frac{1}{1+e^{-f(\mathbf{x}_a^r)}}\right) + \lambda_1 \sum_{i=1}^{N} |w_i| + \lambda_0 \sum_{w_i \neq 0} 1, \quad (4)$$

where λ_1 and λ_0 are regularization constants. The first two terms in the empirical risk function are the logarithms of the probability that the training set is correctly classified. The λ_1 term favors features where small values of w_i distinguish well between the classes. The λ_0 term simply counts the number of non-zero weights.

Given a fixed set of non-zero weights, $L(w_k)$ can be minimized using standard function minimization, while ensuring that any weight that was initially zero remains at zero. A concurrent greedy strategy decides which of these zero weights (if any) should be allowed to vary by choosing the weight with the largest value of $|\partial L/\partial w_i|$. This weight is then added to the set of non-zero weights, and $L(w_k)$ is again minimized. The algorithm begins by allowing only w_0 to vary, then iteratively finding the zero weight with the largest $|\partial L/\partial w_i|$ to add to the empirical risk function. This process is repeated until no more weights can be added. The purpose of the λ_0 term (for this study $\lambda_0 = 10^{-6}$) is to force any w_i that is very near zero to be precisely zero. Conveniently, $L(w_k)$ is a convex function (ignoring the λ_0 term), so the only minimum is a global minimum.

CLASSIFIER FEATURES

For each evoked response, a large number of features are generated by first linearly transforming the raw EEG signal using principal components analysis (PCA), and then calculating windowed fast Fourier transforms (FFT) for a variety of window sizes and starting times relative to the stimulus onset. We typically include window sizes of w = 128, 256, and 512 ms in the feature set. For memory capacity reasons, features corresponding to frequencies greater than 25 Hz are discarded. All other non-zero frequency components are split into real and imaginary parts. For each epoch, the features for all times, window sizes and frequencies are assembled into a large feature vector $\mathbf{x} \in \mathbf{R}^N$.

RESULTS

The experimental task used in this study (**Figure 1**) is a simple extension of the traditional RSVP target-detection paradigm (Chun and Potter, 1995). Here, participants had little difficulty detecting the target faces (Presidents Obama, Bush, and Clinton) from amongst the other faces, both familiar, and novel. However, due to the presentation speed (2 Hz) and block duration (2 min), participants reported difficulty in responding to a target face in time (i.e., before the next face appeared). This is likely the primary reason that behavioral performance was not at ceiling. Over the population of

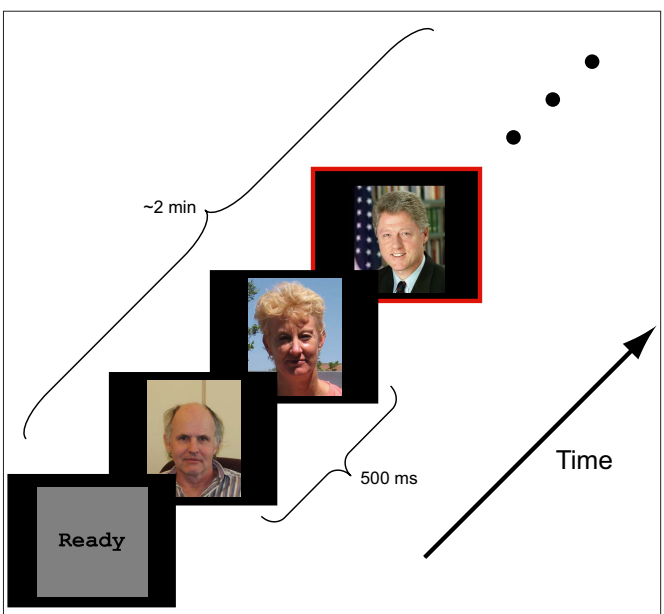


FIGURE 1 | Schematic representation of experimental design. Color photographs of faces were presented in a rapid serial sequence (500 ms per face). Participants were asked to respond, via button press, only when a target face (Presidents Obama, Bush, and Clinton) appeared. Besides the target face, the stimulus ensemble included faces of novel individuals, famous persons selected by the participant and personal friends and family provided by the participant.

22 participants, the average accuracy for the behavioral response was 81% with high and low scores of 92 and 60.4%, respectively. The average reaction time was 598 ± 53 ms (std).

ERP RESULTS

Initial analysis focused on spatial regions of interest (ROIs) that were used in previous studies (Yovel and Paller, 2004; Curran and Hancock, 2007). In particular, we analyzed the same channel groups as Curran and Hancock (2007). The two anterior, superior channel groups located near the standard F_3 and F_4 sites were labeled LAS and RAS (see **Figure 2** montage overlays). The two posterior–superior channel groups, which included the standard P_3 and P_4 sites, were labeled LPS and RPS. Both the familiarity response, or FN400, and recollection response we analyzed over all channel groups. Here the early familiarity response was analyzed from 300 to 500 ms, whereas the recollection response was analyzed from 500 to 800 ms. The increased time window (300 ms) for the recollection response was to compensate for the initial visual response elicited from the following stimuli (presented 500 ms after stimulus onset). **Figure 2** shows the average ERPs for each channel group. In agreement with previous studies (Curran and Hancock, 2007; MacKenzie and Donaldson, 2007), the anterior groups (LAS and RAS) showed a greater difference in the early familiarly response as a function of stimulus class (novel, famous, personal) relative to the posterior groups (LPS and RPS). For the later recollection response, all channel groups showed a clear differentiation with stimulus class. Average amplitudes for each channel group are detailed in **Table 1**.

We compared the mean amplitudes of the frontal LAS and RAS channel groups during the 300–500 ms window. A repeated-measures analysis of variance (ANOVA) with condition (novel, famous, personal) *hemisphere indicated a main effect of stimulus type [$F(2,42) = 46.23$, $p < 0.001$]. A within-subjects contrast confirmed a significant effect between novel and famous faces [$F(1,21) = 23.57$, $p < 0.001$] as well as between personal and novel/famous faces [$F(1,21) = 58.13$, $p < 0.001$]. There was no significant hemispheric difference in the LAS and RAS mean amplitudes [$F(1,21) = 3.67$, $p = 0.07$]. Likewise, there was no significant interaction effect between condition and hemisphere [$F(2,42) = 2.33$, $p = 0.11$]. This result confirms previous research describing the effect of face familiarity on ERP amplitudes over the frontal cortex during the 300–500 ms time window (Curran and Hancock, 2007).

One important difference with our results is the exaggerated effect due to level of familiarity with the particular individual in each stimulus (i.e., faces of friends and family elicited the largest response). To quantify this we calculated the Pearson's correlation coefficient between mean amplitude and stimulus type. Specifically, we calculated the correlation coefficient between stimulus type, sorted from least to most familiar (i.e., novel–famous–personal), and mean amplitude across all participants and each channel group: $r = 0.445$, $p < 0.001$. This indicates that a substantial part of the variance in the mean response is due to stimulus type, or level of familiarity with the stimulus, in addition to individual differences and scalp topography. Typically, familiarity with the stimulus set is carefully controlled during a study phase where participants associate a novel face with some attribute (Yovel and Paller, 2004; Curran and Hancock, 2007; MacKenzie and Donaldson, 2007). This process constrains familiarity with the stimulus to that specific laboratory setting, thereby limiting the strength of subsequent recognition. By using faces of individuals previously known to the participant, we have maximized the amplitude effect and validated the phenomenon in a more realistic context.

The mean amplitudes of the parietal LPS and RPS channel groups were compared over the later 500–800 ms window. Again, a repeated-measure ANOVA with condition (novel, famous, personal) *hemisphere indicated a main effect of stimulus type [$F(2,42) = 50.68$, $p < 0.001$]. Within-subjects contrasts were also significant at all levels [novel versus famous $F(1,21) = 24.37$, $p < 0.001$; personal versus

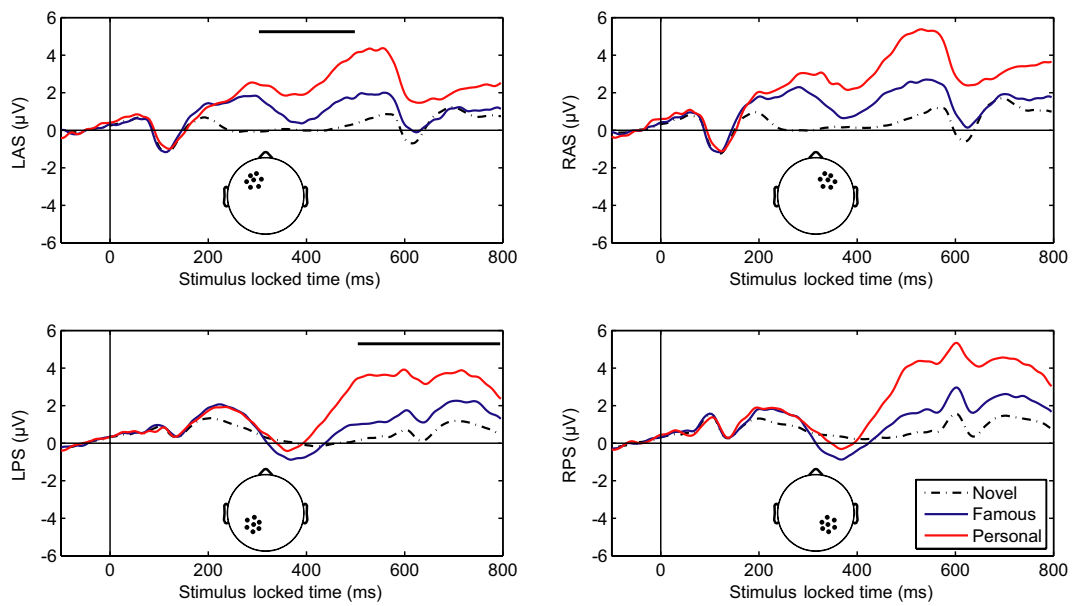


FIGURE 2 | Averaged-referenced, grand-averaged ERPs for depicted channel groups. L, left; R, right; A, anterior; P, posterior; S, superior. Thick overlaid bars in the left column indicate the early (upper) and late (lower) temporal integration widows.

novel and famous $F(1,21) = 58.48$, $p < 0.001$]. The parietal channel groups did show a significant hemispheric difference [$F(1,21) = 5.38$, $p < 0.05$] with a larger mean amplitude over the right (RPS) channel group. However, there was no significant interaction effect between condition and hemisphere [$F(2,42) = 2.34$, $p = 0.11$]. Again, this result confirms and extends the prior work on the recollection response. As with the earlier FN400, the magnitude of the later parietal response is significantly correlated with level of familiarity ($r = 0.554$, $p < 0.001$; Pearson's correlation). Importantly, this parietal recognition response was preserved even with the addition of the early visual response elicited from the following stimulus.

To compare the broad topography between the early and later windows, we again computed the mean amplitudes within the four channels groups shown in **Figure 2**. In this instance however, we used vector normalization (McCarthy and Wood, 1985) and only considered the ERPs associated with the personal stimulus category. Specifically, the average amplitudes for each participant were calculated for the early and late windows described above. These amplitude distributions were normalized, creating 128 dimensional vectors with unit length. The montage averages were then re-calculated using these normalized distributions. Repeated-measures ANOVA indicated that the early and late scalp distributions were significantly different. Both the time (300–500, 500–800 ms) *hemisphere interaction, $F(1,21) = 12.85$, $p = 0.002$, and time *anterior/posterior interaction, $F(1,21) = 53.78$, $p < 0.001$, show a significant effect. This result is consistent with separate-source theory of familiarity and recollection (Yonelinas, 2002; Curran and Hancock, 2007).

