Clinical Neurophysiology 116 (2005) 172–189

www.elsevier.com/locate/clinph

EEG phase synchrony differences across visual perception conditions may depend on recording and analysis methods

Logan T. Trujillo*, Mary A. Peterson, Alfred W. Kaszniak, John J.B. Allen

Department of Psychology, University of Arizona, P.O. Box 210068, Tucson, AZ 85721-0068, USA

Accepted 22 July 2004
Available online 3 September 2004

Abstract

Objective: (1) To investigate the neural synchrony hypothesis by examining if there was more synchrony for upright than inverted Mooney faces, replicating a previous study; (2) to investigate whether inverted stimuli evoke neural synchrony by comparing them to a new scrambled control condition, less likely to produce face perception.

Methods: Multichannel EEG was recorded via nose reference while participants viewed upright, inverted, and scrambled Mooney face stimuli. Gamma-range spectral power and inter-electrode phase synchrony were calculated via a wavelet-based method for upright stimuli perceived as faces and inverted/scrambled stimuli perceived as non-faces.

Results: When the frequency of interest was selected from the upright condition exhibiting maximal spectral power responses (as in the previous study) greater phase synchrony was found in the upright than inverted/scrambled conditions. However, substantial synchrony was present in all conditions, suggesting that choosing the frequency of interest from the upright condition only may have been biased. In addition, artifacts related to nose reference contamination by micro-saccades were found to be differentially present across experimental conditions in the raw EEG. When frequency of interest was selected instead from each experimental condition and the data were transformed to a laplacian 'reference free' derivation, the between-condition phase synchrony differences disappeared. Spectral power differences were robust to the change in reference, but not the combined changes in reference and frequency selection criteria.

Conclusions: Synchrony differences between face/non-face perceptions depend upon frequency selection and recording reference. Optimal selection of these parameters abolishes differential synchrony between conditions.

Significance: Neural synchrony is present not just for face percepts for upright stimuli, but also for non-face percepts achieved for inverted/scrambled Mooney stimuli.

Keywords: Gamma band EEG activity; Neural synchrony; Visual cognition; Face perception; Mooney faces

1. Introduction

Empirical research in vision neuroscience has clearly demonstrated that visual stimulus features are processed at multiple spatially distributed cortical and subcortical brain regions. It has been hypothesized that neural synchronization in the gamma range (20–80 Hz) is the mechanism by which distributed features are integrated into unitary visual percepts (Singer and Gray, 1995; Varela, 1995; von der Malsburg and Singer, 1988).[1] Neural synchronization (or neural synchrony) refers to the phenomenon in which neurons coding for a common representation synchronize or 'phase lock' their (oscillatory) firing activity within a restricted frequency band. The synchronous oscillations have been

* Corresponding author. Tel.: +1-520-621-5543.
E-mail address: logant@u.arizona.edu (L.T. Trujillo).

[1] This definition of the gamma range follows Tallon-Baudry et al. (1998), and encompasses the range investigated by Rodriguez et al. (1999) (see Section 1.1). Other definitions place the lower range of the gamma band as being greater than 30 Hz (e.g. Fries et al., 2001).

1388-2457/$30.00 © 2004 International Federation of Clinical Neurophysiology. Published by Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.clinph.2004.07.025

found to occur in various frequency bands, although the strongest covariations with perception occur within the gamma range. This mechanism may operate whenever component processes subserved by spatially separate brain regions are integrated (e.g. Varela et al., 2001).

The neural synchrony hypothesis has been supported by a large number of multi-unit recording studies in animals (for review, see Singer, 1999; Singer and Gray, 1995). Synchronous firings have been observed to occur within and across cortical areas, hemispheres and sensory/motor modalities. This synchronous behavior can reflect perceptual gestalt criteria and performance. A smaller number of EEG/MEG studies have investigated neural synchrony in humans with most supporting a role for synchrony in neural integration (Singer, 1999; Varela et al., 2001). The presence of gamma-range synchrony has been shown to correlate with the perception of sound and linguistic stimuli (Miltner et al., 1999; Pantev, 1995; Ribary et al., 1991), as well as characterizing REM dream states (Llinas and Ribary, 1993). Gamma range activity has also been linked to attention (Fries et al., 2001; Tiitinen et al., 1993) and working memory (Tallon-Baudry et al., 1998). Gamma range synchrony has been found to accompany object recognition (Gruber and Muller, 2002; Tallon et al., 1995; Tallon-Baudry et al. 1996, 1997), covarying with the binding of visual elements into unitary percepts, although the magnitude of synchrony can be reduced with stimulus repetition (Gruber and Muller, 2002).

Some human studies have failed to support a role for synchrony in human perception and cognition. For example, Menon et al. (1996) found that gamma range synchrony is restricted to less than 2 cm regions across surface cortex. This result argues against functional long-range synchrony in human perception. This study only examined a 7 cm×7 cm region, however; EEG coherence has been found to drop at intermediate cortical distances, but then increase at long range distances (Nunez et al., 1997). Mima et al. (2001) found no between-condition differences for gamma range EEG coherence in response to black and white pictures of real and scrambled objects; although they found greater alpha range coherence for real versus scrambled stimuli. As recording and analysis methods have not been standardized, such differing results may arise from dissimilarities in methodology (e.g. Nunez et al., 1997).

We report an attempt to replicate an experiment reported by Rodriguez et al. (1999) demonstrating neural synchrony in human EEG when participants view upright versus inverted Mooney face stimuli. The pattern of neural synchrony we observed depended critically on our choice of which frequency bands to include in the grand average of the synchrony measures. When we used the criteria used by Rodriguez et al., we replicated their results. We came to believe that their criteria were not appropriate for our data set, however, and when we used a different criteria, we did not replicate their results. We found evidence for neural synchrony, although we observed a different pattern of synchrony than Rodriguez et al. In addition, our investigation revealed that both reference scheme and wavelet size must be carefully considered in gathering and analyzing EEG data for evidence of neural synchrony.

1.1. Rodriguez et al.'s study

Rodriguez et al. (1999) reported results that they took as evidence for the direct involvement of gamma-range synchronous oscillatory activity in human visual perception. Participants were shown 200 ms exposures of fragmented black and white shapes (Mooney, 1956) while EEGs were recorded. When visual closure occurs, these Mooney stimuli are perceived as faces. Face perception is much more likely to occur for upright than for inverted versions of these stimuli (compare samples in Fig. 1a and b). Rodriguez et al. hypothesized that neural synchrony would be more likely to occur when observers perceived faces than when they did not. To investigate this hypothesis, they calculated gamma range (20–60 Hz) *global spectral power* and *phase synchrony* measures (summed across trials, electrodes, and subjects) for the EEG data separately for trials where upright stimuli were perceived as faces (Up/F trials) and trials where inverted stimuli were not perceived as faces (Inv/NF trials) (approximately 70% of the trials in each of these conditions). A note on the types of synchrony measures presented by Rodriguez et al., follows.

Global spectral power measures the consequences of synchronous activity rather than synchronous activity itself. When neuronal populations are synchronized, the resultant constructive summation of their electrical fields yields measurable potentials at the scalp surface. The power of this activity suggests the degree to which synchronization obtains because weakly synchronized activity leads to destructive interference and low measurable power at a given frequency. Thus, spectral power is an indirect index of neural synchrony at best. Nevertheless, it has been used extensively to assess neural synchrony in human electrophysiological studies (for review see Tallon-Baudry and Bertrand, 1999).

Global phase synchrony measures phase synchronization more directly. For instance, in the method developed by Lachaux et al. (1999, 2000), signal phases are extracted by use of wavelet transforms and used to compute the average complex phase difference between two signals across trials.[2] These differences are formed into an index of the variation in phase synchrony between two electrodes across trials, a *phase locking value* (PLV), ranging from 0 (no synchronization) to 1 (perfect synchronization). It should

[2] An additional method to compute synchronization has been developed by Tass et al. (1998) which has been shown to be equivalent to the method of Lachaux and colleagues (Le Van Quyen et al., 2001). It should also be noted that synchronization as indexed in these two methods differs from the more common coherence method to assess frequency-dependent interelectrode correlations, as phase synchrony is independent of relative amplitude covariations while the coherence measure is not (Lachaux et al., 1999).

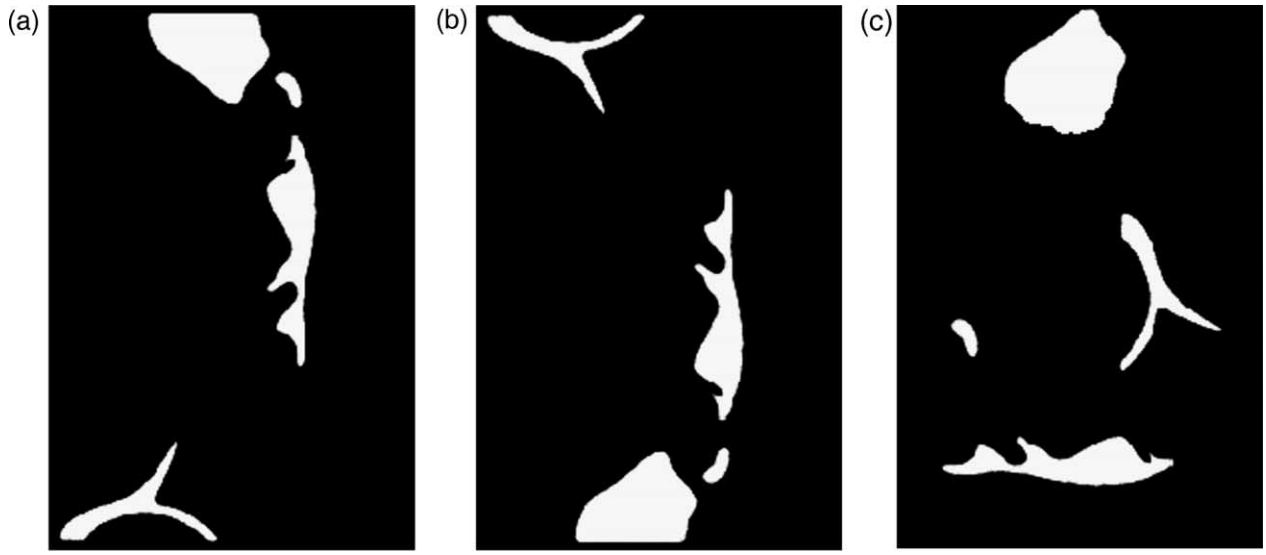

Fig. 1. (a) Example Mooney face stimulus. The image portrays the profile of a women facing to the right, illuminated from the upper right. (b) Inverted, and (c) scrambled versions of the stimulus in a.

be noted at the outset that noise tends to change extracted PLVs away from those representative of the embedded signal of interest, which can have a significant effect on phase synchrony computations (see Section 3).

Rodriguez et al. reported increased global gamma range spectral power, peaking at approximately 230 ms after stimulus onset on both Up/F and Inv/NF trials relative to a pre-stimulus baseline. In addition, they observed significantly greater gamma-band spectral power on Up/F trials than on Inv/NF trials. Gamma range global phase synchrony was found only for Up/F trials, where it reached a maximum 200–260 ms after stimulus onset. These data supported Rodriguez et al.'s hypothesis that neural synchrony would be more likely to occur when observers perceived faces than when they did not. They interpreted the synchrony increase on Up/F trials as reflecting the binding of visual object spatial features into a recognizable object (i.e. a face).

Rodriguez et al. also found a period of de-synchrony after the initial period of synchrony and before a second period of synchrony related to the motor response. This latter period of synchrony was present around the time of response in both experimental conditions and was interpreted as corresponding to the phase interactions underlying the creation and execution of motor representations and responses to the stimuli. The intervening period of decreased synchrony was taken to reflect the disengagement of cells from neural assemblies underlying the perceptual event to allow the creation of new assemblies for the motor act.

1.2. The present study

The present study sought to replicate these results by using presentation, recording, and analysis methods similar to those used by Rodriguez et al. (1999). We were motivated by a number of goals. First, the Rodriguez et al. experiment provided exciting support for the neural synchrony hypothesis, and we were interested in determining whether it could be replicated. Second, we were intrigued and somewhat puzzled by the finding that neural synchrony occurs only on Up/F trials and not on Inv/NF trials. On Inv/NF trials, observers perceived a relatively flat black and white pattern that at times resembled something familiar, albeit not a face. We expected that neural synchrony would reflect the binding together of the features of non-face percepts as well as of face percepts. The global spectral power measure provided indirect evidence for neural synchrony on Inv/NF trials in Rodriguez et al.'s experiment, but the global phase synchrony measure did not. We were interested in further exploring the differences between these types of trials. Third, Rodriguez et al.'s design required repetition of stimuli in both upright and inverted conditions. In pilot studies, we found that observers were aware when they were viewing an inverted version of a stimulus that they had perceived to be a face in the upright orientation. Although they did not attain a global percept of a face for these stimuli, they often recognized small span configurations that allowed them to know which face would be seen in an upright orientation. Because such knowledge could potentially contaminate responses, and hence, interpretation of the EEG patterns, we included scrambled Mooney face stimuli as a third stimulus condition in our experiment. In scrambled stimuli, the features of the upright stimuli were rearranged or 'scrambled,' thereby destroying both small span- and large span-configurations. The scrambled stimuli should not access any preexisting memory representations and thus should provide a more neutral comparison with respect to the upright faces.

2. Methods

2.1. Participants

A total of 12 undergraduates (six males and six females), recruited from the University's undergraduate subject pool participated in this experiment They were all right-handed, with a mean age of 19.3±0.3 years, and a mean education level of 13.3±0.4 years. All participants were fully informed of the recording technique, methods, and procedures before their consent to participate was obtained. The experimental protocol was reviewed and approved by the Institutional Review Board for Human Studies at the University of Arizona.

2.2. Stimuli and procedure

The stimuli consisted of 40 Mooney faces (Mooney, 1956) divided into two groups, one group with an original facing direction to the left (set A), and one group with an original facing direction to the right (set B) (see Fig. 1). In addition, 20 scrambled Mooney faces were generated from chosen members of the two stimulus sets, 10 from set A and 10 from set B (see Fig. 1). The scrambled stimuli were created by selecting various stimulus features from the upright stimulus, rotating them 45–90° in a commercial graphics program (Adobe PhotoShop, Adobe Systems Inc., San Jose, CA, USA), and then spatially rearranging the rotated parts. The rotations preserved the total number of horizontal and vertical contours in the transformation, thus ensuring that the spatial frequencies of the scrambled stimuli were approximately equal to those of the original stimuli. However, in some cases, changes were also made to the curvature relationships of the rotated stimulus features in order to minimize the spontaneous perception of a 'facelike' stimulus in the scrambled stimuli. There were no spatial frequency differences between the original and scrambled stimuli.[3]

Four versions of each stimulus were created by reflecting the stimuli around a vertical axis (creating a left-facing and a right-facing version of each stimulus), and then reflecting each of these stimuli around a horizontal axis (creating inverted versions of each of the left and right-facing stimuli). All stimuli were 7.2 cm wide×10.5 cm high. At the 60 cm viewing distance employed, the stimuli subtended a visual angle of approximately 7°×10°.

The experiment took place in a sound attenuated room. Stimulus materials were presented to the participants on a 19″ computer CRT screen. Stimuli were displayed against a white background within a black rectangular outline centered on the screen. The rectangle outlined the perimeter of the stimuli and was displayed throughout the experiment. Participants were instructed to fix their eyes within the center of the rectangle during each experimental block. Following a training block of 20 trials (five stimuli not included in the experimental set shown facing both left and right in upright and inverted orientations), participants received eight experimental blocks of 60 trials. Each block consisted of 20 upright, 20 inverted, 20 scrambled (10 upright, 10 inverted) stimuli. The stimuli were presented in a controlled random order such that each original Mooney and scrambled stimulus were presented only once (upright or inverted, original facing direction or mirrored facing direction) in each block. Each stimulus was presented once in each combination of orientation and facing direction over the first four blocks. The next four blocks were exact repetitions of the first four. Each upright, inverted, and scrambled stimulus was presented four times during the experiment, twice facing left, and twice facing right. Four different running orders were created that rotated the stimulus orientation of the initial presentation of each of the stimulus groups (A and B), ensuring that each group was presented across all four different ordering possibilities over the four running orders. Subjects were assigned to each running order in a sequential manner.

The stimuli were presented for 200 ms, with an inter-stimulus interval of 2000 ms.[4] The participants' task was to report whether or not they perceived a face, whatever its orientation, by pressing one of two response buttons mounted on joysticks held in each hand. Participants pressed the button in their dominant right hand if they perceived a face, and the button in the left hand if they did not perceive a face. They were instructed to report their perception at first glance as quickly as possible after they had determined their initial perception. The instructions stressed that they should press the 'face' button only if they perceived a complete face, not if they knew from features that were present that the stimulus would look like a face if they viewed it in a different orientation. Correct/incorrect responses and reaction time to stimuli were recorded by an IBM-compatible PC with a Pentium microprocessor and analyzed offline.

Our experimental procedure was the same as that used to collect the data analyzed by Rodriguez et al. (whose data were originally collected in an ERP study by George et al., 1997, see Section 3.3) except for the computer presentation of the stimuli and the inclusion of the scrambled faces.

[3] Possible spatial frequency differences for the scrambled stimuli were assessed by performing two-dimensional Fourier transforms on the original and scrambled stimuli. The resulting spatial frequency power values were summed across horizontal and vertical frequency ranges to yield summed values for each horizontal and vertical frequency. These two sets of values were then squared and summed, creating two global spatial frequency power statistics for horizontal and vertical frequencies. A T test performed on these statistics showed no overall differences in horizontal/vertical frequencies between the original and scrambled stimuli (horizontal: $t(19)=0.15$, ns; vertical: $t(19)=0.02$, ns).

[4] Our use of a constant ISI was not intentional but a consequence of a software error in the stimulus presentation program discovered after data collection. Since Rodriguez et al. used a variable ISI, it is possible, although unlikely, that any differences between our results and theirs are due in part to this discrepancy. This issue will be discussed in more detail in Section 4.

2.3. Behavioral data analysis

The proportions of 'face' responses made in the upright, inverted, and scrambled conditions, respectively, were calculated for each subject and then averaged across subjects Reaction times for both 'face' and 'no face' responses were calculated for all three conditions for each subject and then averaged within each condition across subjects. Reaction times <200 ms were not included in the averages. Between-condition differences were assessed via one-way repeated-measures analysis of variance (ANOVA) performed on the calculated proportions of face responses and the reaction time data. Since the ANOVA F-ratios had 2 degrees of freedom in the numerator, all effects were subjected to the Greenhouse–Geisser degrees of freedom correction for violations of sphericity when appropriate (Vasey and Thayer, 1987). Post hoc pairwise comparisons between conditions were calculated via within-subject student's t-tests for all significant outcomes.

2.4. EEG recordings

Thirty channels of EEG were recorded over the scalp. The recording apparatus consisted of 30 Ag/AgCl electrodes mounted in an electrode cap. Recording sites in the electrode cap included FZ, CZ, PZ, OZ, FC1, FC2, FC5, FC6, C3, C4, T7, T8, CP1, CP2, CP5, CP6, P3, P4, P7, P8, PO3, PO4, M1, M2, P9, P10, PO9, PO10, O9, O10. Recordings were made with respect to a nose reference, with a ground lead placed on the midline anterior to Fz. Vertical and horizontal blinks and eye movements were recorded from two pairs of Ag/AgCl bipolar recording leads, affixed to the outer canthi and also to the superior and inferior orbit of the left eye. All channels were amplified using a Grass model 12 neurodata system and were bandpass filtered between 0.01 and 100 Hz. The recording sampling rate was 1024 Hz. The data were later decimated offline to 512 Hz in order to reduce data analysis computational time. All electrode impedances were kept below 5 kΩ. All muscle artifacts were removed from the raw EEG record by visual inspection, and an automatic correction algorithm was implemented for the correction of eye movements and blinks (Neuroscan, Neurosoft Inc., Sterling VA, USA). Two second epochs ranging from -750 ms pre-stimulus to 1250 ms post-stimulus were extracted from the raw EEG record. In order to minimize the influence of edge-effects as a result of the temporal extension of the wavelet window (see Section 2.5), epochs were truncated after wavelet transformation to -500 ms prior to 1000 ms after stimulus onset.

2.5. Electrophysiological data analysis method of Rodriguez et al. (1999)

Following Rodriguez et al. (1999), analyses were restricted to epochs in which upright stimulus presentations were perceived as faces (Up/F), and inverted/scrambled presentations were not perceived as faces (Inv/NF and Scr/NF, respectively). As for Rodriguez et al., analyses of the other possible cases (such as upright presentations perceived as non-faces and inverted presentations perceived as faces) was not possible due to the small number of trials obtained for these conditions.