LINEAR CLASSIFIER

In addition to the post-experiment ERP analysis described above, the participant's neural response was scored via a classification algorithm, in real-time, after each stimulus presentation (see Materials and Methods for details). In our case, the classifier consisted of a linear discriminant function that was applied to the evoked response to generate a score for each face. **Figure 3** shows the grand-average topographic ERP for personal faces and the general classifier model overlaid on the same ERP. Essentially, the model is a set of linear weights applied to each channel at each time point in the evoked response. The weighted evoked response is then summed to generate the score. Here, the general model (used in every participant's initial session) primarily isolates regions over the parietal cortex during the later epoch of the evoked response (500–600 ms). These spatial–temporal regions align well with the maximal recognition response shown in **Figure 2**. Interestingly, the general model was built from a prior target-detection experiment (Curran et al., 2009; Touryan et al., 2010). In that experiment participants were asked to detect targets (people and vehicles) in an RSVP sequence of natural images. A machine learning algorithm (Perkins et al., 2003) was then used to identify features in the evoked response that maximally separated target from background images. Data from all participants, both right- and left-handed, was used to build this general classifier model, which captured the relevant components of the P300 complex.

We quantified the performance of the general classifier model in the face recognition task with receiver operating characteristics (ROC) analysis (Green and Swets, 1966). For each participant, we calculated the areas under the curve (AUC) for two conditions. First, famous and personal faces were combined and considered as familiar and compared against novel faces. Second, only personal faces were compared against novel faces. **Figure 4** shows the performance of each participant using these two calculations. The classifier performed well for the majority of participants with an average AUC for the familiar versus novel of 0.827 (min. = 0.583, max. = 0.957). The comparison between personal and novel was even more compelling with and average AUC of 0.858 (min. = 0.580, max. = 0.998). This difference was statistically significant ($p < 0.05$, paired t-test).

Overall, the classifier performance was not correlated with behavioral accuracy ($r = -0.17$, $p = 0.46$; Pearson's correlation), rather it reflected the individual differences in the underlying response. **Figure 5** illustrates this diversity in the recognition response over the right-parietal electrodes (RPS channel group). This area in particular, is critical to the general classifier model (see **Figure 3**). Not surprisingly, the performance of the general model was good for the first participant (participant 121, AUC = 0.91) but not for the second (participant 107, AUC = 0.69). It is clear from the average ERPs that the first participant had a large recognition response by 500 ms, whereas the second participant's recognition response was substantially delayed. The general classifier model, built from multiple participants, is heavily weighted toward the earlier response

Table 1 | Amplitude means and SE.

Channel group	Latency (ms)	Novel	Famous	Personal
LAS	300–500	0.0 (0.0)	1.1 (0.1)	2.5 (0.1)
	500–800	0.6 (0.1)	1.1 (0.1)	2.6 (0.2)
RAS	300–500	0.2 (0.0)	1.4 (0.1)	3.1 (0.2)
	500–800	0.9 (0.1)	1.7 (0.2)	3.6 (0.2)
LPS	300–500	0.0 (0.0)	–0.1 (0.1)	0.9 (0.2)
	500–800	0.6 (0.1)	1.6 (0.1)	3.5 (0.1)
RPS	300–500	0.3 (0.0)	0.1 (0.2)	1.1 (0.3)
	500–800	1.0 (0.1)	2.1 (0.1)	4.3 (0.1)

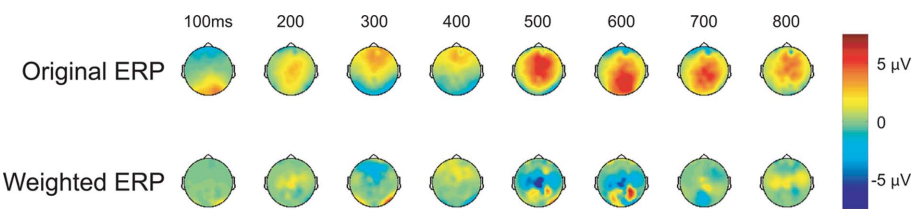


FIGURE 3 | Topographic maps of the evoked response. Upper column: grand-averaged ERP for personal (faces of friends and family) stimulus class. Lower column: same grand-averaged ERP weighted by the general classifier model (linear discriminant function).

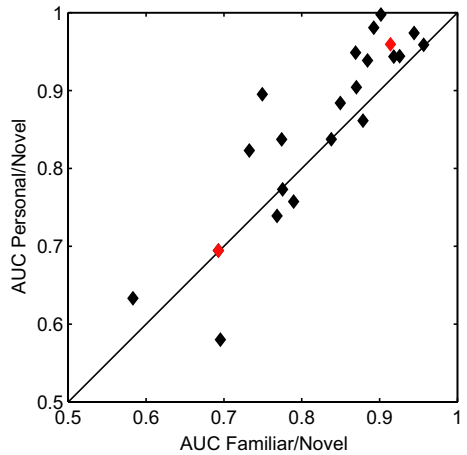


FIGURE 4 | Receiver operating characteristics analysis of classifier performance. Scatter plot of the area under the curve (AUC) for each participant. The *x* and *y* axes represent the two comparison conditions: familiar (personal and famous) versus novel and personal versus novel. Red diamonds indicate the individual participants whose ERPs are shown in **Figure 5**.

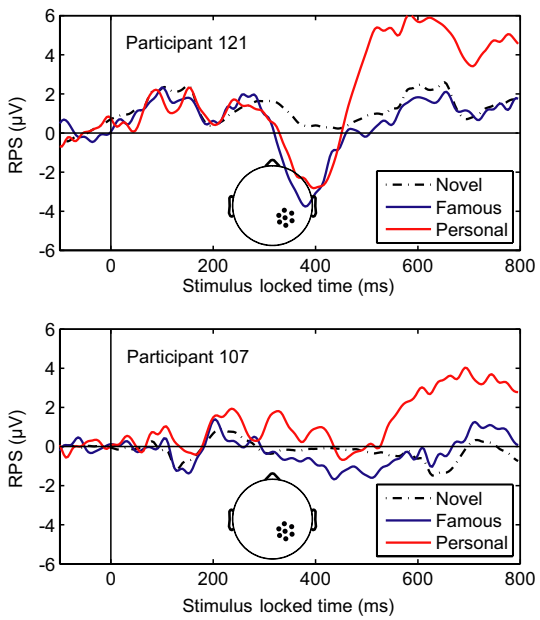


FIGURE 5 | Averaged-referenced ERPs for two individual participants. Top panel: ERP for participant 121 with high AUC (0.91). Bottom panel: ERP for participant 107 with low AUC (0.69). The AUC values were calculated using the general classifier model. Inset: the RPS channel group.

(500–600 ms) and is thus negatively affected by this type of delay. These types of individual variations in the evoked response are a primary reason for the suboptimal performance of the general model.

To accommodate the individual variations in the evoked response, we built customized models for each participant after their initial session. The custom models were built in the same manner as the general model described above (see Materials and Methods for details). To compare classification performance between the

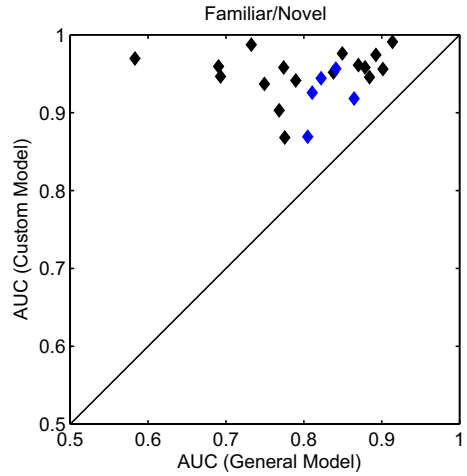


FIGURE 6 | Receiver operating characteristics analysis of classifier models. Scatter plot of the AUC for each participant. All AUC values were calculated for the familiar (personal and famous) versus novel condition. The *x* and *y* axes represent the two model conditions: a general model (same for all participants) and a custom model (unique to each participant). Blue diamonds indicate the individual participants whose custom model was validated with a second, independent experimental session.

models, we again calculated the AUC for the familiar versus novel condition (**Figure 6**). In each case, the customized model resulted in a substantial improvement in classifier performance. Here, the average AUC was 0.945 (min. = 0.868, max. = 0.991) a significant improvement over the general model ($p < 0.001$, paired *t*-test). For an additional validation of these customized models, a subset of participants ($n = 5$) returned for a second experimental session. During this second session, the custom model was used for the real-time classification of the evoked response. Again, the customized model significantly improved classifier performance ($p < 0.05$, paired *t*-test) with an average improvement of 0.06 AUC. The separate sessions controlled for any effects of over-fitting or sensitivity to exact electrode placement. In sum, while the general model was able to capture the universal neural correlate of the recognition response, the customized models improve the performance by accounting for some of the individual variations in the evoked response.

DISCUSSION

In this study we quantified the neural correlates of recognition evident the evoked response (Yovel and Paller, 2004; Curran and Hancock, 2007; MacKenzie and Donaldson, 2007). Unlike previous studies, the components of recognition, namely familiarity, and recollection, were not independently manipulated. However, it is clear from these results that both the mid-frontal FN400 (familiarity) and later parietal component (recollection) are influenced by knowledge of the individual depicted in the photograph (Paller et al., 2007). A common hypothesis describes familiarly as a sub-threshold process that does not achieve the level of contextual memory recall associated with recollection (Yonelinas, 2002; Wixted, 2007). If this is the case, one might expect this process would saturate once recollection is achieved. Our results, however, indicate that the frontal FN400 is significantly larger for individuals

personally known by the participant than for celebrities selected by the participant. Here, the entire recognition response is a not a binary operation but rather a process that reflects the participant's level of experience with the stimulus. While this study does not resolve the debate over neural correlates of familiarity and recollection, it supports the conclusion that the early and late components of the recognition response behave in a similar continuous or graded fashion.

Our results also indicate that magnitude and time course of the recognition response remains robust even when the stimuli, color photographs of faces, are much less controlled relative to previous studies. In our experiment, faces were of different genders and ethnicities, with large variations in lighting, angle, background, and resolution. Indeed, this diverse ensemble was necessary to control for the variation in participant-provided imagery. Likewise, the ERP associated with each stimulus category clearly shows the early and late components of the recognition response, even though the following stimulus was presented before the complete evolution of the response. These results offer an increased level of ecological validity to the previously described recognition response and demonstrate the utility of the signal for applications outside the laboratory.

In previous studies the level of familiarity was carefully controlled within the laboratory session. Typically, participants would encode novel stimuli (words or faces) during an explicit study phase. This process mitigates, to some degree, influences external to the experimental session and is often necessary for controlling or counterbalancing various cognitive processes (Gabrieli, 1998; Paller et al., 2007). However, one consequence of this is an attenuation of the recognition signal due to the capacity or limits of human memory. In the real world, recognition is often clear and unambiguous, especially the recognition of individuals we encounter on a daily basis (friends, coworkers, family members, etc.). For applications that seek to use the recognition signal, evidenced in the evoked response, it is important to quantify the magnitude of the effect in a more realistic setting.

Brain–machine interfaces are one of the more common applications that utilize the evoked response (Sellers et al., 2006). Many BMI technologies rely on classification of the occipital–parietal P300 signal, similar to the ERPs shown here (Krusienski et al., 2006). While this classification is typically designed to identify stimuli that are task-relevant (e.g., a particular letter or type of object), the recognition response produces a similar signal. While the BMI applications of spelling out letters to form words (Krusienski et al., 2006; Klobassa et al., 2009) or finding relevant targets in imagery is clear (Luo and Sajda, 2009; Touryan et al., 2010), one can also imagine applications that employ the recognition response as either a memory probe or a metric for learning (i.e., what items has the participant committed to memory). This study indicates that, not only is the recognition response distinguishable within single-trials (Parra et al., 2002), but that EEG classification techniques developed for visual categorization tasks can be easily adapted to recognition tasks.

Likewise, the differential neural activity based on memory (referred to as *Dm*) could be adapted for the same purpose. Previous studies have shown an enhanced posterior positivity in the ERP (with latency similar to the recognition response) for word stimuli that are subsequently remembered (Paller and Wagner, 2002). Yovel and Paller (2004) showed a similar *Dm* effect during a face-occupation memory task. Here there was a clear distinction in the evoked response for faces that were later remembered as opposed to faces later forgotten. In other words, the evoked response elicited from the initial exposure to a stimulus is predictive of subsequence memory for that stimulus. This intriguing signal offers the prospect for applications that directly measure memory encoding. Ideally, such an application could use this signal to identify information that will either be remembered or forgotten. In addition, this type of tool could be useful for exploring the neural mechanisms of memory at the level of a single object, word or atom of information. Such a memory encoding tool could be useful when designing BMI systems that seek to maximize or augment human performance.

REFERENCES

Chun, M. M., and Potter, M. C. (1995). A two-stage model for multiple target detection in rapid serial visual presentation. *J. Exp. Psychol. Hum. Percept. Perform.* 21, 109–127.

Curran, T. (2000). Brain potentials of recollection and familiarity. *Mem. Cogn.* 28, 923–938.

Curran, T., Gibson, L., Horne, J. H., Young, B., and Bozell, A. P. (2009). Expert image analysts show enhanced visual processing in change detection. *Psychon. Bull. Rev.* 16, 390–397.

Curran, T., and Hancock, J. (2007). The FN400 indexes familiarity-based recognition of faces. *Neuroimage* 36, 464–471.

Gabrieli, J. D. (1998). Cognitive neuroscience of human memory. *Annu. Rev. Psychol.* 49, 87–115.

Green, D. B., and Swets, J. A. (1966). *Signal Detection Theory and Psychophysics.* New York: Wiley.

Guo, C., Voss, J. L., and Paller, K. A. (2005). Electrophysiological correlates of forming memories for faces, names, and face-name associations. *Brain Res. Cogn. Brain Res.* 22, 153–164.

Klobassa, D. S., Vaughan, T. M., Brunner, P., Schwartz, N. E., Wolpaw, J. R., Neuper, C., and Sellers, E. W. (2009). Toward a high-throughput auditory P300-based brain-computer interface. *Clin. Neurophysiol.* 120, 1252–1261.

Krusienski, D. J., Sellers, E. W., Cabestaing, F., Bayoudh, S., McFarland, D. J., Vaughan, T. M., and Wolpaw, J. R. (2006). A comparison of classification techniques for the P300 Speller. *J. Neural Eng.* 3, 299–305.

Luo, A., and Sajda, P. (2009). Comparing neural correlates of visual target detection in serial visual presentations having different temporal correlations. *Front. Hum. Neurosci.* 3:5. doi: 10.3389/neuro.09.005.2009

MacKenzie, G., and Donaldson, D. I. (2007). Dissociating recollection from familiarity: electrophysiological evidence that familiarity for faces is associated with a posterior old/new effect. *Neuroimage* 36, 454–463.

McCarthy, G., and Wood, C. C. (1985). Scalp distributions of event-related potentials: an ambiguity associated with analysis of variance models. *Electroencephalogr. Clin. Neurophysiol.* 62, 203–208.

Paller, K. A., Voss, J. L., and Boehm, S. G. (2007). Validating neural correlates of familiarity. *Trends Cogn. Sci.* 11, 243–250.

Paller, K. A., and Wagner, A. D. (2002). Observing the transformation of experience into memory. *Trends Cogn. Sci.* 6, 93–102.

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

Perkins, S., Lacker, K., and Theiler, J. (2003). Grafting: fast, incremental feature selection by gradient descent in function space. *J. Mach. Learn. Res.* 3, 1333–1356.

Rugg, M. D., Mark, R. E., Walla, P., Schloerscheidt, A. M., Birch, C. S., and Allan, K. (1998). Dissociation of the neural correlates of implicit and explicit memory. *Nature* 392, 595–598.

Sellers, E. W., Krusienski, D. J., McFarland, D. J., Vaughan, T. M., and Wolpaw, J. R. (2006). A P300 event-related potential brain-computer interface (BCI): the effects of matrix size and inter stimulus interval on performance. *Biol. Psychol.* 73, 242–252.

Smith, M. (1993). Neurophysiological manifestations of recollective experience during recognition memory judgments. *J. Cogn. Neurosci.* 5, 1–13.

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

Touryan, J., Gibson, L., Horne, J. H., and Weber, P. (2010). "Real-time classification of neural signals corresponding to the detection of targets in video imagery," in *International Conference*

on Applied Human Factors and Ergonomics, Miami, FL, 60.
Voss, J. L., Lucas, H. D., and Paller, K. A. (2010). Conceptual priming and familiarity: different expressions of memory during recognition testing with distinct neurophysiological correlates. *J. Cogn. Neurosci.* 22, 2638–2651.
Voss, J. L., and Paller, K. A. (2006). Fluent conceptual processing and explicit memory for faces are electrophysiologically distinct. *J. Neurosci.* 26, 926–933.
Voss, J. L., Reber, P. J., Mesulam, M. M., Parrish, T. B., and Paller, K. A. (2008). Familiarity and conceptual priming engage distinct cortical networks. *Cereb. Cortex* 18, 1712–1719.
Wixted, J. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychol. Rev.* 114, 152–176.
Yonelinas, A. (2002). The nature of recollection and familiarity: a review of 30 years of research. *J. Mem. Lang.* 46, 441–517.
Yovel, G., and Paller, K. A. (2004). The neural basis of the butcher-on-the-bus phenomenon: when a face seems familiar but is not remembered. *Neuroimage* 21, 789–800.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 29 November 2010; paper pending published: 08 January 2011; accepted: 28 February 2011; published online: 11 March 2011.
Citation: Touryan J, Gibson L, Horne JH and Weber P (2011) Real-time measurement of face recognition in rapid serial visual presentation. Front. Psychology ***2****:42. doi:10.3389/fpsyg.2011.00042*
This article was submitted to Frontiers in Perception Science, a specialty of Frontiers in Psychology.

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 02 August 2011
doi: 10.3389/fpsyg.2011.00185

Physiological signal variability in hMT+ reflects performance on a direction discrimination task

Magdalena G. Wutte[1,2]*†, Michael T. Smith[3,4†], Virginia L. Flanagin[2,5] and Thomas Wolbers[3,6]

[1] Graduate School of Systemic Neurosciences, Ludwig-Maximilians-University, Munich, Germany
[2] Department of Neurology, Klinikum Großhadern, Ludwig-Maximilians-University, Munich, Germany
[3] Centre for Cognitive and Neural Systems, University of Edinburgh, Edinburgh, UK
[4] School of Informatics, University of Edinburgh, Edinburgh, UK
[5] Integrated Centre for Research and Treatment of Vertigo (IFB), Ludwig-Maximilians-University, Munich, Germany
[6] Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK

Edited by:
Guillaume A. Rousselet, University of Glasgow, UK

Reviewed by:
David J. Freedman, University of Chicago, USA
Lars Muckli, University of Glasgow, UK

***Correspondence:**
Magdalena G. Wutte, Department of Neurology, Klinikum Großhadern, Ludwig-Maximilians-University, Marchioninistrasse 23, 81377 Munich, Germany.
e-mail: mwutte@lrz.uni-muenchen.de

†Magdalena G. Wutte and Michael T. Smith have contributed equally to this work.

Our ability to perceive visual motion is critically dependent on the human motion complex (hMT+) in the dorsal visual stream. Extensive electrophysiological research in the monkey equivalent of this region has demonstrated how neuronal populations code for properties such as speed and direction, and that neurometric functions relate to psychometric functions within the individual monkey. In humans, the physiological correlates of inter-individual perceptual differences are still largely unknown. To address this question, we used functional magnetic resonance imaging (fMRI) while participants viewed translational motion in different directions, and we measured thresholds for direction discrimination of moving stimuli in a separate psychophysics experiment. After determining hMT+ in each participant with a functional localizer, we were able to decode the different directions of visual motion from it using pattern classification (PC). We also characterized the variability of fMRI signal in hMT+ during stimulus and rest periods with a generative model. Relating perceptual performance to physiology, individual direction discrimination thresholds were significantly correlated with the variability measure in hMT+, but not with PC accuracies. Individual differences in PC accuracy were driven by non-physiological sources of noise, such as head-movement, which makes this method a poor tool to investigate inter-individual differences. In contrast, variability analysis of the fMRI signal was robust to non-physiological noise, and variability characteristics in hMT+ correlated with psychophysical thresholds in the individual participants. Higher levels of fMRI signal variability compared to rest correlated with lower discrimination thresholds. This result is in line with theories on stochastic resonance in the context of neuronal populations, which suggest that endogenous or exogenous noise can increase the sensitivity of neuronal populations to incoming signals.

Keywords: inter-individual differences, motion perception, direction sensitivity, hMT+/V5, fMRI, BOLD signal variability, multi-voxel pattern classification

INTRODUCTION

Accurate perception of visual motion is a key function of the human brain, enabling us to interpret the world around us, to predict trajectories of moving objects and to steer vehicles and control locomotion. While many psychophysical and neurophysiological studies have revealed common processing of visual motion information across participants, perceptual capabilities can differ substantially between individuals (Halpern et al., 1999). Though classical behavioral experiments average these difference to focus on the mean tendency, heterogeneity in visual motion perception can provide information on perceptual functioning. Describing performance profiles of motion perception might, for example, help to distinguish subgroups in phenomena like dyslexia or describe aging processes in the visual system (Talcott et al., 2000; Slaghuis and Ryan, 2006; Bennett et al., 2007; Billino et al., 2008). Exploring the relation between differences in performance on motion tasks and physiological signals in the visual dorsal stream can shed light on the relationships between cortical processing and perception.

Motion perception in humans critically depends on area hMT+ (also known as V5, for a review see Born and Bradley, 2005). Extensive research on its equivalent in monkeys (MT) has shown that neurons in this region are selective for the direction and speed of moving stimuli. Direction sensitive neurons show columnar organization, with columns of smoothly changing preferred directions abutting columns of the opposite preferred direction (Born and Bradley, 2005). Relating neuronal characteristics to behavior, neurometric functions of single-neurons were shown to correlate with psychometric functions in a direction discrimination task (Britten et al., 1992). More evidence for a direct link between MT neuronal properties and perception comes from studies which show that microstimulation can considerably bias performance (Cohen and Newsome, 2004) and that deteriorated neuronal speed and direction selectivity accompanies aging (Yang et al., 2009; Liang et al., 2010).

In humans, hMT+ lies in an anatomically variable region and shows variation in histological and functional anatomy across individuals (Dumoulin et al., 2000; Huk et al., 2002; Malikovic et al.,

2007). Studies exploring neurophysiological properties of hMT+ have worked with exogenous variation of the stimulus (e.g., coherence of movement) to describe related modulations of the blood oxygen level dependent (BOLD) signal. Other studies have considered endogenous signal changes in hMT+ during the presentation of ambiguous stimuli, reflecting switches between percepts (for example Castelo-Branco et al., 2002; Muckli et al., 2002). The latter line of research shows the informative value of looking at endogenous fluctuations in hMT+, an approach we took in the current study to describe inter-individual physiological differences. While structural differences in the visual stream have been shown to correlate with individual psychophysical thresholds (Kanai and Rees, 2011), the connection between individual physiological properties of hMT+ and inter-individual differences in psychophysical tasks is less explored.

On a neuronal level, a possible reason for different perceptual sensitivity for direction could be the relative width of directional tuning curves. Sharper tuning curves lead to an unambiguous population signal in hMT+, which could be reflected in more distinct patterns for different directions of motion. On the behavioral level, this might translate into lower psychophysical thresholds when an individual has to make fine discrimination between different directions of motion (Purushothaman and Bradley, 2005; Liang et al., 2010). A potential candidate for revealing such physiological differences in fMRI is multi-voxel pattern analysis (MVPA) which is able to resolve fine grain patterns of hMT+ organization invisible to univariate techniques (Kamitani and Tong, 2006). Individual differences in decoding accuracy might indicate the distinctiveness of the hMT+ population pattern and correlate with perceptual performance.

Another method which has been recently suggested as a good gauge for inter-individual comparisons is variability analysis of the BOLD signal (Garrett et al., 2010; Mohr and Nagel, 2010; Samanez-Larkin et al., 2010; Mennes et al., 2011). Measurements of variability aim to describe endogenous background fluctuations in the signal, which appear independent of the timecourse of the experimental manipulation. An important confound for accurately measuring such endogenous variability is that the relationship between the stimulus and the BOLD signal has to be described as precisely as possible. Only if this is achieved can one investigate if the observed physiological variability has functional significance. A growing body of studies suggests that neurophysiological variability patterns can be understood as (functional relevant) "signal" rather than (function disturbing) "noise" (Faisal et al., 2008; McDonnell and Abbott, 2009; Garrett et al., 2010). Population signal variability in hMT+ could have different effects on performance accuracy: higher overall variability levels in hMT+ could be detrimental for discrimination performance if they would have an destabilizing effect on the hMT+ population signal as some authors suggest for the dopamine system (Winterer et al., 2006; Samanez-Larkin et al., 2010). Alternatively, a certain level of variability has been described to improve the sensitivity of systems, e.g., by stabilizing synchronized oscillating populations (Ermentrout et al., 2008), an observation described as stochastic resonance (Emberson et al., 2007; McIntosh et al., 2008; McDonnell and Abbott, 2009; Garrett et al., 2010).