Gamma-range (20–60 Hz) spectral power and phase synchrony were computed by a wavelet-based time-frequency analysis (Lachaux et al., 1999, 2000) using Matlab computing software (The Math Works, Inc., Natick, MA, USA). Raw EEG data were bandpass filtered around a selected frequency of $f\pm2$ Hz (see information regarding frequency selection below), using a finite impulse response (length 300 ms) zero-phase shift time-domain filter. The filtered signal was convoluted offline by a complex Morlet wavelet with a Gaussian shape in both frequency and time (Mallat, 1999). Given the raw signal $S(t)$ and the wavelet $\Phi(t,f)$, the convolution of the signal and the wavelet is defined as

$$W(t,f) = \int_{-\infty}^{\infty} S(t)\Phi(t-t',f)\,\mathrm{d}t' \tag{1}$$

where $\Phi(t,f) = (\sigma_t^2\pi)^{-1/2}\exp(-t^2/2\sigma_t^2)\exp\{i2\pi ft\}$ and σ_t is the standard deviation in the time domain at the central frequency f. This wavelets $\Phi(t,f)$ are normalized such that their total energy is equal to unity. $W(t,f)$ is the wavelet transform of $S(t)$ and represents the portion of the signal that is oscillating at a given frequency at a given time in the time–frequency plane. The wavelet used in this method gives the best compromise in the necessary trade-off between time and frequency resolution (Mallat, 1999), with more precise frequency resolution at lower frequencies and better temporal resolution at higher frequencies. The wavelet standard deviations in the time and frequency domains are related by the uncertainty relationship $\sigma_f = 1/(2\pi\sigma_t)$ (Mallat, 1999), and together with the frequency form a wavelet family characterized by a constant ratio $\alpha = (f/\sigma_f)$ (Tallon-Baudry et al., 1997). The value of this ratio should be chosen to be greater than five (Grossman et al., 1989) to ensure convergence of the convolution integral in numerical approximations.

Following Lachaux et al. (1999) and Rodriguez et al. (1999), we chose a ratio parameter of $\alpha/2\pi = 7$, or $\alpha = 7\times2\pi \approx 44$. This choice yields a wavelet duration of $(2\sigma_t) = 700$ ms at frequency $f = 20$ Hz with a frequency spread of $(2\sigma_f) = 0.9$ Hz, while $f = 60$ Hz yields a duration of $(2\sigma_t) = 233$ ms with a frequency spread of $(2\sigma_f) = 1.4$ Hz.

A wavelet transform was calculated for each electrode and epoch followed by calculation of the real function $P(t,f) = |W(t,f)|^2$, which represents the spectral power of the signal. Rodriguez et al. calculated spectral power via a pseudo Wigner–Ville transformation. The latter has better time–frequency resolution than the wavelet transform, but is

known to be susceptible to spurious power distortions due to the presence of cross-term interference arising from the quadratic form of the transformation (Mallat, 1999). We chose the wavelet transformation as our spectral power measure, in part to avoid the difficulties associated with the Wigner transformation, but mainly because the wavelet transformation acts as the initial basis of the phase synchrony measure used by Rodriguez et al. Our primary motivation was to facilitate comparison between spectral power and phase synchrony responses. It is not clear what differences, if any, may manifest when comparing phase synchrony computed via one spectral measure with the power responses determined via a different spectral measure.

Wavelet transforms were performed on 2 s epochs (−750 ms prior to 1250 ms after stimulus onset) which were then truncated to 1500 ms (−500 ms prior to 1000 ms after stimulus onset) in order to minimize the influence of edge-effects as a result of the temporal extension of the wavelet window. The data were resampled by retaining every eighth data point in order to reduce computation time (a similar data decimation was used by Rodriguez et al.). Over the 1500 ms observation period, this procedure yielded a total of 96 data points for each electrode and frequency. The results of this transformation were then normalized with respect to a pre-stimulus baseline. For each electrode, an average power μ_{base} and standard deviation σ_{base} was computed over the 500 ms baseline period between −500 to 0 ms prior to stimulus onset (a total of 32 resampled data points).[5] Thus the normalized frequency power was calculated for each frequency band as

$$P(t,f)_{\text{norm}} = \frac{P(t,f) - \mu_{\text{base}}}{\sigma_{\text{base}}} \tag{2}$$

The normalized values were then averaged across electrodes and subjects to produce a grand average spectral power map of the global EEG responses over time for each condition.

Phase synchrony was determined by extracting the phase $\phi(t,f,n)$ of the convolution for each electrode, at all time points t, trials n [1,…,N], and frequencies of interest f. The degree of phase locking, called phase-locking value (PLV), between two electrodes at frequency f and time t was then calculated as

$$\text{PLV}_{jk}(t,f) = \frac{\left|\sum_{n=1}^{N} \exp(\text{i}\theta_{jk}(t,f,n)\right|}{N} \tag{3}$$

where $\theta_{jk}(t,f,n)$ is the phase difference $\phi_j(t,f,n) - \phi_k(t,f,n)$ between the electrode pair jk. As in the Rodriguez et al. (1999) study, the PLVs were normalized to a pre-stimulus baseline as was done for the power measure (Eq. (2)). The normalized values were then averaged across electrodes and participants to produce a grand average PLV map of the global EEG synchrony over time.

Following Rodriguez et al. (1999), a restricted global measure was used for the purpose of statistically testing and visualizing the time course of the spectral power and PLVs. For each of the upright, inverted, and scrambled conditions, the spectral power and PLVs associated with the frequency f_0 of each subject's maximal spectral power activity in the upright condition (averaged across electrodes) were extracted from the grand average spectral power and PLV maps for the entire pre- and post-stimulus time period. The maximal power activity was evaluated over the middle portion of the perception period (ranging from 150 to 250 ms).[6] These values were then averaged within each condition across all subjects to yield a single grand-averaged global measure of spectral power across time.

This aspect of Rodriguez et al.'s analysis method—restricting both the spectral power and the PLV analysis to analysis of the frequency at which maximum spectral power was observed in the upright condition—is problematic, as will be shown in the present paper. We discuss our problems with this aspect of the analysis after first reporting the results we obtained via the analysis methods used in the original study.

Time regions of significant global PLV and spectral power differences between conditions were indexed via three one-tailed Wilcoxon signed-rank T tests (one for each condition pair) performed at all time points exhibiting above random background synchrony (see below). These tests assessed the apriori hypothesis, established on the basis of the results of Rodriguez et al. (1999), that spectral power and phase synchrony are significantly greater during Up/F epochs than Inv/NF epochs. We extended the same prediction to Scr/NF epochs as to the Inv/NF epochs. This procedure differed slightly from that of Rodriguez et al. (1999); the earlier study only tested time points at which substantial peaks presented themselves qualitatively in the global PLV and spectral power waveforms. The present procedure also assessed the statistical significance of the differences between the Inv/NF and Scr/NF epochs. Performing a statistical test at each time point allows the assessment of the significance of synchrony

[5] Note that post-stimulus activation may have contaminated the pre-stimulus power values due to the temporal spread of the wavelet function. Such contaminants could artificially raise the baseline mean and variance and thus introduce error into the calculation of post-stimulus values. As a check, the normalization was also computed using baselines from −500 to 100 ms prior to stimulus onset so as to minimize the contamination by post-stimulus activity. These analyses yield an earlier apparent pre-stimulus activation timecourse (as expected from the temporal spread of the wavelet plus any expectation effects), but negligible differences otherwise. We report the baseline parameters of the original Rodriguez et al. study for ease of comparison between the two studies.

[6] Rodriguez et al. used the entire 1000 ms observation period to determine the maximal values of phase synchrony. The present study used the period between 150 and 250 ms because perception-related differences between conditions were of primary interest, and because of their data that showed a peak at 230 ms (see Section 4).

as it emerges from baseline behavior. Bonferroni corrections for multiple comparisons were not performed, as successive data points do not meet the requirements of statistical independence (due to the inherent spread of the wavelet). Instead, the total area under the global spectral power and PLV curves was calculated for each comparison between conditions that exhibited time periods of four or more significant sequential time points at the $P<0.05$ level. For 12 subjects, this significance level required the Wilcoxon test statistic to be at most $T=17$ or less (with the maximum allowable value of the statistic depending upon the number of zero between-condition differences among subjects for a given time point). A one-way ANOVA was computed between these area measures for each significant between-condition comparison (no corrections for sphericity violation were computed as these effects had fewer than 2 degrees of freedom in the F-ratio numerator).

To test the significance of measured PLVs from random background synchronization, bootstrapping techniques were used (Lachaux et al., 1999; Wasserman and Bockenholt, 1989). These techniques approximate the underlying statistical distribution of a data set when the distribution parameters are unknown and/or non-normal. For all distinct electrode pairs $\{(j,k)|\ j,k=1\ldots N;\ j\neq k\}$, a total of 200 new series of variables were created that had the same characteristics of the signal measured from electrode j, yet were independent of electrode k. This was achieved by randomly permuting or shuffling the trials recorded from electrode k to make a new series of trials $y'_k(n')$ = permutation of $\{y_k(1)\ldots y_k(N)\}$, where $y_k(n)$ is the signal from electrode k during trial n. The shuffles occurred across the original trial orderings as recorded, so that trials from different conditions were sometimes paired together. These series were then used to compute 200 PLV time series between all electrode pairs at the selected frequency for each subject. The 200 PLV time series were averaged across electrodes for each participant, and then across participants for each time series. This procedure yields an estimate of the sampling distribution of the grand-average global PLVs for each time point, representing the average randomly coincident or background phase synchrony across trials and conditions. For a given time point the 95% confidence interval of the corresponding distribution is an interval that contains $.95\times200=190$ of the grand-average PLVs centered at the distribution median. PLVs for each experimental condition that fell outside of these confidence limits were interpreted as being significantly different from the sampled synchrony baseline at the $P<0.05$ (two-tailed) level. The median and upper/lower 95% confidence limits of the bootstrapped distributions were plotted for each time point alongside the PLVs calculated for each experimental condition.[7]

[7] Note that this method of creating the shuffled bootstrapped distribution differs from that used by Rodriguez et al.; see Section 4.

3. Results

3.1. Behavioral results

An ANOVA performed on the percentage of face responses revealed a main effect of condition, $F(1.15,12.65)=58.17$, $P<0.001$ Post hoc tests showed that participants reported seeing faces more often when viewing upright ($72\pm4\%$) rather than inverted ($31\pm6\%$), or scrambled ($5\pm2\%$) Mooney stimuli, within-subjects $t(11)=11.67$ and $t(11)=19.67$, respectively, Ps <0.01 (two-tailed). In addition, they reported seeing faces significantly more often when they viewed inverted ($31\pm6\%$) rather than scrambled ($5\pm2\%$) displays, $t(11)=5.67$, $P<0.01$ (two-tailed).

An ANOVA performed on the reaction time data only revealed a trend for the main effect of condition, $F(1.30,14.30)=3.313$, $P<0.08$; (RTs for UP/F stimuli $=627\pm25$ ms; Inv/NF $=688\pm34$ ms; Scr/NF $=625\pm28$ ms).

No differences in either percentage of face reports or response latencies were found between the four stimulus-orientation running orders, as indicated by two 3×4 (condition by running order) ANOVA interactions (all Ps >0.80).

3.2. Results: Rodriguez et al.'s analysis method

Fig. 2a shows the global spectral power time course obtained using frequency selection criteria based upon the Up/F epochs and a wavelet parameter of $\alpha\approx44$ (the same value used by Rodriguez et al.). The spectral power (Fig. 2a) exhibited two local maxima over an average frequency band of 42 ± 3 Hz. The first peak occurred at approximately 300 ms post-stimulus onset and was contained within time periods where the Wilcoxon T test procedure (see Section 2) indicated that the spectral power was significantly greater during Up/F epochs than for either Inv/NF or Scr/NF epochs over time points spanning 16–422 and 0–469 ms, respectively. ANOVAs performed on the spectral power area measures during this time period confirmed these results: ($F(1,11)=13.81$; $P<0.003$ and F $(1,11)=8.44$; $P<0.01$ for Up/F versus Inv/NF and Scr/NF, respectively. The second peak was broader in morphology and spanned most of the interval from 500 to 1000 ms. No significant differences between conditions were found for time points in this interval.