In the present study, we set out to characterize brain activity that correlates with inter-individual variability in the accuracy of visual motion perception. We used multivariate pattern classification (PC) to describe hMT+ population patterns and we characterized the variability of the hMT+ BOLD signal during perception of motion in different directions. We investigated if these measures can serve as sensitive indicators for inter-individual performance differences on a motion direction discrimination task.

MATERIALS AND METHODS

PARTICIPANTS

Fifteen healthy subjects gave written informed consent to participate in this study. The study was performed in accordance with the Declaration of Helsinki and approved by the ethics committee of the medical faculty of the Ludwig-Maximilians University Munich. Handedness was determined according to a 10-item excerpt of the "Handedness Inventory," coding the degree of handedness (+100: exclusively right handed, −100: exclusively left handed; Oldfield, 1971). It resulted in +100 in 13 subjects, one with +64 and one with +81. All subjects had normal or corrected-to-normal visual acuity as determined binocularly with a Snellen table (0.8 on 6 m or better). None of the subjects were taking medication or had any history of neurological disease. All subjects understood the instructions without difficulty. One subject was excluded from the MR analysis due to excessive motion resulting in a final cohort of 14 subjects (age range: 21–27, 6 female). These 14 subjects consecutively also took part in the psychophysical task on direction discrimination. Three subjects were excluded from psychophysical data analysis, as their measurements did not fulfill stability criteria as described below.

PSYCHOPHYSICS

Apparatus

Stimuli were generated by a Fujitsu Siemens Pentium(R) 4 CPU at a frame rate of 85 Hz and displayed on a 40-cm × 30-cm Conrac Elektron CRT monitor driven by a NVIDIA Quadro Pro2 graphics card. The monitor resolution was set to 1280 × 1024. White and black pixel had a luminance of 25.3 and 0.1 cd/m^2, respectively, resulting in a maximum Michelson contrast of 99%. Experiments were conducted in a darkened room and subjects were seated in 60 cm distance from the monitor.

Stimulus

Stimuli were programmed in Matlab 7.3 using the Psychophysics Toolbox extensions (Brainard, 1997). Coherent translational flow fields were presented in a circular aperture (11.4° × 11.4°), containing 300 white dots (diameter: 0.1°) at a time on a black background. All dots of one stimulus moved in a upward direction either vertically or at a small tilt from the vertical with a speed of 8°/s. Dots moving out of the aperture reappeared at new random positions (at the bottom of the aperture). Stimulus intensity was defined as the degree of tilt of the match stimulus (clockwise or anticlockwise) in respect to the upward (0°) reference stimulus.

Procedure

A two-alternative-forced-choice task was used to determine individual thresholds and psychometric functions of direction sensitivity. Reference stimulus and match stimulus were presented consecutively

(stimulus duration: 1.5 s, inter stimulus interval: 0.25 s, intertrial interval: 1.25 s). While fixating on the center of the aperture, subjects indicated with a buttonpress whether the second stimulus (match) was tilted clockwise or anticlockwise with respect to the first, upward moving reference stimulus (compare **Figure 1A**). After initial training with feedback (60 trials), preliminary thresholds were determined by two repetitions of a 3-down-1-up adaptive double-staircase method (140 trials). The staircase measure was defined as stable if the slope of the linear fit from the last 12 reversals was less than 0.02. All but one subject achieved stable staircase measurements (this subject belonged also to the outliers in the measurement of constant stimuli, defined as subjects whose threshold exceeded the fourth quartile, see 2.2.4). Consecutively, the method of constant stimuli was used to sample the psychometric function, the range of sampling was set around the threshold determined by the staircase measurements. Tilt was varied between seven different intensities and each intensity was presented in 30 trials, resulting in a total of 210 trials. Subjects answered following the second stimulus and both speed and accuracy of the response were emphasized. Response times were measured from the moment the second stimulus ended until the moment of response. No feedback was given in staircase or constant stimulus measurements.

Data analysis

Data was analyzed using psignifit toolbox (Wichmann and Hill, 2001a,b) in Matlab 7.3. Final thresholds were obtained by fitting the percentage of correct responses determined by the method of constant stimuli with a cumulative Weibull distribution using a maximum likelihood procedure. Free parameters were threshold, slope, and lapse rate, which was kept variable between 0 and 0.5 (Wichmann and Hill, 2001b). Thresholds were taken as the 0.5 cut-off from the fitted function, corresponding roughly to a performance level of 75% correct (see **Figure 1B**).

To ensure data reliability, those subjects whose thresholds exceeded the fourth quartile were excluded from further analysis (2 of 14). Subjects were furthermore excluded if the fit of their psychometric function did not meet goodness-of-fit criteria in the sensitivity analysis. Summary statistics yielded good fits between the psychometric function and the data for 11 of the 12 remaining subjects. Ninety-five percentage confidence intervals (CI) were calculated for the thresholds of each subject using the bootstrapping method (sampling with replacement, 1999 repetitions).

A one-way Kruskal–Wallis ANOVA tested for inter-individual differences in the behavioral thresholds, using the bootstrapped results.

Averaged reaction times (RT) were calculated as the arithmetic mean over the whole constant stimuli experiment. RT consistency was calculated as the SD over the experiment.

MAGNETIC RESONANCE IMAGING

Experimental stimulus and procedure

Visual stimuli were projected with a LCD projector on a screen placed behind participants in the MR-scanner, which they viewed through a mirror placed above them at 45°. Vizard 3.0 (Worlviz)[1], was used to produce coherent translational flow fields presented in a circular aperture (300 dots per display, aperture size 11.4° × 11.4°). Participants watched flow fields in one of four possible directions (0°, 90°, 180°, and 270°), shown in a randomized order, while fixating on a cross in the middle. Using a block design, 18 s task periods were interleaved with 10 s rest periods, during which subjects continued fixating. One block consisted of four trials, in which direction of motion was kept constant. Subjects performed a two-alternative forced-choice speed discrimination task, to keep their attention directly related to the movement of the stimulus while incidentally coding stimulus direction. In each trial, two consecutive

[1]http://www.worldviz.com/

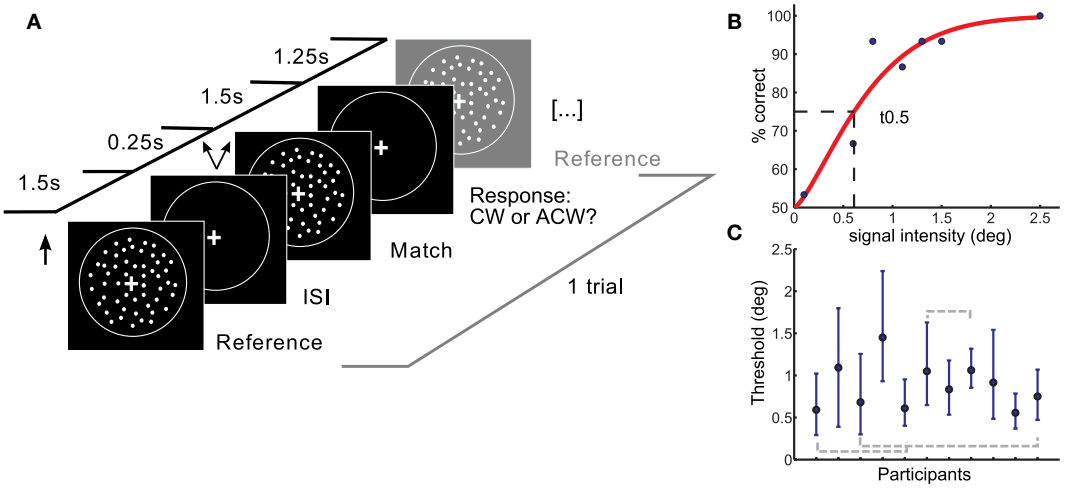


FIGURE 1 | Setup and results of the psychophysical experiment. (A) A coherent flow field of white dots was moving in an upward vertical direction followed by a match stimulus moving upward with a slight tilt clockwise or anticlockwise from vertical. Participants responded following the match stimulus. One trial is depicted. **(B)** Psychometric function of one representative subject. Percent correct responses are plotted as a function of the deviation (in degree) of the match stimulus from the reference stimulus. Thresholds were taken at the mid-point of the psychometric function (t0.5, shown by dotted lines). **(C)** Subjects yielded different thresholds. Shown are individual t0.5 thresholds with errorbars depicting the 95% confidence intervals. Subjects which were not found to be different in a *post hoc* test are shown with gray dotted lines. deg, degree tilt from vertical; ISI, interstimulus interval; CW, clockwise; ACW, anticlockwise.

stimuli were shown, a reference speed of 8°/s and a match stimulus of faster speed randomly distributed to the first or second presentation (stimulus duration: 1.5 s, interstimulus interval: 0.25 s, intertrial interval: 1.25 s, as for the psychophysical stimulus). Subjects reported the order-position of the faster stimulus with a buttonpress (see **Figure 2A**). For keeping task difficulty constant, individual speed discrimination thresholds were kept at a task performance of about 80% correct with an adaptive staircase procedure (QUEST, Watson and Pelli, 1983). Subjects performed 8 runs for a total of 32 repetitions per direction. Participants practiced the task outside the MR-scanner until they reached a satisfactory performance level (2 runs in which participants had to be error-free for 12 trials (fixed velocity difference) after which a staircase procedure started, on which subjects had to demonstrate a stable 80% correct threshold for at least 12 trials). They also practiced inside the bore of the MR-scanner, until they were comfortable conducting the task in a supine position.

A separate fMRI experiment was conducted to functionally localize hMT+ in each subject, according to previously established procedures (Morrone et al., 2000; Huk et al., 2002). Briefly, a stimulus of alternating moving and stationary dot patterns was presented in a circular aperture with interleaved rest periods. Moving dots (velocity: 17.1°/s) traveled toward and away from the fixation cross for 16 s, followed by a 16-s stationary dot field, and a 20-s blank screen. Subjects fixated at all times.

fMRI acquisition

Imaging data were acquired on a 3T MR-Scanner (GE Sigma HDx) with a standard 8 channel head coil using an echo-planar imaging sequence (TR: 2 s, echo time: 40 ms, flip angle: 70°) to acquire 25 slice volumes (interleaved acquisition, no gap), centered on the area of interest (medial temporal lobe). Voxel size was 1.75 mm × 1.75 mm × 2.4 mm. In total, 8 runs of 225 volumes for