These findings are in qualitative agreement with those of Rodriguez et al. (1999), who interpreted the first gamma-band peak as being a correlate of perceptual processes and the second peak as reflecting post-perceptual processes and/or the motor response. In our data, the two maxima were not cleanly separated from each other in time as in the original study. However, we did observe a clear separation of the two maxima when we used a wavelet transform with a smaller value of α ($\alpha=10$). We also noticed that the power magnitudes tended to be lower when using a small versus

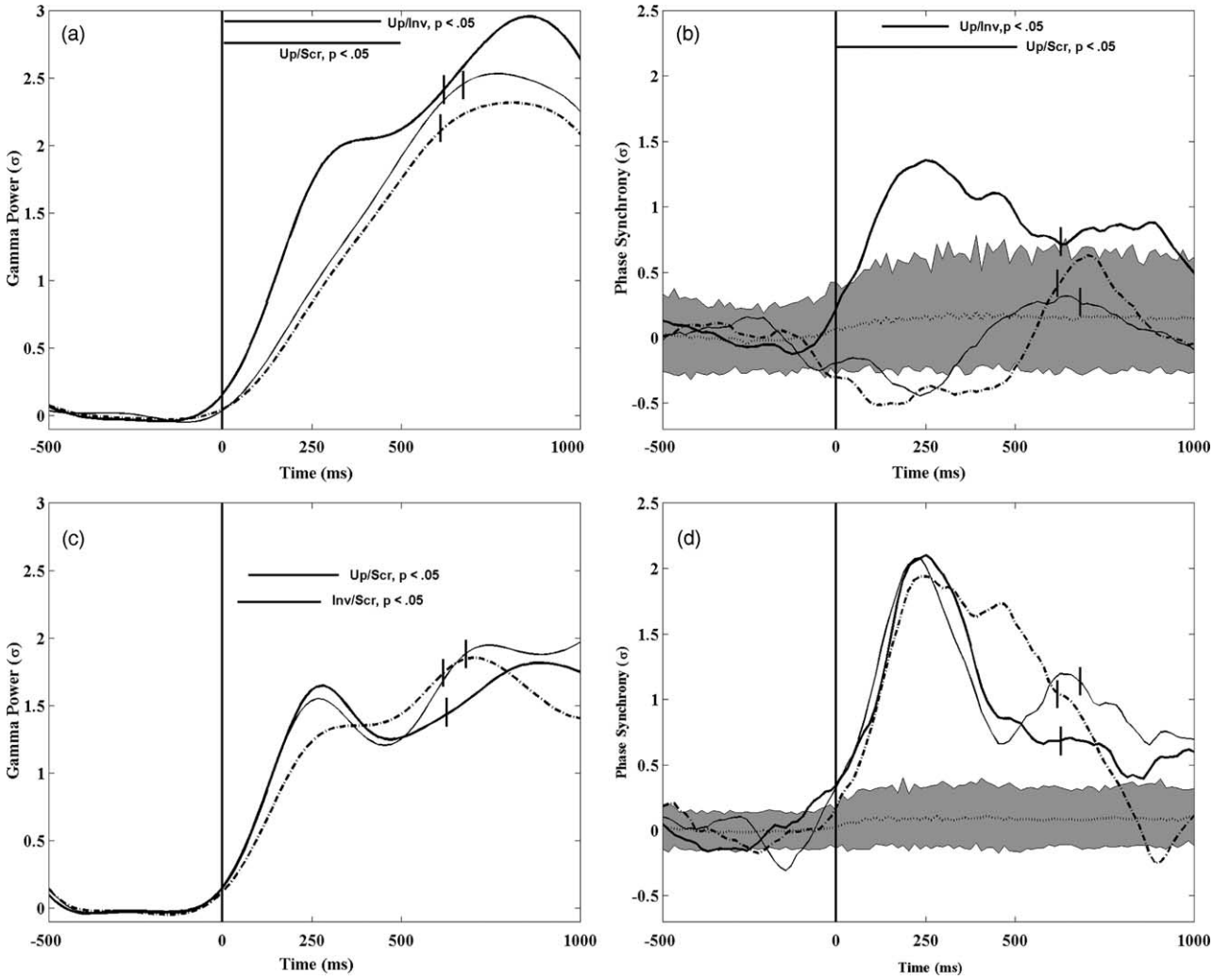

Fig. 2. Time courses for nose reference-based spectral power (a and c) and phase synchrony (b and d) indicated in standard deviations from the 500 ms pre-stimulus baseline. The time courses are computed for large parameter wavelets. All graphs are grand averages over electrodes, trials, and subjects in the Up/F (thick line), Inv/NF (thin line), and Scr/NF (dash-dot line) conditions. The dotted lines represent the median of the shuffled data distributions across time; the gray strips indicate the 95% confidence intervals. The horizontal lines represent the regions of significance ($P < 0.05$) for the indicated comparison conditions. The vertical bars intersecting the data curves indicate mean reaction time. (a) Nose reference-based spectral power grand averaged across each subject's maximum γ-band activity in the Up/F condition. (b) Nose reference-based phase synchrony grand-averaged across each subject's maximum γ-band activity in the Up/F condition. (c) Nose reference-based spectral power grand averaged across each subject's maximum γ-band activity *as determined independently for each condition*. (d) Nose reference-based synchrony grand averaged across each subject's maximum γ-band activity *as determined independently for each condition*. For all spectral power and synchrony graphs, the temporal spread of the wavelet plus expectation effects due to the constant interstimulus interval (ISI) account for results that obtain before stimulus onset.

large wavelet. This observation was confirmed by averaging the spectral power across conditions for each wavelet size and then performing a single Wilcoxon T test comparison between the perception period peak values (Wilcoxon $T(11) = 13.0$, $P = 0.05$, two-tailed). This suggests that large wavelets may be more sensitive to neural synchrony, while smaller wavelets may be more suited to revealing differences between processes that occur close together in time.

Following Rodriguez et al., we computed gamma-band phase synchrony on large parameter wavelets using the same frequency of interest as for the spectral power analysis (Fig. 2b). As in the original study, during the perception period, we observed a single peak for the Up/F epochs that was outside the confidence interval set by the baseline data. Synchrony for Up/F epochs reached a maximum at approximately 250 ms after stimulus onset. In contrast, for Inv/NF and Scr/NF epochs, synchrony decreased to a minimum during the perception period (Inv/NF ≈ 125 ms, Scr/NF ≈ 250 ms); the inverted response of Rodriguez et al. also displayed a similar decrease during the perception period. After the perception period, synchrony on the Up/F epochs decreased gradually, then maintained a slow increase for several hundred milliseconds before it decreased sharply immediately before the end of the epoch. This pattern is similar to that reported by Rodriguez et al. (1999), although

they had observed a period of significant desynchronization between the two peaks. Although we observed qualitative evidence for desynchronization during Up/F epochs, a Wilcoxon T test only revealed a trend towards significance, ($T(11)=19$, $P<0.07$).

The Wilcoxon T test procedure revealed that the synchrony was significantly greater for the Up/F epochs than for Inv/NF epochs over timepoints spanning 16–422 ms at the $P<0.05$ level. This difference was confirmed by ANOVA ($F(1,11)=9.65$; $P<0.01$). Synchrony was significantly greater during Up/F epochs than Scr/NF epochs from 0 to 469 ms ($F(1,11)=8.26$, $P<0.02$). No differences were found between the Scr and Inv conditions. There were no significant differences between any experimental conditions during the response period.

The shuffled baseline data displayed small oscillations throughout the epoch with limited variance (on the order of 0.1σ) (See Fig. 2). However, the confidence intervals spanned a range of approximately -0.35σ to 0.6σ, larger than reported by Rodriguez et al. This difference is most likely due to the different method of computing the shuffled baseline (see Section 2). Despite the larger confidence range, synchrony (or desynchrony) was observed beyond the 95% confidence intervals during the perception period in all three epochs types, suggesting that the findings were not due to random phase coincidences.

The overall results of these analyses essentially replicate those of Rodriguez et al. (1999). Nevertheless, during the course of data analysis, we uncovered two methodological problems that make interpretation of the present results difficult. The problems were (i) questionable criteria for relevant frequency selection, and (ii) possible reference electrode activity contamination of the EEG. Because of these problems, the between-condition differences in neural synchrony obtained above may not reflect differences in neural processing; instead, they may be artifacts of the methods of recording and analysis. These problems are discussed in more detail in the following two sections.

3.3. The influence of frequency band choice on computation of synchrony

In the course of our investigation, we discovered that a bias was introduced into the data by choosing our frequencies of interest based on each subject's maximal spectral power activity in Up/F epochs (averaged across electrodes) during the center portion of the perception period (ranging from 150 to 250 ms) for the final grand averages of both the spectral power and global synchrony measures. We have two reasons for making this claim: First, as discussed previously, spectral power is not necessarily a good index of phase synchrony (see also Lachaux et al., 1999; Rodriguez et al., 1999). It follows that different criteria should be used for choosing frequency band(s) of interest from spectral power and phase synchrony maps, respectively. Second, synchrony may be a dynamical phenomenon across both time and frequency. The individual and grand average maps computed from our data show qualitative evidence of synchrony across several bands simultaneously with variations in active bands from condition to condition (see Fig. 3). Functionally relevant phase synchrony might be present in one frequency band during face perception but in another band during non-face perception, perhaps reflecting the creation of separate large-scale networks associated with the two different perceptions (Varela, 1995; Varela et al., 2001). A more accurate assessment of synchrony variations across epoch types may be one in which frequencies of interest are chosen for each epoch type individually, and for the time period most relevant for the cognitive process investigated (e.g. perception and motor responses, for instance).

We recomputed global spectral power and phase synchrony using the frequency with the maximum synchrony in the 150–250 ms perception period for each epoch type individually (Fig. 2c and d) using large size wavelets. The spectral power (Fig. 2c) exhibited a local maxima for each epoch type (average frequencies, Up/F $=42\pm3$, Inv/NF $=40\pm2$, Scr/NF $=41\pm3$), with the first peak occurring at approximately 250–300 ms post-stimulus onset. Wilcoxon T tests indicated that the spectral power was significantly greater on Up/F epochs than on Scr/NF epochs (Up vs. Scr: 78-328 ms) which was confirmed by ANOVA ($F(1,11)=5.55$, $P<0.04$). Spectral power was also greater for Inv/NF epochs than for Scr/NF epochs (Inv vs. Scr: 47–266 ms; ($F(1,11)=5.48$, $P<0.04$). No power differences between the upright and inverted conditions were found.