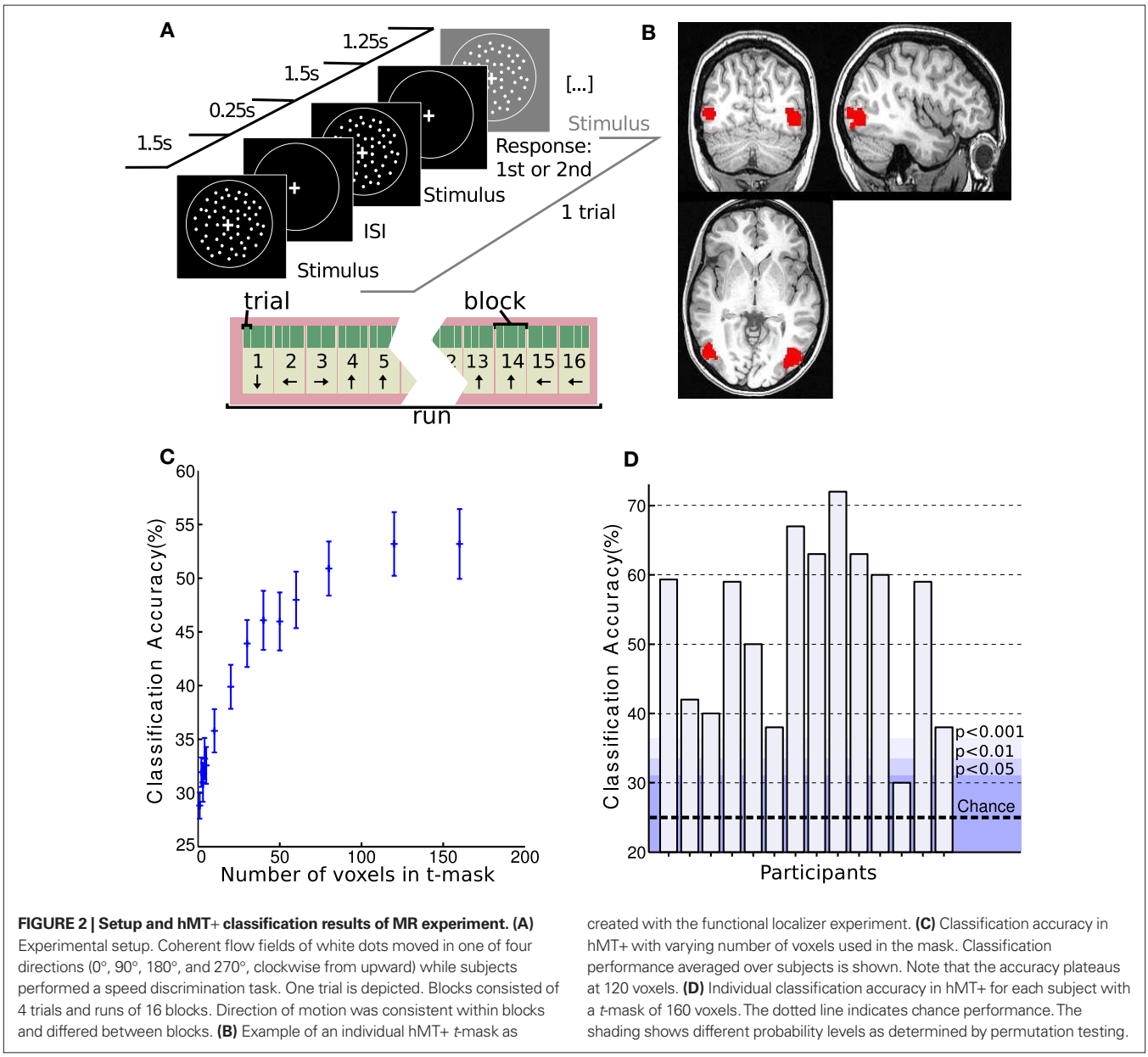


FIGURE 2 | Setup and hMT+ classification results of MR experiment. (A) Experimental setup. Coherent flow fields of white dots moved in one of four directions (0°, 90°, 180°, and 270°, clockwise from upward) while subjects performed a speed discrimination task. One trial is depicted. Blocks consisted of 4 trials and runs of 16 blocks. Direction of motion was consistent within blocks and differed between blocks. **(B)** Example of an individual hMT+ *t*-mask as created with the functional localizer experiment. **(C)** Classification accuracy in hMT+ with varying number of voxels used in the mask. Classification performance averaged over subjects is shown. Note that the accuracy plateaus at 120 voxels. **(D)** Individual classification accuracy in hMT+ for each subject with a *t*-mask of 160 voxels. The dotted line indicates chance performance. The shading shows different probability levels as determined by permutation testing.

the experimental condition and 1 run of 132 volumes for the functional hMT+ localizer were acquired in each subject. In addition, a T1-weighted anatomical volume was acquired.

Defining hMT+ and V1 masks

To define functional regions of interest, fMRI data from the functional localizer were realigned to the first volume of the timeseries and smoothed with a kernel of 4 mm FWHM as implemented in SPM8 (Wellcome Department of Imaging Neuroscience, London, UK). Data were processed in individual space. A general linear model analysis comprising regressors for motion and stationary conditions was performed. Contrasting motion and stationary regressors identified clear delineated clusters for hMT+ (FWE, $p < 0.05$ in all but two subjects, who showed hMT+ clusters only at $p < 0.001$ uncorrected). See **Figure 2B** for an example. The clusters from the two hemispheres were combined to make a hMT+ mask of voxels for further analysis.

The V1 mask was created using anatomical and functional constraints. V1 was determined anatomically using FreeSurfer's cortical parcelation algorithms in every subject, based on anatomical constrains described by Hinds et al. (2008). The final mask consisted of voxels within this anatomically defined V1 which showed significant activation in the functional localizer, using the motion–stationary contrast.

Multivariate pattern classification and preprocessing

We used the Princeton Multi-Voxel Pattern Analysis Toolbox (MVPA)[2], to test whether voxels within hMT+ or V1 contained information about the direction of the stimulus. Data were prepared by unwarping, realigning (SPM8), and detrending (MVPA) the timeseries to remove linear trends and high-pass filtering (cut-off: 128 s) to remove low frequency noise. Z-scoring of response amplitudes for stimulus periods of individual voxels was applied to minimize baseline differences across runs and to reduce the impact of outliers. To account for the latency of the hemodynamic response, all stimulus onset times were shifted forward in time by 4 s as described previously (Kamitani and Tong, 2006). Data were neither smoothed nor spatially normalized, to avoid signal degradation and preserve inter-individual differences. The nine image volumes from each block of four trials were combined to generate a single average volume for each block.

The 160 voxels with the highest *t*-values in the functional localizer experiment were selected from the hMT+ or V1 masks respectively for decoding analysis. We tested different mask sizes, but found no improvement in classification accuracy beyond 160 voxels (see **Figure 2C**).

The lSVM (linear support vector machine) classifier was chosen as it provided stable results across participants without overfitting. It was used with a fixed cost, $c = 1$. Classification used standard leave-one-out cross-validation, in which the data set was divided, with seven runs in the training set and one run in the testing set. The test was repeated eight times, with each different run being the test set (Pereira et al., 2009). The accuracy scores reported represent the proportion of blocks in which the classifier correctly decoded directions.

[2]www.pni.princeton.edu/mvpa

Generating a stability index to quantify head motion

An index was designed to assess data stability for individual subjects. Head-movement causes image shifts between classifier training and test periods which are detrimental for MVPA. Specifically, a movement in the middle of the acquisition is more detrimental than a movement at its start or end because there will be more cross-validation iterations in which the training set contains volumes misaligned with the test set's volumes. Our stability index (SI) roughly represents the longest stable stretch of head orientation during data acquisition. For each volume, the location of the center of hMT+ is estimated from the realignment parameters generated during image preprocessing. Each volume is compared with all others. At each comparison (e.g., between volumes i and j), the distance, d_{ij} between the estimated locations of hMT+ is calculated and a number, A_{ij}, assigned describing how aligned the pair of volumes are. This alignment score is

$$A_{ij} = \frac{1}{1 + d_{ij}} \tag{1}$$

The similarity S_i of each volume with all the other volumes is summarized by summing over all of its alignment scores:

$$S_i = \sum_j A_{ij} \tag{2}$$

Finally, the whole recording session is given a SI, which is the score for the volume with the highest similarity score:

$$\mathrm{SI} = \max_i S_i \tag{3}$$

Figure 3A illustrates how the similarity value varies for different time points over a fictitious series of volumes. The example shows little head motion during the longest part of data acquisition and a single large head motion toward the end. Similarity values for volumes in the long stable period are higher than for those after the movement, because the volumes in the former are similar to many more timepoints than a volume taken after the movement.

Estimating BOLD signal statistics using a generative model

Variability of the timecourses of the 160 voxels from the above described hMT+ and V1 masks was assessed with a generative model for stimulus (SDstim) and rest periods (SDrest) (see **Figure 4B** for an illustration of the model). In addition, variability was estimated in a white matter region to quantify the contribution of non-physiological variability to noise, as those regions show little change in local metabolism (Rostrup et al., 2000). Spheres of 80 voxels in each hemisphere were selected from the anterior portion of the corona radiata (CR), as determined by the Harvard–Oxford structural atlas (see **Figure 3C** for an example).

Timecourses were high-pass filtered before model analysis. The temporal properties of the BOLD signal were described by modeling all eight events within a stimulus block as box-cars (1.25 s duration), which is similar to modeling them as delta functions as used in event-related designs. Box-cars were then convolved with the canonical hemodynamic response function (HRF), to account for the latency of the BOLD signal. A mixing parameter α_i was generated by this function and assigned to each volume i, describing the proportion of the signal recorded at that timepoint that was provided by the stimulus periods.

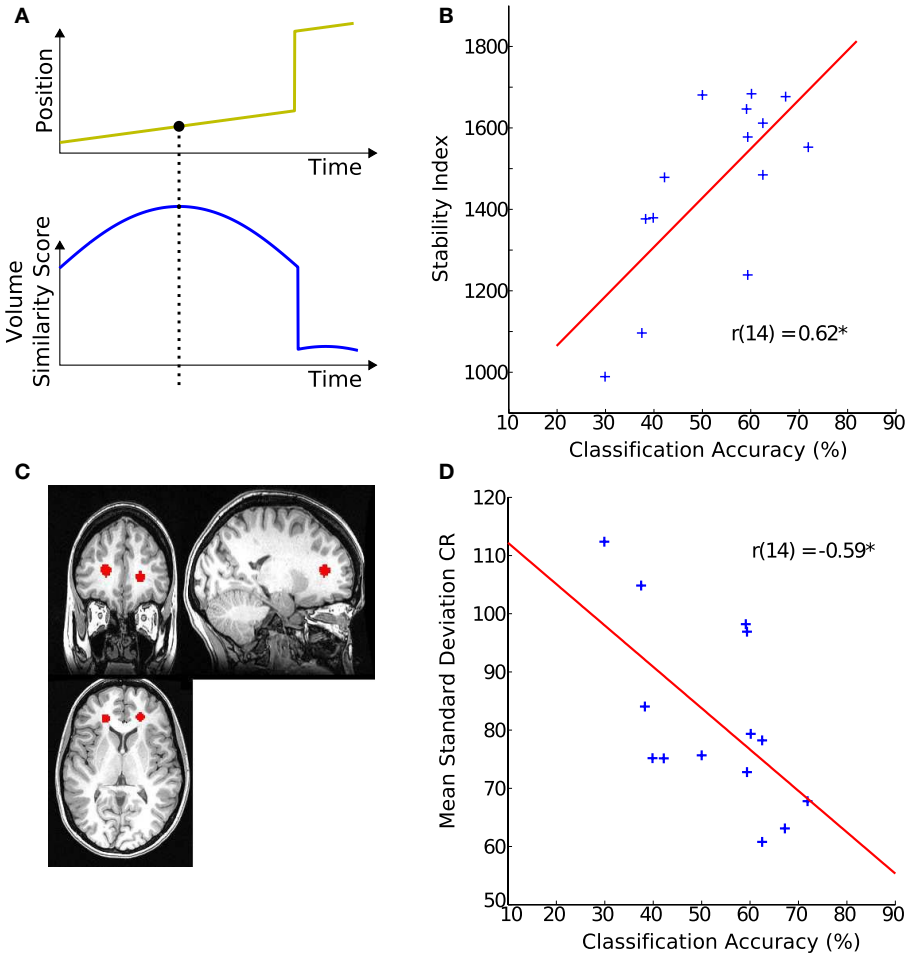


FIGURE 3 | Origins of classification variability. (A) Schematic of the stability index. The upper plot shows simulated movement of a participant across a scanning session. The lower plot indicates how similar each of the volumes is to all the other volumes in the scan. For example the last few volumes are very dissimilar to the rest of the scan. The value of the volume with the highest similarity score is used as this participant's overall stability index. **(B)** Stability index and hMT+ classification accuracy. The stability index was positively correlated with decoding accuracy. **(C)** Region of interest in white matter. The region is defined by placing two spheres of 80 voxels each in the anterior portion of the corona radiata (CR). **(D)** Signal variance in white matter and hMT+ classification accuracy. The individual SD during block-periods in CR was negatively correlated with the decoding accuracy in their hMT+.

For the generative model, both stimulus and rest periods were modeled as gaussian distributions (stimulus: mean μ_s, variance σ_s^2; rest: mean μ_r, variance σ_r^2). The proportion of each of these distributions included in the final signal was estimated using maximum likelihood estimation (MLE). The estimate of the signal for a particular time point was calculated by finding the weighted sum of the two distributions. The mean and variance of the sum of two independently distributed gaussian random variables was found by adding the means and variances of the two distributions. So the mean and variance of the new distribution for time point *i* could be written as:

$$\mu_{ci} = \alpha_i \mu_s + (1-\alpha_i)\mu_r \tag{4}$$

$$\sigma_{ci}^2 = \alpha_i^2 \sigma_s^2 + (1-\alpha_i)^2 \sigma_r^2 \tag{5}$$

this allows one to write the probability of value x_i as

$$p(x_i) = N(\mu_{ci}, \sigma_{ci}^2) \tag{6}$$

Assuming independent and identically distributed sampling, the likelihood of the whole timeseries is:

$$p(x) = \prod_i N(\mu_{ci}, \sigma_{ci}^2) \tag{7}$$

The log likelihood therefore is:

$$LL(x) = \sum_i log\, N(\mu_{ci}, \sigma_{ci}^2) \tag{8}$$

$$LL(x) = -\sum_i \tfrac{1}{2} log\, \sigma_{ci}^2 - \sum_i \frac{(x_i - \mu_{ci})^2}{2\sigma_{ci}^2} \tag{9}$$

The four parameters were estimated by maximizing this function with respect to each of them.

Finally, the arithmetic difference between the SD within stimulus periods (σ_s) and within rest periods (σ_r) was calculated for each participant (SDdiff).

Adaptation model. The above model is not the only conceivable description of the signal timecourse. An alternative model was tested to assess the stability of our results yielded with the first approach. In this alternative model, possible signal adaptation in hMT+ over a block was accounted for by introducing a exponential decay term with a time constant of 5 s. This reduced the (pre-HRF convolved) box-car signal exponentially while the stimulus was applied, and allowed it to recover using the same exponential function during the stimulus-off periods. A maximum reduction of 14% in the BOLD response due to the adaptation was assumed, based on electrophysiological studies (Petersen et al., 1985; Krekelberg et al., 2006). The model was tested at four values of the time constant: 5, 10, 20, and 40 s. All other parameters of the model were kept constant.

ASSESSING EYE MOVEMENTS FROM fMRI DATA

Although subjects were instructed to fixate, we were concerned that systematic eye movements occurred. It has been shown previously that eye movements can be estimated from fMRI data by analyzing the timecourse of fMRI signal in the vitreous of the eye (Beauchamp, 2003). We took this retrospective approach in those subjects in which the eyeball was partially contained in the field of view (FOV; in 3 of 11 participants the eyeballs were to 33, 40, and 46% contained in the FOV, see **Figure 5**). We defined a region of interest for the available section of the eyeball using FreeSurfer. The mean timecourse was extracted using marsbar in SPM8. To estimate the dependency between eyeball signal and the rest of the brain, we used the eyeball timecourse as regressor in a GLM, as has been described previously (see Muckli et al., 2009 supplementary material).

CORRELATION OF BEHAVIORAL DATA WITH MR MEASUREMENTS

A Pearson correlation was calculated between individual thresholds from the behavioral experiment (t0.5) and the individual noise difference between block and rest periods as determined by the generative model (SDdiff). Additionally, a Spearman correlation was performed which also showed a significant correlation. The robustness of the significant Pearson correlation was estimated using bootstrapping, sampling with replacement with 2000 iterations, to produce 95% CI for the r distributions.

RESULTS

INTER-INDIVIDUAL VARIABILITY IN DIRECTION DISCRIMINATION

On average, direction discrimination thresholds were found to be similar to previous results (Westheimer and Wehrhahn, 1994). We observed significant differences in discrimination thresholds between subjects (Kruskal–Wallis ANOVA, $p < 0.001$). *Post hoc* analysis also revealed similarities in subgroups of subjects, in three subject pairs (see **Figure 1C**: there was no significant difference between subject 1 and 4, between subject 3 and 11 and between subject 6 and 8). Note that data stem from 11 subjects, as three subjects did not reach reliability criteria explained in Materials and Methods. Slopes of the individual psychometric functions were heterogeneous as well and showed a negative correlation with threshold (the higher the slope, the lower the threshold). The width of subjects 95% CI also differed between subjects. Average RT and RT consistency varied between subjects (max: 460 ms, min: 176 ms, SD: 67 ms, and SD max: 149 ms, SD min 57 ms respectively). RT means or variability did not correlate with individual direction discrimination thresholds.

PATTERN CLASSIFICATION IS CONFOUNDED BY RESIDUAL HEAD MOTION AND CANNOT EXPLAIN PERCEPTUAL DIFFERENCES

Replicating previous results (Kamitani and Tong, 2006), the linear SVM was able to discriminate between the four motion directions in hMT+ with above chance accuracy ($\mu = 53 \pm 13\%$, $p < 0.002$ using permutation testing) in all but one participant (see **Figure 2D**). Also consistent with previous results, classification accuracy was still higher in V1 ($\mu = 65 \pm 12\%$, $p < 0.001$).

To test if individual classification scores in hMT+ or V1 were related to performance on the direction discrimination tasks, a correlation analysis between scores and psychophysical thresholds (t0.5) was performed which showed no significant effect (hMT+: $r = 0.15$, $p = 0.64$; V1: $r = 0.16$, $p = 0.64$).

To investigate possible reasons for inter-individual differences in classification scores, we looked at its correlation with non-physiological noise of the MR signal. Classification accuracy correlated significantly with variability (SDstim) in the white matter region CR ($r = -0.59$, $p < 0.03$, **Figure 3D**), from which we concluded that the level of global noise determined the differences in decoding success rather than local hMT+ noise.

To test this, we looked at one of the largest methodological contributors to variability in MR signal: head-movement (Friston et al., 1996; Lund et al., 2005). A strong correlation was observed between the SI reflecting stability of the signal and classification accuracy ($r = 0.62$, $p < 0.02$, **Figure 3B**).

This implies that noise induced by subject movement is the predominant cause for differential classification accuracies in subjects. Being this sensitive for head-movement artifacts, PC differences between subjects are unlikely to be a viable method to investigate physiological differences between subjects.

A GENERATIVE MODEL FOR ASSESSING BOLD SIGNAL VARIABILITY

We used the arithmetic difference between SD of block and rest periods (SDdiff) to look at variability of the MR signal in hMT+ and V1 in individual participants. Being a relative measure, it was assumed to be largely resistant to movement induced artifacts and background scanner noise, as those would influence both periods to the same extend.

Considerably more variability was found in the hMT+ region than in a white matter region (CR), both within stimulus blocks, and rest periods (SD was 30% higher in hMT+ and V1 than in CR). The SDdiff was also found to be larger in hMT+ and V1 than in CR (36%).

Importantly, subjects with a larger noise difference in hMT+ between rest and blocks did not have larger SI scores ($r = -0.4810$, $p = 0.0695$) which demonstrates that SDdiff is less affected by head motion.

VARIABILITY PATTERNS IN hMT+, BUT NOT V1, CORRELATE WITH DIRECTION SENSITIVITY

In the final analysis, we tested whether inter-individual variability of perceptual performance was correlated with variability characteristics of the hMT+ signal. As can be seen in **Figure 4A**, we observed a significant correlation between psychophysical threshold and SDdiff: participants with a greater SDdiff showed better behavioral performance (smaller thresholds) compared to participants with a smaller SDdiff ($r = -0.61$, $p < 0.046$, bootstrap CI 95% for r: −0.87 to −0.23). In other words, the larger the difference in variability (stimulus block minus rest), the lower the threshold the respective subject achieved. Similar correlation results were found for estimating SDdiff with an alternative model taking into account adaptation effects within blocks ($r = -0.59$, $p < 0.058$, bootstrap CI 95% for r: −0.84 to −0.20).

To investigate the specificity of this effect, we also correlated SDdiff in the CR with the psychophysical thresholds which was not significant ($r = -0.35$, $p = 0.29$, **Figure 4C**, lower panel). To test another region involved in direction coding, we correlated SDdiff of V1 with psychophysical thresholds. We did not observe a significant correlation in V1 neither ($r = -0.44$, $p = 0.181$, **Figure 4C**, lower panel).

When the MR-blocks were split into those with stimuli of different directions, the effect remained significant for vertical but not horizontal motion (see **Figure 4C**, upper panel). Given that the stimulus in the psychophysics experiment were visual flow fields moving vertically upward, this might indicate that we are observing a phenomenon specific for vertical motion. Alternatively, one could interpret this observation as showing a general bias for vertical versus horizontal motion in hMT+. Further studies are necessary to clarify this point.

EYE MOVEMENT ANALYSIS

The hMT+ is known to be influenced by eye movements (Dukelow et al., 2001; Acs and Greenlee, 2008). For this reason we instructed subjects to fixate, with which they reported no difficulties. We can not exclude however, that eye movements occurred. To investigate this, we used a retrospective approach to assess, if the signal timecourse of the eyeballs taken from the EPI images correlates with fluctuations in hMT+. In the three subjects analyzed, we did not observe significant correlations of eyeball signal timecourse with fluctuations in area hMT+ (see **Figure 5**).

DISCUSSION

We demonstrate in the current study that inter-individual differences in performance on a direction discrimination task of visual motion are correlated with signal variability characteristics of hMT+ but not V1. We furthermore show that PC, though being able to decode direction from hMT+ within subjects, is a poor tool to describe inter-individual differences. Assessing individual BOLD signal variability difference in stimulus and rest periods is shown to be a better measure for such comparisons, being less influenced by non-physiological noise.

Differences in psychophysical thresholds between subjects show that perceptual sensitivity for motion direction is variable even within a homogeneous sample. Worse or better perception of motion stimuli in subjects with normal visual acuity has been suggested to reflect changes in higher level visual cortical areas rather than in the peripheral apparatus (Halpern et al., 1999).

Relatively little is known about hMT+'s contribution to worsening of direction perception (Bennett et al., 2007; Billino et al., 2008), although concepts like the "magnocellular theory" behind learning disorders like dyslexia attribute a partial cause of the phenomenon to perceptual malfunctioning in the dorsal visual stream (Stein, 2001). Other authors already suggested that BOLD signal variability over the whole brain (Garrett et al., 2010) or in specific regions like the nucleus accumbens (Samanez-Larkin et al., 2010) might have predictive value for degradation of function during aging. Our method of characterizing signal variability in hMT+ could help the clinical understanding of degraded motion perception in aging or disorders like dyslexia.

Better performance in the psychophysical task suggests higher perceptual sensitivity in that particular participant and thereby most likely more effective processing in the brain. Our results show that variability characteristics in hMT+ but not V1 correlate with psychophysical thresholds. This might indicate that we observe individual differences not at the initial encoding of the visual information in V1, but rather during a more complex motion processing step in hMT+, an area thought to drive perceptual decisions in higher cortical areas.

We find lower thresholds correlating with larger variability differences between stimulus and rest periods which mean higher variability levels in stimulus periods (but see the below discussion on model bias as a limitation to this claim). How could increased random physiological signal be beneficial for the sensitivity of a system? An influential theory based on the phenomenon of stochastic resonance advertises "[...]randomness that makes a non-linearity less detrimental to a signal." (McDonnell and Abbott, 2009). The theory asserts that a certain level of noise can actually be beneficial for signal transmission. Studies have shown that a certain level of endogenous noise can make synchronized oscillating populations more stable (Ermentrout et al., 2008; Ghosh et al., 2008) and benefits the emergence of fast oscillations in local field potentials (Brunel and Wang, 2003). For us this means that detecting higher levels of endogenous variability in the hMT+ population signal might actually reflect a more robust signal.

Other fMRI and EEG studies have described lower levels of cortical noise in senior subjects (Garrett et al., 2010) and children (McIntosh et al., 2008) compared to young adults. This has been discussed as neurophysiological noise being inversely related to the well described U-shaped function of performance during the lifespan (MacDonald et al., 2006; McIntosh et al., 2008).