Phase synchrony (Fig. 2d) was present for all epoch types, with the same morphology and amplitude (average frequencies, Up/F $=31\pm3$, Inv/NF $=32\pm2$, Scr/NF $=30\pm2$). The Wilcoxon T test procedure failed to reveal any differential synchrony across epoch type for any time point at the $P<0.05$ level. Shuffled baseline distributions were computed for each individual epoch type (using the individual maximum frequencies) and then collapsed across epoch types to yield an overall baseline distribution. The baseline data displayed small, limited variance oscillations, with smaller confidence intervals than found using Rodriguez et al.'s analysis method.

Thus, although perception-related synchrony may occur at different frequencies during different types of epochs, the amplitude of the synchronous activity is approximately equivalent across epoch type. This finding provides evidence against the hypothesis that synchronous activity is observed only when faces are perceived. Instead, synchrony is instantiated in different frequency bands when faces are perceived versus when they are not perceived.

3.4. EEG contamination by common reference electrode artifact

We discovered another problem using the original methods that, although not as important as the problem

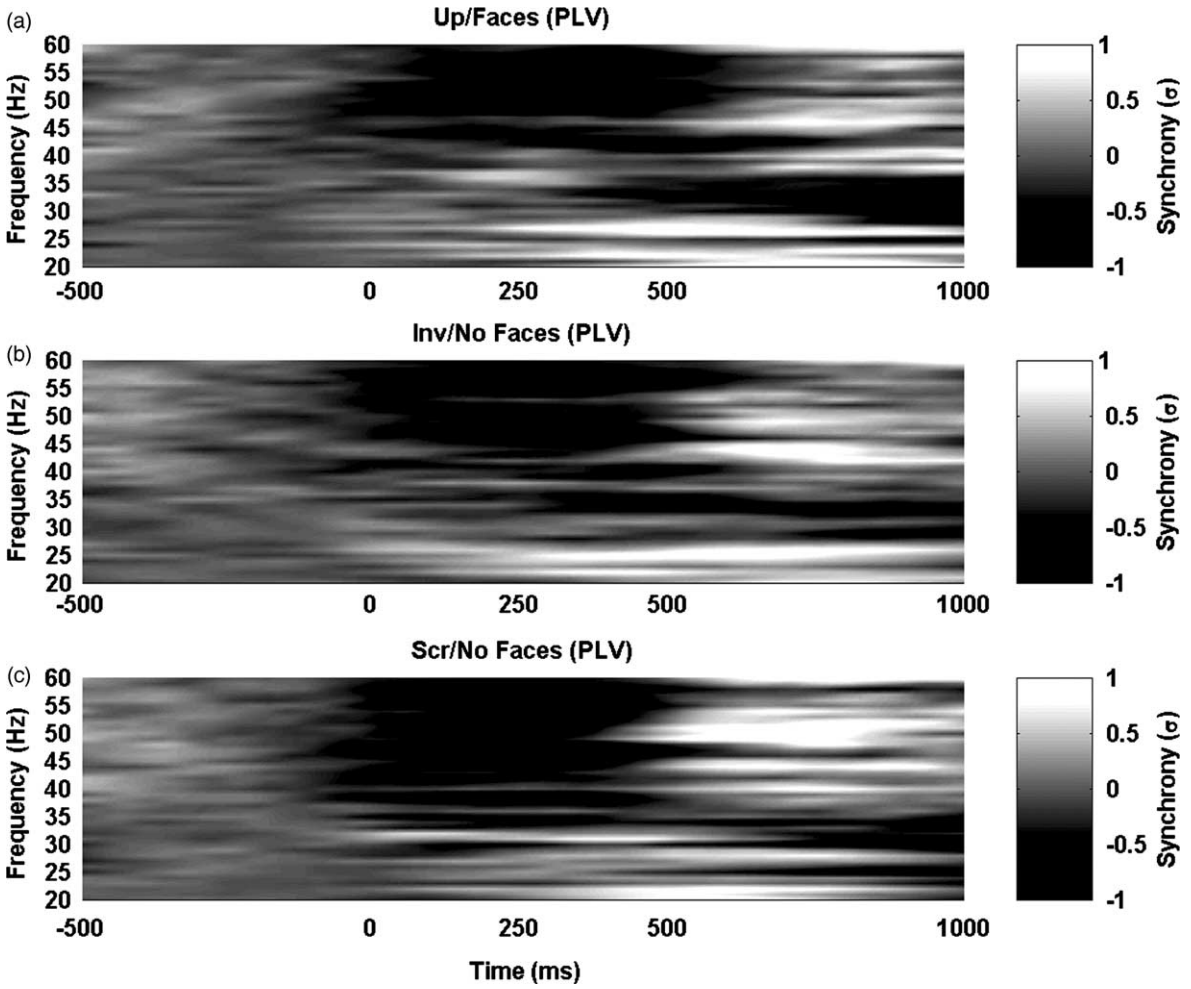

Fig. 3. Time–frequency map of phase synchrony averaged across electrodes, trials, and subjects in all three epoch types. Frequencies shown are in gamma range. Power values have been normalized with respect to 500 ms pre-stimulus baseline; gray color scale shows regions of increase (light) and decrease (dark) indicated in standard deviations from the baseline.

introduced by the choice of frequency band, is worth noting. The problem concerns contamination of EEG data obtained with a nose reference by eye movements that are not removed by standard means.

The data analyzed in Rodriguez et al. (1999) study were gathered in an earlier ERP study (George et al., 1997) that used a nose reference during recording. In order to fully replicate the Rodriguez et al. study, we used a nose reference in the recording procedures of the present study. During data acquisition, a persistent artifact waveform presented itself across all subjects. This artifact resembled a transient high frequency positive spike that was present across all electrode channels simultaneously (Fig. 4a). At first, it was hypothesized that the artifact may have been the result of facial electromyographic (EMG) activity in the vicinity of the nose area where the common reference electrode was located. However, deliberate movement of this facial area by participants failed to systematically produce the observed artifact. Instead, trial and error investigation determined that the artifact consistently occurred during abrupt, discontinuous changes in eye movements, as described next.

Since the common nose reference used in this study was located very close to the eyes, it was hypothesized that the artifact was the result of small, sharp saccadic eye movements recorded by the reference electrode. This hypothesis is consistent with the observed simultaneous multi-channel behavior in that any eye-movement related fluctuations in the signal recorded at the nose electrode would be transmitted to all electrodes to which it served as a reference (see Fig. 4a). It is thus possible that the synchrony observed in the present study might, at least in part, reflect eye movement related fluctuations present at all electrode sites rather than neural synchrony. Although large-scale eye movements had been removed from the present data via an automatic eye movement correction algorithm, these eye movement correction coefficients were based primarily upon eye blinks and not small-scale saccadic activity.

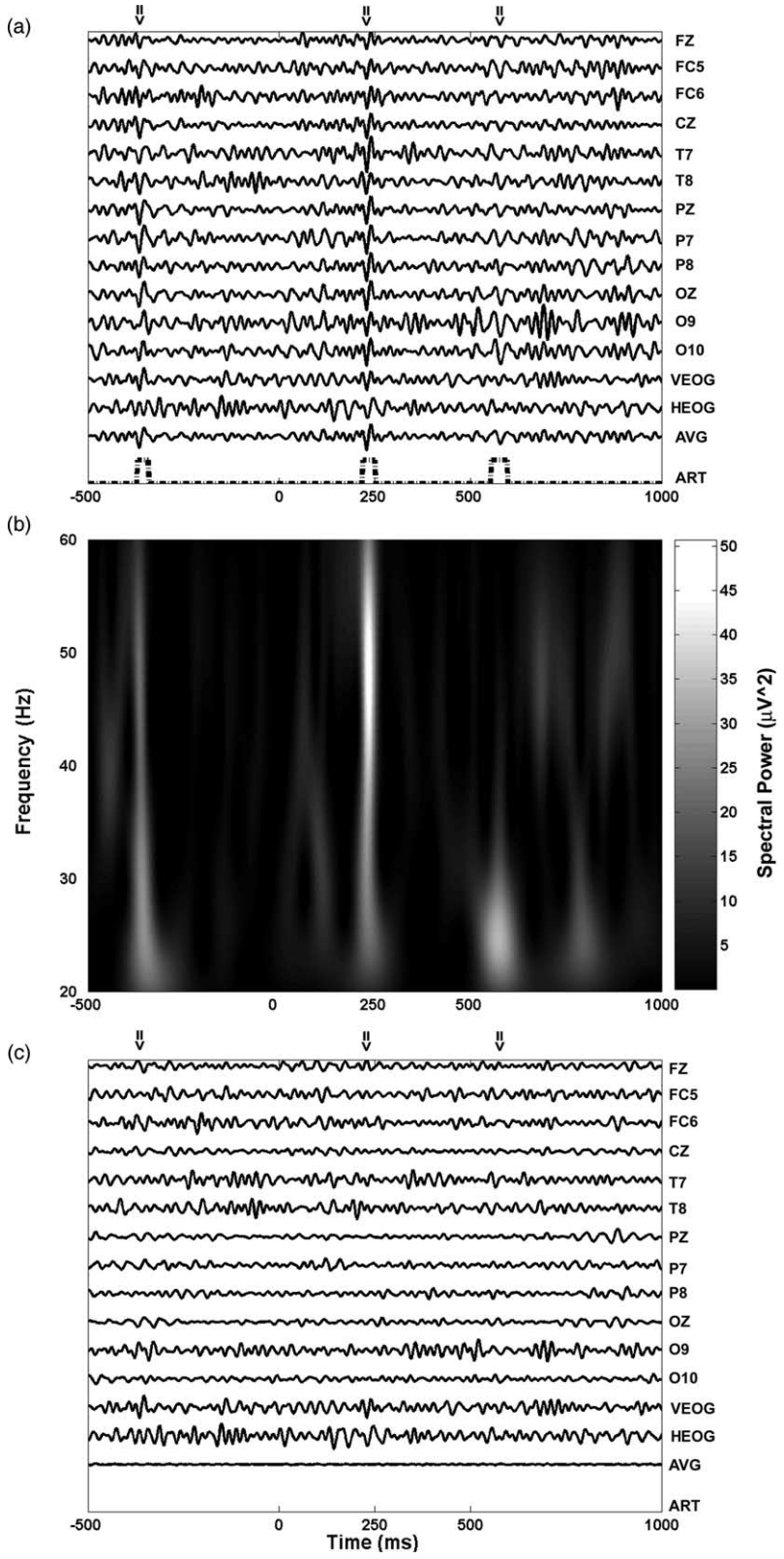

Fig. 4. (a) An example epoch from a single subject (12 representative signals out of the 30 recording electrodes are shown, plus EOG and artifact markers; AVG, averaged signal across electrodes; ART, artifact markers). The arrows at the top of the graph point to representative saccadic artifacts. Notice the strong similarity in waveform morphology across channels due to contamination of common reference activity. (b) The spectral power representation of the epoch shown in a. It is clear that periods of substantial power coincide with the across-electrode artifact. (c) The same data epoch after transformation to a laplacian reference scheme. Arrows show approximate locations of artifact in the untransformed epoch. Notice how the saccadic artifact is reduced or eliminated and the morphologies of the individual waveforms are now greatly differentiated from one another.