More specifically for our case of signal variability in the visual system, Bair et al. (2001), recording single-neurons in macaque MT, describe that those neuron pairs with high signal correlations also showed an increase in the correlation of noise. Clearly, given the coarse MR resolution, correlated noise would be more detectable at the fMRI level than uncorrelated noise. Our results suggest that greater variability differences between stimulus and rest periods might be beneficial for perceptual sensitivity in hMT+. The basis for signal variability could be caused by individual neurophysiological characteristics of hMT+.

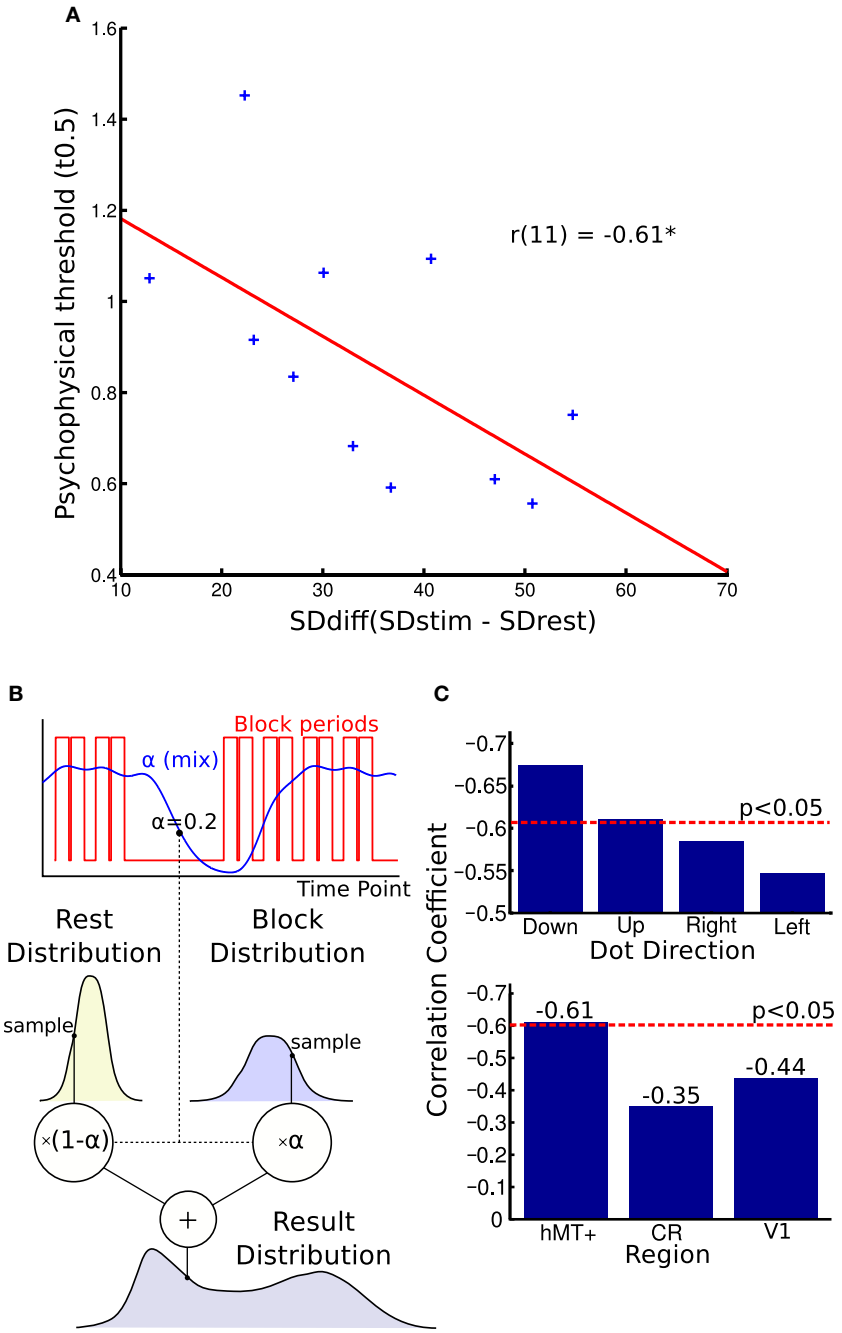


FIGURE 4 | Blood oxygen level dependent (BOLD) signal variability and behavioral performance. (A) Correlation of BOLD signal variance and direction discrimination threshold. The difference in individual SD between the blocks and rest periods correlated with single-subject thresholds from the psychophysics experiment. A larger variability difference is correlated with lower direction discrimination thresholds. **(B)** This figure illustrates the generative model used to estimate the parameters of the two distributions. The graph shows how the alpha "mix" values are calculated from the block times. Each volume's alpha value is used to estimate what proportion of the signal is from the stimulus and what proportion is from the rest period. These two distributions are sampled and their weighted sum is found. This is used to generate the distribution. The log likelihood of the real distribution being generated in this way is calculated. The parameters of the block and rest distributions are then altered to maximize this log likelihood. **(C)** Top graph: Comparing correlations for different stimulus directions. Splitting the block and rest periods in the four directions shown during the MR experiment, we observed small differences in correlation strength. Bottom graph: Comparing correlations over different brain regions. The correlation between noise difference and psychophysical threshold was smaller and not significant in the white matter region CR and V1. CR, corona radiata.

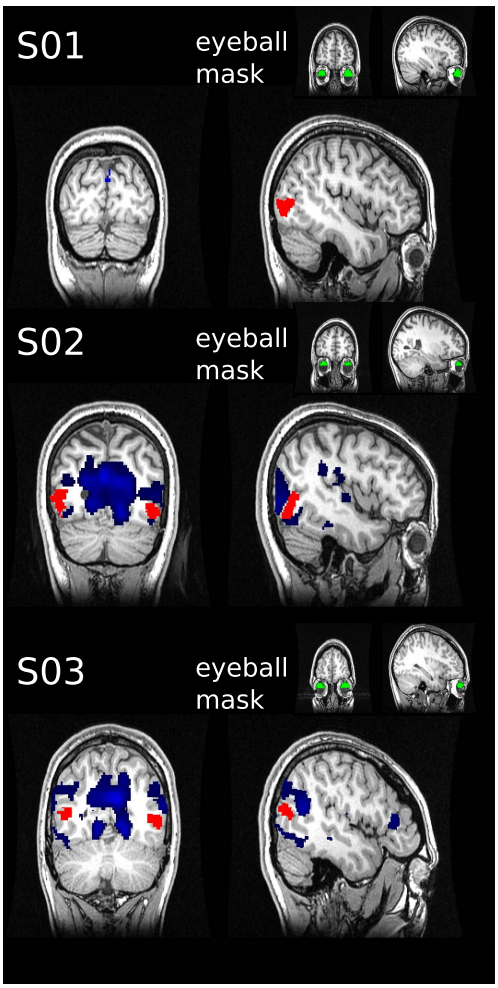


FIGURE 5 | Retrospective eye movement analysis. The mean timecourse was extracted from the eyeball ROIs (green) defined for three subjects. The timecourses were used as regressors in general linear models to assess correlated activity in the rest of the brain (blue). No overlap was found with the hMT+ masks (red).

A confound that must be considered before interpreting our variability signal is signal fluctuations in hMT+ caused by eye movements. Participants were instructed to fixate, but as we used translation stimuli, an automatic smooth pursuit must have been suppressed which individual subjects might have achieved with more or less success over the timecourse of stimulation. However, we did not find that the signal timecourse from the eyeball ROIs as measured in a subgroup of subjects correlated with hMT+ signal fluctuations. Participants furthermore did not report difficulties fixating. Although we cannot exclude an influence of eye movements on the hMT+ signal, we believe it is not the strongest component causing the observed inter-individual differences in fluctuation of the hMT+ signal.

Also non-perceptual phenomena like individual motivation or attentional levels could explain our results, influencing both physiological and perceptual measurements. It has been shown that hMT+ BOLD signal is modulated by attention (Berman and Colby, 2002; Liu et al., 2011; Stoppel et al., 2011), and those participants able to apply attention most accurately to the stimulus are not only likely to do well in the psychophysics direction discrimination task, but may also show the BOLD signal variability we observe. Top-down control by areas described for internally evoked attention processes like the intraparietal cortex and superior frontal cortex could play a role in inducing the individual hMT+ signal variability we observe (Corbetta and Shulman, 2002).

From the methodological point of view, we demonstrate that PC is a poor method to determine between subject differences. Although it could decode directional information from hMT+ activity in individual subjects, its ability to describe the relative difference between subjects was confounded by individual head-movement and scanner artifact differences. Filtering out movement artifacts has been a challenge in the field of MR, as it contributes the greatest amount of non-physiological noise (Friston et al., 1996; Lund et al., 2005). Although successful methods have been established for reducing the effect of head-movement in univariate analysis based on the general linear model (Friston et al., 1996; Andersson et al., 2001), the specific influence of residual artifacts on new methods like PC is less well documented. Beyond this methodological confound, other evidence exists that classification accuracy may not be an appropriate metric to compare experimental conditions, brain regions, or participants. Smith et al. (2011) for example suggest that classifier performance is influenced by other factors besides neural specificity such as response amplitude. Using MVPA for between subject comparison might therefore require further corrections to guarantee comparability.

Head-movement artifacts can also confound measures of signal variability. Garrett et al. (2010) show that the predictability of a noise measure was greatly improved by the extensive preprocessing of the data, beyond the conventional steps of realignment and normalization. Their methods included artifact correction via independent component analysis (Beckmann and Smith, 2004) and regressing out motion parameters. For future analysis of both PC and BOLD signal variability, this seems to be a fruitful approach. In the current study we used the relative value of noise differences between stimulus and rest periods, which minimizes the movement confound, as both periods should be equally affected by movement.

Critically, all assumptions on signal variability characteristics depend on the validity of our method to estimate the variability in the hMT+ signal. We used a generative model to estimate variability in the fMRI signal, modeling all eight events within a stimulus blocks separately as box-cars convolved with the HRF. The model furthermore accounted for the HRF-induced overlap of stimulus blocks and rest periods by assigning mixing values to each individual volume, based on the estimation of the relative contribution of stimulus and rest periods to the signal in that particular volume. Compared to other methods to assess variability in the BOLD signal, our method is quite complex. Garrett et al. (2010) for example directly calculated the SD over blocks. Considering that the physiological response in hMT+ to our stimulus periods probably consisted of a sustained elevation in BOLD signal, overlaid with single spikes evoked by the eight single events, simply calculating the SD would have not allowed us to separate the endogenous from the stimulus induced variability. The current model is

designed to account for the stimulus induced modulation of the BOLD signal, leaving us with the endogenous variability. Certain stimulus induced modulation of the BOLD signal might still not have been accounted for, such as repetition suppression which might occur due to repetitive stimulus display during a block. An alternative model taking this adaptation effect of the signal into account yielded similar results as our initial model. Extending our model to include an adaptation effect therefore seems to have little consequence for our measure of variability.

The model, in its current form, has also important limitations. By using non-uniformly distributed mixing parameters (e.g., the stimulation and rest periods), a bias is introduced as the maximum likelihood estimator assigns more of the variance in the data to the more frequent parameter (the stimulus period). Critically, though, this does not affect their use to compare subjects, as the bias will influence all subjects equally.

Another point to be considered is that the mixing parameter was calculated by convolving the stimulus events with the HRF, while the remaining signal was assumed to stem from the rest periods. Other ways to model the data are conceivable, e.g., convolving both rest and block-periods with the HRF. The ratio or sum of the two could then be used to model the data. Different modeling schemes remain to be explored systematically, to find which best estimates the contributions of the two distributions.

Taking the relative difference as a measure and not absolute variance, we are however confident that we observed physiological differences in hMT+ correlating with perceptional sensitivity. We conclude by suggesting that modeling variability difference between rest and stimulus cycles is a promising method to investigate physiological differences between subjects. We furthermore suggest that perceptual sensitivity in direction discrimination might be associated with noise characteristics in hMT+. This could ultimately help to understand normal and pathological changes in visual motion perception.