Since the morphology of eye blinks and saccades are not identical, saccadic activity was most likely not entirely eliminated from the data.[8] These artifacts might have influenced computed PLVs by decreasing the overall signal-to-noise ratio. In addition, any significant differential distribution of artifacts across conditions could have contributed to the differential synchrony we observed using the original analysis of Rodriguez et al.

There are reasons to believe that different amounts of saccadic activity might have been present across conditions. For instance, experiments by Henderson and Hollingworth (1999) and Yarbus (1967) show that different eye movement patterns are obtained when observers view faces versus scenes. Faces provide salient configurations that may lead to a greater degree of stereotyped saccadic activity than other stimuli.

To investigate this possibility, we devised a procedure to detect the presence of these artifacts and compute their distribution across the presentation period for each condition. Individual epochs were bandpass filtered between 20 and 60 Hz to restrict analysis to only those artifact signal frequency components that could contribute to the synchrony observed in the original analysis. The filtered signals were averaged across electrodes within each epoch. Since the artifacts occurred across all electrodes simultaneously, this averaging should lead to constructive summation of the artifacts and destructive summation of the remaining waveform. The latter was not perfect (Fig. 4a, second from bottom trace), but it was enough to allow sufficient isolation of the artifacts from the background signal average via wavelet transform for each epoch (Fig. 4b). A threshold was applied to each individual epoch spectral map identifying points greater than half of the maximum value of the map (across time and frequency). Time points having eight or more contiguous threshold indices along the frequency axis were assigned a value of unity. These values acted as artifact markers which were mapped against time to yield a measure of the temporal distribution of the central portion of the artifact over the epoch (see bottom trace, Fig. 4a). The criterion used in creating the artifact markers from the spectral maps required the artifact to have a significant presence along the frequency dimension; eight contiguous threshold indices span the wavelet frequency spread at 40 Hz ($2\sigma_f = 8$ Hz), the center of the frequency range examined in this study.

Each individual epoch distribution was divided into 50 ms bins from 0 to 300 ms. Any bin containing a nonzero time point was deemed contaminated by a substantial portion of the artifact. The across-trial frequency of contaminated bins was computed separately for each condition and participant, and was placed into a one-way chi-square analysis across conditions for each bin. Artifact counts were also collapsed across bins to yield a measure of the overall across-trial frequency of artifact occurring within the 300 ms perception period. Artifacts were counted for the 500 ms baseline period as a comparison. These frequency counts were subjected to one-way chi-square analyses.

We also assessed what effect the removal of reference electrode contamination from the EEG would have upon phase synchrony by transforming the data offline to a laplacian (or scalp current density [SCD]) reference scheme. This reference scheme is known to reduce the effects of physical reference activity and to reduce volume conduction (Law et al., 1993). The Laplacian reference scheme transforms the raw EEG potentials into a measure of the radial current density at the scalp (hence the alternative name, SCD). For this analysis, the surface laplacian was computed from the spherical harmonic expansion of the recorded scalp potentials via the method of Lagerlund et al. (1995). The spherical electrode coordinates for the electrodes in standard 10–20 positions were obtained from Pascual-Marqui et al. (1988). The coordinates of all nonstandard electrode positions were calculated from the 10–20 positions by interpolation.

Laplacian-derived data were also subjected to the artifact distribution analysis outlined above with the exception that the maxima values used for thresholding individual epochs were the same as those calculated for the corresponding nose-referenced epochs. This ensured that only Laplacian-derived trials containing substantial deviations in the across-electrode average (on the order of those found for the nose reference) would be included in the frequency counts.

3.5. Effects of reference artifact on synchrony

Fig. 4c shows an original nose-referenced epoch transformed to a Laplacian reference. The saccadic artifact was no longer present in the data, consistent with the hypothesis that the artifact in the raw data was due to contamination by reference electrode activity. Furthermore, the across electrode average of Laplacian transformed epochs showed limited magnitude and variance. The spectral power maps of these signals (not shown) failed to show any responses near the scale of activity found for the nose reference derived across electrode averages.

The artifact distribution analysis for the perception period for the nose reference revealed only one 50 ms

[8] A concern could be raised that the method of eye blink correction could introduce differences in to the synchrony computations, and these differences might account for the differences between our findings and those of Rodriguez et al. (1999). We do not know the method used by the earlier study; we used a built-in algorithm of the Neuroscan software (Compumedics, El Paso, TX, USA) which is based on a procedure due to Semlitsch et al. (1986). A theoretical analysis (Trujillo, in preparation) of the effects of this correction procedure on synchrony computation predicts that any differential synchrony introduced by eyeblink correction will vary as a function of the variability of eyeblinks across trials. We identified the presence of significant EOG activity in a given trial ($|EOG| > 75$ μV) and computed the frequency of eyeblinks across trials during the perception and baseline periods; the frequency of eyeblinks across trials did not differ between conditions for either perception ($\chi^2(2) = 0.16$, ns) or baseline periods ($\chi^2(2) = 1.12$, ns), suggesting that our results are not the result of the chosen eyeblink correction method.

time bin (200–250 ms) where the artifact counts were significantly different across conditions ($\chi^2(2)=9.94$, $P<0.01$). This bin was right in the middle of the perception period, in the time frame when greater neural synchrony was observed in Up/F epochs than in Inv/NF or Scr/NF epochs (as determined by Rodriguez et al.'s original analysis method). Indeed, artifact counts were greater during Up/F epochs ($23\pm2\%$) than either Inv/NF ($20\pm2\%$) or Scr/NF ($18\pm2\%$) epochs. Furthermore, when collapsed across epoch type, more artifact was present in the 300 ms perception period than the 500 ms baseline period ($\chi^2(2)=35.82$, $P<0.01$). The percentage of trials contaminated by post-stimulus eye movement artifacts was large ($\sim$90%) and significantly differed between the baseline and perception periods. The difference in pre-versus post-stimulus contamination was large enough ($\sim$10%) to potentially introduce significant noise into PLV computations and thus obscure any true synchrony present. In contrast to the nose reference, the analysis discovered no artifacts for any of the Laplacian time bins for any conditions. Since we used maximum threshold values determined by the nose-reference for the Laplacian analysis, this confirms that the artifact was removed by the laplacian transformation (Fig. 4c).

Fig. 5a and b shows the time course of gamma-band spectral power and phase synchrony (average frequency 49 ± 2 Hz) for the Laplacian reference scheme. The spectral power time course was attenuated, but similar in morphology to that found for the nose reference across all conditions (Fig. 5a). Significant spectral power differences were obtained using the large wavelets between the Up/F and the Inv/NF and Scr/NF epochs

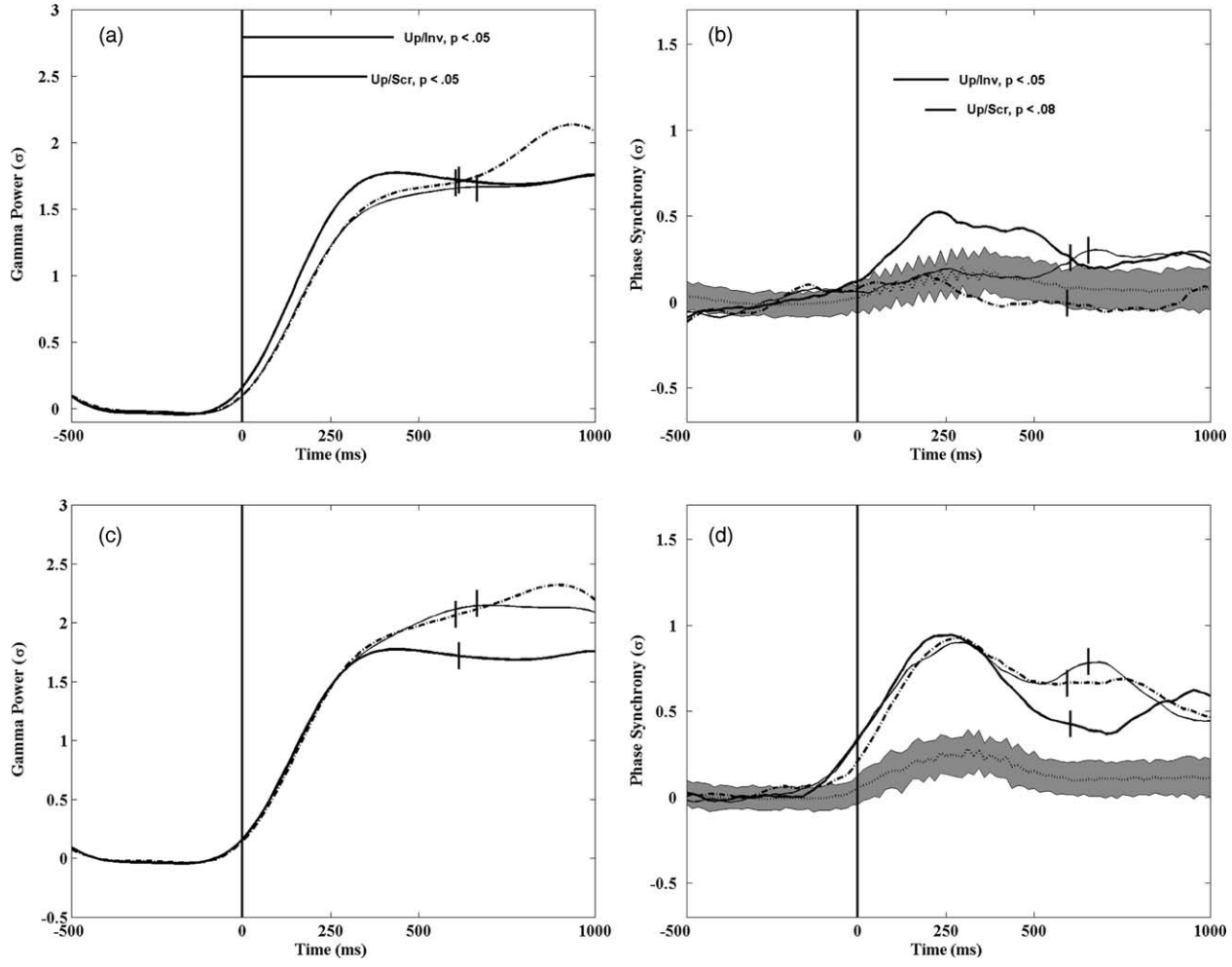

Fig. 5. Time courses for laplacian-based spectral power (a and c) and phase synchrony (b and d) indicated in standard deviations from the 500 ms pre-stimulus baseline. The time courses are computed for large parameter wavelets. All graphs are grand averages over electrodes, trials, and subjects in the Up/F (thick line), Inv/NF (thin line), and Scr/NF (dash–dot line) conditions. The dotted line represents the median of the shuffled data distribution across time; the *gray strip* indicates the 95% confidence intervals. The horizontal lines represent the regions of significance ($P<0.05$) for the indicated comparison conditions. The vertical bars intersecting the data curves indicate mean reaction time. (a) Laplacian-based spectral power grand averaged across each subject's maximum γ-band activity. (b) Laplacian-based phase synchrony grand-averaged across each subject's maximum γ-band activity. (c) Laplacian-based spectral power grand averaged across each subject's maximum γ-band activity *as determined independently for each condition*. (d) Laplacian-based synchrony grand averaged across each subject's maximum γ-band activity *as determined independently for each condition*. For all spectral power and synchrony graphs, the temporal spread of the wavelet plus expectation effects due to the constant interstimulus interval (ISI) account for results that obtain before stimulus onset.

(0–422 ms, Up/Inv: $F(1,11)=9.926$, $P<0.01$; and 0–328 ms, Up/Scr: $F(1,1)=9.45$, $P<0.01$). Phase synchrony exhibited a maximum at approximately 200 ms in the upright condition (Fig. 5b). Synchrony for the Inv/NF and Scr/NF conditions remained at nearly equal levels until deviating from one another at approximately 250 ms. Although neither the inverted nor scrambled responses deviated from the shuffled confidence intervals during the perception period, we computed Wilcoxon T test differences between the upright and latter conditions. The T test comparisons revealed a significant difference between the Up/F and Inv/NF epochs between 109 and 266 ms; this result was confirmed by ANOVA ($F(1,11)=4.77$, $P<0.05$). Surprisingly the sequential T tests only revealed a trend difference between the Up/F and Scr/NF conditions ($P<0.08$) for the period between 203 and 295 ms.

Shuffled data showed smaller confidence intervals for the laplacian reference (ranging from approximately -0.1σ to 0.3σ) than for the nose reference in the original Rodriguez et al. analysis but slightly more pronounced oscillations than the latter (Fig. 5b). To test whether these oscillations might reflect the amplification of some between-condition differences introduced by the laplacian transformation, we performed the Wilcoxon T test procedure on shuffled data created from individual epoch types (using the original frequency choice criteria of Rodriguez et al.). This analysis found no significant difference between any epoch types at any time points at the $P<0.05$ level.

As for the nose reference, we recomputed Laplacian-based global spectral power and phase synchrony using the frequency with the maximum synchrony in the 150–250 ms perception period for each epoch type individually (Fig. 5c and d). Spectral power and phase synchrony were present for all epoch types, with the same morphology and amplitude (average frequencies, power—Up/F = 49 ± 1, Inv/NF = 48 ± 2, Scr/NF = 47 ± 3; synchrony Up/F = 41 ± 3, Inv/NF = 41 ± 3, Scr/NF = 38 ± 3). A Wilcoxon T test procedure failed to reveal any significant differential power or synchrony across epoch type for any time point at the $P<0.05$ level. As for the analysis presented in Section 3.3, shuffled baseline distributions were computed for each individual epoch type (using the individual maximum frequencies). The baseline data displayed small, limited variance oscillations, on the order of that reported in section 3.3; however there was a slight increase in the magnitude of shuffled synchrony during the perception period (Fig. 5d). To test whether this increase might reflect the amplification of some between-condition differences introduced by the Laplacian transformation, we performed the Wilcoxon T test procedure on the shuffled data created from individual epoch types. Our analysis found no significant difference between any epoch types at any time points at the $P<0.05$ level.

To conclude this section: ocular-related electrode artifacts were identified in the raw nose referenced EEG data that were found to contaminate a large percentage of epochs across all three conditions. More artifact contamination was found in Up/F epochs than in the other epochs during a 200–250 ms post-stimulus perception interval. This time-period also contained the peak synchrony found for the Up/F epochs, with less or no synchrony found for other epochs. Application of a Laplacian transformation removed the reference related artifacts, and yielded a much smaller degree of synchrony relative to the nose-referenced data, although significant between-condition differences remained during the perception condition when Rodriguez et al.'s frequency selection criterion was used. Thus, although eye movements can contaminate synchrony measures computed based on a nose reference, they do not appear to account for all of the between condition differences reported by Rodriguez et al. Nevertheless, we have shown that synchrony measures based on a nose reference are subject to contamination by differential eye movements made in different epoch types. Furthermore, the between-condition differences were removed when analysis was based on our less-biased method of choosing the relevant frequencies of interest. This suggests that neural synchrony can be present across all perceptual conditions.

4. Discussion

4.1. Proper frequency choice for computation of global synchrony

The present study compared EEG phase synchrony responses to upright, inverted, and scrambled Mooney faces recognized as faces (upright) and non-faces (inverted/scrambled). Faces were seen on 72% of upright, 31% for inverted, and 5% for scrambled. Hence, it is clear that at least 31% of the time, inverted stimuli, like upright stimuli, engage processes that lead to face perception. We included scrambled stimuli in which all the large and small span configurations leading to face perception were destroyed to serve as a control that was less likely to engage face processing mechanisms. Our results obtained using the original synchrony analysis method of Rodriguez et al. produced results quite similar to those of the original study when comparable methods of analysis were employed. Larger spectral power and greater phase synchrony were present when faces were perceived (Up/F epochs) than when faces were not perceived (Inv/NF and Scr/NF epochs); the synchrony morphology of inverted and scrambled epochs were similar. Rodriguez et al. took their data as evidence that neural synchrony is greater when faces are perceived than when they are not perceived. Our initial results seemed to support this interpretation.

We discovered, however, that this pattern of results was observed in our data only when the summary spectral power and phase synchrony measures for all epoch types were based on the frequency at which maximum spectral power was observed during Up/F epochs. We created additional

summary analyses comparing the three epoch types, based on the frequency at which phase synchrony was largest for each of the three conditions independently for spectral power and phase synchrony. Spectral power differences were maintained between the Up/F and Scr/NF conditions and the Inv/NF and Scr/NF conditions, but not between the Up/F and Inv/NF conditions. In contrast, phase synchrony measures based on the maximum frequency for each synchrony measure for each epoch type were approximately equal for the different epoch types. Thus, widespread neural synchrony was observed for all perceptual conditions, not just when faces were perceived. Between-condition differences in neural synchrony may occur across different frequencies.

We restricted this analysis to a single maximal frequency for each condition; one could ask why even restrict analysis to one frequency only?[9] The analysis can be done on all frequencies within a chosen range (e.g. the gamma range). This also makes sense because important co-variations may be expressed across larger scales in the frequency space. There are drawbacks to such a process, however. A simple analysis would be to do two-dimensional statistical comparisons of the synchrony maps across all time and frequency ranges, perhaps by computing between-condition test maps (Duffy et al., 1981). However the size of the synchrony maps we used prohibit such an analysis due to the problem of multiple comparisons (41 frequencies $\times$ 96 time points = 3936 data points). The probability of type-I statistical error becomes very great when making such a large number of comparisons. If such an analysis were to be performed for each electrode pair individually, the number of comparisons would increase by a factor equal to the total number of distinct electrode pairs ($N_{pairs} = 435$ for $N = 30$ electrodes) to yield over 1.7×10^6 comparisons. Furthermore, nearby points on the time–frequency maps are not completely independent because of the smoothing effects of the wavelet transform; in this situation, ordinary Bonferroni corrections are overly conservative. This is analogous to the situation in PET and fMRI imaging where between-condition comparisons are made across densely sampled functional–anatomical maps of brain activation. Complex theoretical models have been developed to properly correct statistical comparisons of these data (for example, the approach developed by Friston et al., 1991). In the absence of such theoretical development for phase synchrony (a viable program for future research), a proper summary statistic can reduce the number of experimental comparisons while yielding useful information about synchrony dynamics across time and condition that represents the data in a meaningful way.

We believe that the summary measure used here provides a more accurate assessment of the across-condition distribution of synchrony present in our data than employed in the original analysis of Rodriguez et al. (1999). To be fair, we must state that Rodriguez et al. were guided in their frequency choice by a different trend in their data, as the frequency of maximum spectral power was approximately the same in their upright and inverted conditions (Eugenio Rodriguez, personal communication). It is a reasonable first step to examine the synchrony associated with particular spectral power responses, especially if the latter are observed to covary with the perceptual conditions. However, our data show that synchrony may manifest within different frequency bands than spectral power. As such, our analyses suggest a different interpretation of what role neural synchrony may play in cognition and perception. A prominent interpretation of the neural synchrony hypothesis (Varela, 1995; Varela et al., 2001; see also Section 1) has been that any cognitive or motor event should be associated by the creation of a dominant cell assembly distributed over many brain areas; the establishment of the assembly should then be associated with a large degree of global synchrony that 'binds' the distributed structure into a unified network. A simple hypothesis based in this interpretation would be that neural synchrony should be present in the case of face perception, but absent otherwise.

However this hypothesis is ambiguous with respect to what degree of synchrony should be present when multiple assemblies are competing for dominance, as the component cells/regions of each assembly should be synchronized with respect to each other. It can be argued that some degree of binding must occur on a local visual level when viewing scrambled Mooney faces because although no face-like pattern is usually detected, the black and white component shapes of the image are organized visually into figures and grounds, each of which is perceived as a perceptual 'whole'. If synchrony indexes binding, then some synchrony should be present during the perception of the scrambled stimuli. It is also not clear what degree of synchrony should obtain in the case of a weakly active perceptual representation. For example, inverted Mooney faces often appear as meaningless, although face perception may occur (face reports occurred on 31% of inverted trials in the present experiment; see Section 3). In addition to any possible synchrony for the perception of the stimulus components (as in the scrambled case), a weaker degree of synchrony may also be established as the result of a partial perception of the face-like (albeit inverted) patterns in the stimuli.

These considerations suggest that neural synchrony should be present when any stimulus is perceived, although perhaps to a differential extent in magnitude or topography. This is supported by the fact that our nose- and Laplacian-referenced data both showed equivalent amounts of synchrony for all perception conditions when the frequency of analysis was chosen independently across conditions (see Section 3). Furthermore this synchrony can be expressed across different frequencies depending upon condition (Fig. 5b and d). We are currently developing measures

[9] We thank an anonymous referee for raising this issue.

that attempt to quantify widespread synchronization (and perhaps desynchronization) across multiple frequencies as a function of condition (Trujillo et al., in preparation).

4.2. Correcting for reference electrode contamination

We also found that eye movement artifacts can contaminate nose-referenced data. Moreover, those artifacts were more numerous in conditions where neural synchrony was larger. We did not have enough EEG amplifiers at the time of experiment to record signals from other physical reference points that could be used to re-reference the data offline, and so are unable to make any statement about the effect of other physical reference schemes. We chose not to re-reference to scalp sites (such as CZ) because they may be highly active during a cognitive/perceptual event, especially the standard mastoid references which are close to visual regions in nearby ventral temporal cortex. Average references were avoided because they can be highly active depending on electrode configuration (Bertrand et al., 1985) and brain activation with a magnitude that is impossible to predict (Rappelsberger, 1990)[10]; hence we were lead to choose the Laplacian reference.