ACKNOWLEDGMENT

This work was supported by the Deutsche Forschungsgemeinschaft (GRK1091, JA1087/1-1) and the Neuroinformatics and Computational Neuroscience Doctoral Training Centre, School of Informatics, University of Edinburgh. The authors thank Stefan Glasauer for helpful input, all the participants, and the scanning assistants for their support.

REFERENCES

Acs, F., and Greenlee, M. W. (2008). Connectivity modulation of early visual processing areas during covert and overt tracking tasks. *Neuroimage* 41, 380–388.

Andersson, J. L., Hutton, C., Ashburner, J., Turner, R., and Friston, K. (2001). Modeling geometric deformations in EPI time series. *Neuroimage* 13, 903–919.

Bair, W., Zohary, E., and Newsome, W. T. (2001). Correlated firing in macaque visual area MT: time scales and relationship to behavior. *J. Neurosci.* 21, 1676–1697.

Beauchamp, M. S. (2003). Detection of eye movements from fMRI data. *Magn. Reson. Med.* 49, 376–380.

Beckmann, C. F., and Smith, S. M. (2004). Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE Trans. Med. Imaging* 23, 137–152.

Bennett, P. J., Sekuler, R., and Sekuler, A. B. (2007). The effects of aging on motion detection and direction identification. *Vision Res.* 47, 799–809.

Berman, R. A., and Colby, C. L. (2002). Auditory and visual attention modulate motion processing in area MT+. *Brain Res. Cogn. Brain Res.* 14, 64–74.

Billino, J., Bremmer, F., and Gegenfurtner, K. R. (2008). Differential aging of motion processing mechanisms: evidence against general perceptual decline. *Vision Res.* 48, 1254–1261.

Born, R. T., and Bradley, D. C. (2005). Structure and function of visual area MT. *Annu. Rev. Neurosci.* 28, 157–189.

Brainard, D. H. (1997). The psychophysics toolbox. *Spat. Vis.* 10, 433–436.

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

Brunel, N., and Wang, X. J. (2003). What determines the frequency of fast network oscillations with irregular neural discharges? I. Synaptic dynamics and excitation-inhibition balance. *J. Neurophysiol.* 90, 415–430.

Castelo-Branco, M., Formisano, E., Backes, W., Zanella, F., Neuenschwander, S., Singer, W., and Goebel, R. (2002). Activity patterns in human motion-sensitive areas depend on the interpretation of global motion. *Proc. Natl. Acad. Sci. U.S.A.* 99, 13914–13919.

Cohen, M. R., and Newsome, W. T. (2004). What electrical microstimulation has revealed about the neural basis of cognition. *Curr. Opin. Neurobiol.* 14, 169–177.

Corbetta, M., and Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215.

Dukelow, S. P., DeSouza, J. F., Culham, J. C., van den Berg, A. V., Menon, R. S., and Vilis, T. (2001). Distinguishing subregions of the human MT+ complex using visual fields and pursuit eye movements. *J. Neurophysiol.* 86, 1991–2000.

Dumoulin, S. O., Bittar, R. G., Kabani, N. J., Baker, C. L., Goualher, G. L., Pike, G. B., and Evans, A. C. (2000). A new anatomical landmark for reliable identification of human area V5/MT: a quantitative analysis of sulcal patterning. *Cereb. Cortex* 10, 454–463.

Emberson, L., Kitajo, K., and Ward, L. M. (2007). "Endogenous neural noise and stochastic resonance", in *Proceedings SPIE Noise and Fluctuations in Biological, Biophysical, and Biomedical Systems*, ed. S. M. Bezrukov, Florence, 6602.

Ermentrout, G. B., Galán, R. F., and Urban, N. N. (2008). Reliability, synchrony and noise. *Trends Neurosci.* 31, 428–434.

Faisal, A. A., Selen, L. P. J., and Wolpert, D. M. (2008). Noise in the nervous system. *Nat. Rev. Neurosci.* 9, 292–303.

Friston, K. J., Williams, S., Howard, R., Frackowiak, R. S., and Turner, R. (1996). Movement-related effects in fMRI time-series. *Magn. Reson. Med.* 35, 346–355.

Garrett, D. D., Kovacevic, N., McIntosh, A. R., and Grady, C. L. (2010). Blood oxygen level-dependent signal variability is more than just noise. *J. Neurosci.* 30, 4914–4921.

Ghosh, A., Rho, Y., McIntosh, A. R., Kötter, R., and Jirsa, V. K. (2008). Noise during rest enables the exploration of the brain's dynamic repertoire. *PLoS Comput. Biol.* 4, e1000196. doi: 10.1371/journal.pcbi.1000196

Halpern, S. D., Andrews, T. J., and Purves, D. (1999). Interindividual variation in human visual performance. *J. Cogn. Neurosci.* 11, 521–534.

Hinds, O. P., Rajendran, N., Polimeni, J. R., Augustinack, J. C., Wiggins, G., Wald, L. L., Rosas, H. D., Potthast, A., Schwartz, E. L., and Fischl, B. (2008). Accurate prediction of V1 location from cortical folds in a surface coordinate system. *Neuroimage* 39, 1585–1599.

Huk, A. C., Dougherty, R. F., and Heeger, D. J. (2002). Retinotopy and functional subdivision of human areas MT and MST. *J. Neurosci.* 22, 7195–7205.

Kamitani, Y., and Tong, F. (2006). Decoding seen and attended motion directions from activity in the human visual cortex. *Curr. Biol.* 16, 1096–1102.

Kanai, R., and Rees, G. (2011). The structural basis of inter-individual differences in human behaviour and cognition. *Nat. Rev. Neurosci.* 12, 231–242.

Krekelberg, B., van Wezel, R. J. A., and Albright, T. D. (2006). Adaptation in macaque MT reduces perceived speed and improves speed discrimination. *J. Neurophysiol.* 95, 255–270.

Liang, Z., Yang, Y., Li, G., Zhang, J., Wang, Y., Zhou, Y., and Leventhal, A. G. (2010). Aging affects the direction selectivity of MT cells in rhesus monkeys. *Neurobiol. Aging* 31, 863–873.

Liu, T., Hospadaruk, L., Zhu, D. C., and Gardner, J. L. (2011). Feature-specific attentional priority signals in human cortex. *J. Neurosci.* 31, 4484–4495.

Lund, T. E., Nørgaard, M. D., Rostrup, E., Rowe, J. B., and Paulson, O. B. (2005). Motion or activity: their role in intra- and inter-subject variation in fMRI. *Neuroimage* 26, 960–964.

MacDonald, S. W. S., Nyberg, L., and Bäckman, L. (2006). Intra-individual variability in behavior: links to brain

structure, neurotransmission and neuronal activity. *Trends Neurosci.* 29, 474–480.

Malikovic, A., Amunts, K., Schleicher, A., Mohlberg, H., Eickhoff, S. B., Wilms, M., Palomero-Gallagher, N., Armstrong, E., and Zilles, M. K. (2007). Cytoarchitectonic analysis of the human extrastriate cortex in the region of V5/MT+: a probabilistic, stereotaxic map of area hOc5. *Cereb. Cortex* 17, 562–574.

McDonnell, M. D., and Abbott, D. (2009). What is stochastic resonance? Definitions, misconceptions, debates, and its relevance to biology. *PLoS Comput. Biol.* 5, e1000348. doi: 10.1371/journal.pcbi.1000348

McIntosh, A. R., Kovacevic, N., and Itier, R. J. (2008). Increased brain signal variability accompanies lower behavioral variability in development. *PLoS Comput. Biol.* 4, e1000106. doi: 10.1371/journal.pcbi.1000106

Mennes, M., Zuo, X. N., Kelly, C., Martino, A. D., Zang, Y. F., Biswal, B., Castellanos, F. X., and Milham, M. P. (2011). Linking inter-individual differences in neural activation and behavior to intrinsic brain dynamics. *Neuroimage* 54, 2950–2959.

Mohr, P. N. C., and Nagel, I. E. (2010). Variability in brain activity as an individual difference measure in neuroscience? *J. Neurosci.* 30, 7755–7757.

Morrone, M. C., Tosetti, M., Montanaro, D., Fiorentini, A., Cioni, G., and Burr, D. C. (2000). A cortical area that responds specifically to optic flow, revealed by fMRI. *Nat. Neurosci.* 3, 1322–1328.

Muckli, L., Kriegeskorte, N., Lanfermann, H., Zanella, F. E., Singer, W., and Goebel, R. (2002). Apparent motion: event-related functional magnetic resonance imaging of perceptual switches and states. *J. Neurosci.* 22, RC219.

Muckli, L., Naumer, M. J., and Singer, W. (2009). Bilateral visual field maps in a patient with only one hemisphere. *Proc. Natl. Acad. Sci. U.S.A.* 106, 13034–13039.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113.

Pereira, F., Mitchell, T., and Botvinick, M. (2009). Machine learning classifiers and fMRI: a tutorial overview. *Neuroimage* 45, S199–S209.

Petersen, S. E., Baker, J. F., and Allman, J. M. (1985). Direction-specific adaptation in area MT of the owl monkey. *Brain Res.* 346, 146–150.

Purushothaman, G., and Bradley, D. C. (2005). Neural population code for fine perceptual decisions in area MT. *Nat. Neurosci.* 8, 99–106.

Rostrup, E., Law, I., Blinkenberg, M., Larsson, H. B., Born, A. P., Holm, S., and Paulson, O. B. (2000). Regional differences in the CBF and BOLD responses to hypercapnia: a combined PET and fMRI study. *Neuroimage* 11, 87–97.

Samanez-Larkin, G. R., Kuhnen, C. M., Yoo, D. J., and Knutson, B. (2010). Variability in nucleus accumbens activity mediates age-related suboptimal financial risk taking. *J. Neurosci.* 30, 1426–1434.

Slaghuis, W. L., and Ryan, J. F. (2006). Directional motion contrast sensitivity in developmental dyslexia. *Vision Res.* 46, 3291–3303.

Smith, A. T., Kosillo, P., and Williams, A. L. (2011). The confounding effect of response amplitude on MVPA performance measures. *Neuroimage* 56, 525–530.

Stein, J. (2001). The magnocellular theory of developmental dyslexia. *Dyslexia* 7, 12–36.

Stoppel, C. M., Boehler, C. N., Strumpf, H., Heinze, H. J., Noesselt, T., Hopf, J. M., and Schoenfeld, M. A. (2011). Feature-based attention modulates direction-selective hemodynamic activity within human MT. *Hum. Brain Mapp.* doi: 10.1002/hbm.21180. [Epub ahead of print].

Talcott, J. B., Hansen, P. C., Assoku, E. L., and Stein, J. F. (2000). Visual motion sensitivity in dyslexia: evidence for temporal and energy integration deficits. *Neuropsychologia* 38, 935–943.

Watson, A. B., and Pelli, D. G. (1983). QUEST: a Bayesian adaptive psychometric method. *Percept. Psychophys.* 33, 113–120.

Westheimer, G., and Wehrhahn, C. (1994). Discrimination of direction of motion in human vision. *J. Neurophysiol.* 71, 33–37.

Wichmann, F. A., and Hill, N. J. (2001a). The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Percept. Psychophys.* 63, 1314–1329.

Wichmann, F. A., and Hill, N. J. (2001b). The psychometric function: I. Fitting, sampling, and goodness of fit. *Percept. Psychophys.* 63, 1293–1313.

Winterer, G., Musso, F., Vucurevic, G., Stoeter, P., Konrad, A., Seker, B., Gallinat, J., Dahmen, N., and Weinberger, D. R. (2006). COMT genotype predicts BOLD signal and noise characteristics in prefrontal circuits. *Neuroimage* 32, 1722–1732.

Yang, Y., Liang, Z., Li, G., Wang, Y., and Zhou, Y. (2009). Aging affects response variability of V1 and MT neurons in rhesus monkeys. *Brain Res.* 1274, 21–27.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 17 January 2011; accepted: 21 July 2011; published online: 02 August 2011.
Citation: Wutte MG, Smith MT, Flanagin VL and Wolbers T (2011) Physiological signal variability in hMT+ reflects performance on a direction discrimination task. Front. Psychology ***2****:185. doi: 10.3389/fpsyg.2011.00185*
This article was submitted to Frontiers in Perception Science, a specialty of Frontiers in Psychology.