The number of recording electrodes ($N=30$) used in this study was at the low end of the accepted range for minimal error produced in the laplacian transformation (Law et al., 1993). It is possible that the Laplacian transformation used here may have introduced error into the synchrony computation as has been found for the calculation of interelectrode coherence (Rappelsberger, 1990). Lachaux et al. (1999), using data recorded from a larger number of intracranial electrodes ($N=32$), found that an SCD reference computed via spline interpolations (Pernier et al., 1988) was inconsistent in its effects on synchrony. Overall, the procedure reduced the magnitude and sharpened the spatial topography of synchrony (as compared to an infinite reference), but occasionally it increased synchrony. It should be noted that other studies have found the spline-based SCD algorithm of Pernier et al. (1988) to produce artificial increases in interelectrode coherence (Fein et al., 1988; Biggins and Fein, 1993; Biggins et al., 1992), although this result has been contested (Nunez et al., 1997; Perrin, 1992; Pascual-Marqui, 1993); so the synchrony distortions reported by Lachaux et al. (1999) might have been related to the SCD algorithm used. The present study used a spherical harmonic approach to SCD calculation that has been shown to be more resistant to artificial coherence inflation or reduction and to provide good estimates of the source current topography (Lagerlund et al., 1995).

Of course, we cannot claim that any reference-related artifact was present in Rodriguez et al.'s data; nose reference contamination will vary with many factors, including electrode and recording characteristics, the position on the nose, individual tissue conductance differences between participants, and ocular behavior of individual participants. However, our results do suggest that, when there are reasons to suspect that different eye movements might be made in different conditions, it would be wise to use another, quieter reference scheme when looking for between condition differences in neural synchrony.

4.3. Differences between our procedure and Rodriguez et al.'s

There were methodological differences between our study and Rodriguez et al.'s that may account for some of the differences in results obtained here. First, we used a constant ISI between stimulus presentations, whereas Rodriguez et al. used a variable ISI. The constant ISI might have led to expectancy effects. Indeed, our global measures do show some evidence of pre-stimulus activity, which may reflect expectation effects. It is unlikely that expectancy might have caused cognitive operations leading to perception of a face to occur earlier in time; the perception interval peak latencies of spectral power ($\sim$300 ms) and phase synchrony ($\sim$250 ms) found using the original Rodriguez et al. procedure are in the order of, or are slightly longer than, those reported in the original study ($\sim$230–250 ms for spectral power and phase synchrony). The change in pre-stimulus activity might have affected subsequent normalization of the PLVs, perhaps accounting for the absence of significant desynchronization between perception and motor responses in the present study and/or the large discrepancy between our effect sizes (present max synchrony $\sim 2\sigma$; Rodriguez et al. ~ 4–10σ).[11] However, this effect size difference could also reflect the fact that our stimuli were presented on a computer screen against a white luminous background, while the stimuli of the original George et al. (1997) paper were presented from slides. The presence of the luminous background might have elevated pre-stimulus baseline activity, and thus lowered any differential responses between the pre-and post-stimulus periods.

We believe that these differences do not detract from our basic conclusions presented in this paper. Although the use of a constant ISI and white luminous presentation background may have reduced overall synchrony magnitudes relative to those reported by Rodriguez et al., we still found significant deviations from baseline synchrony, as well as finding between-condition differences similar to those found by the original study when using their original reference scheme and analysis methods. This suggests that our study is directly comparable to theirs despite these procedural disparities.

Our method of creating the shuffled bootstrap distributions also differed from the procedures used by Rodriguez

[10] We thank all three anonymous referees for making us aware of the difficulties when using the average reference to study EEG correlations across the scalp.

[11] We thank an anonymous referee for raising this possibility.

et al. and Lachaux et al. (1999). We computed a separate distribution from shuffled data for each individual time point, and then used these distribution's 95% confidence intervals to ascertain statistical significance of measured PLVs from random background synchronization. The original procedure of Lachaux et al. (1999) retained the maximal PLVs of each permuted time series across the entire epoch to create a distribution of 200 values which were then used to statistically test measured synchrony at each time point. Rodriguez et al.'s method of creating the grand-average shuffled plots differed in that they averaged across shuffles and participants, and computed confidence limits of ± 3 standard errors of the across-participant distribution for each time point. While this method yields information about the variability of shuffled PLVs across subjects, it is not clear if it can be used as a correct statistical test and it was not used as such by Rodriguez et al. for their global synchrony data. Our method is based in an established technique (Wasserman and Bockenholt, 1989) and provides a correct statistical test for above baseline synchrony while simultaneously allowing visualization of the computed baselines and their confidence limits alongside with the experimental values.

Finally, we chose our frequencies of interest based on activity within the perception period; Rodriguez et al. considered the whole post-stimulus period. It is clear that choice of another time-period, for example during the motor response or the entire observation window, could possibly yield different maximal PLVs and therefore different between-condition results. However, it is unlikely that the choice of the entire window would have produced results similar to those obtained by Rodriguez et al. unless they were an artifact of adding in synchrony during the response window rather than the perception window. Hence we believe that our frequency choice procedure is more appropriate for testing perception related differences in phase synchrony, which was the goal of the present study.

4.4. Guidelines for future research

It is apparent that correction of the methodological issues revealed by various analytical approaches changes the synchrony observed in our data; these changes then lead to alternative interpretations of the data than that suggested by Rodriguez et al. (1999). To prevent these issues from arising in future research, the following guidelines are proposed for designing and analyzing studies using measures of EEG synchrony:

(1) Synchrony may be present at different times and frequencies across conditions; thus frequencies of interest must be selected by objective criteria that do not favor one condition over another.
(2) Scalp potentials should be recorded with respect to quiet references and/or transformed into a 'reference-free' derivation in order to minimize contamination of the data with reference electrode activity; the use of a nose reference should be avoided.
(3) Wavelet sizes should be chosen appropriately for the timing characteristics of the stimulus-response sequence. If possible, experiments should be designed so that the temporal spacing between successive perceptual, cognitive, and/or motor events minimizes any temporal blurring of spectral power and phase synchrony due to the extended nature of the wavelet.

Proper application of these guidelines should serve to reduce error and increase reliability in EEG synchrony studies.

Note added in proof

After this paper was sent for publication we discovered a published method (Burgess and Gruzelier, 1999) designed to correct for multiple statistical comparisons in topographical mapping of EEG and event-related desynchronization data. This method has been successfully used to detect between-condition spectral power differences across time and frequency during word discrimination under conditions of backward masking (Summerfield et al., 2002). We are currently investigating this method's usefulness for correcting two-dimensional statistical analyses of phase synchrony across time and frequency, which may provide a statistically principled means to select relevant frequencies of interest (see Section 4.1).

Acknowledgements

We thank Eugenio Rodriguez, Jean-Philippe Lachaux, and the late Francisco J. Varela for providing invaluable advice concerning how to measure neural synchrony and for providing Matlab software scripts that served as the basis for the time-frequency and synchrony algorithms used to analyze our data. We also thank three anonymous referees for their valuable comments on an earlier version of this paper. This research was supported by a US Department of Education Jacob K. Javits Fellowship to LTT and by a National Science Foundation grant (BNS 99-06063) to MAP.

References

Biggins CA, Fein G. Spline computation of scalp current density and coherence: a reply to Pascual-Marqui (letter to the editor). Electroenceph Clin Neurophysiol 1993;87:65–6.

Biggins CA, Ezekiel F, Fein G. Spline computation of scalp current density and coherence: a reply to Perrin (letter to the editor). Electroenceph Clin Neurophysiol 1992;83:172–4.

Bertrand O, Perrin F, Pernier J. A theoretical justification of the average reference in topographic evoked potential studies. Electroenceph Clin Neurophysiol 1985;62:462–4.

Burgess AP, Gruzelier, JH. Methodological advances in the analysis of event-related desynchronization data: reliability and robust analysis. In: Pfurtscheller G, Lopes de Silva FH, editors. Handbook of Electroencephalography and Clinical Neurophysiology, Revised Series, Vol. 6. Amsterdam: Elsevier; 1999. p. 139–58.

Duffy FH, Bartels PH, Burchfiel JL. Significance probability mapping: an aid in the topographic analysis of brain electrical activity. Electroenceph Clin Neurophysiol 1981;51:455–62.

Fein G, Raz J, Brown FF, Merrin EL. Common reference coherence data are confounded by power and phase effects. Electroenceph Clin Neurophysiol 1988;69:581–4.

Fries P, Reynolds JH, Rorie AE, Desimone R. Modulation of oscillatory neuronal synchronization by selective attention. Science 2001;291:1560–3.

Friston KJ, Frith CD, Liddle PF, Frackowiak RSJ. Comparing functional (PET) images: the assessment of significant change. J Cereb Blood Flow Metab 1991;11:690–9.

George N, Jemel B, Fiori N, Renault B. Face and shape repetition effects in humans: a spatio-temporal ERP study. Neuroreport 1997;8:1417–23.

Grossman A, Kronland-Martinet R, Morlet J. Reading and understanding continuous wavelet transforms. In: Combes JM, Groomsman A, Tchamitchian P, editors. Wavelets, time–frequency methods, and phase space. Berlin: Springer; 1989. p. 2–20.

Henderson JM, Hollingworth A. High-level scene perception. Annu Rev Psychol 1999;50:243–71.

Lachaux JP, Rodriguez E, Martinerie J, Varela FJ. Measuring phase synchrony in brain signals. Hum Brain Mapp 1999;8:194–208.

Lachaux JP, Rodriguez E, Le Van Quyen M, Lutz A, Martinerie J, Varela FJ. Studying single-trials of phase-synchronous activity in the brain. Int J Bifurc Chaos 2000;10:2429–41.

Lagerlund TD, Sharbrough FW, Busacer NE, Cicora KM. Interelectrode coherences from nearest-neighbor and spherical harmonic expansion computation of laplacian of scalp potential. Electroenceph Clin Neurophysiol 1995;95:178–88.

Law SK, Nunez PL, Wijesinghe RS. High resolution EEG using spline generated surface laplacians on spherical and ellipsoidal surfaces. IEEE Trans Biomed Eng 1993;40:145–53.

Le Van Quyen M, Foucher J, Lachaux JP, Rodriguez E, Lutz A, Martinerie J, Varela FJ. Comparison of Hilbert transform and wavelet methods for the analysis of neuronal synchrony. J Neurosci Methods 2001;111:83–98.

Llinas R, Ribary U. Coherent 40-hs oscillation characterizes dream state in humans. Proc Natl Acad Sci USA 1993;90:2078–81.

Mallat S. A wavelet tour of signal processing. London: Academic Press; 1999.

Menon V, Freeman WJ, Cutillo BA, Desmond JE, Ward MF, Bressler SL, Laxer KD, Barbaro N, Gevins AS. Spatio-temporal correlations in human gamma band electrocorticograms. Electroenceph Clin Neurophysiol 1996;98:89–102.

Miltner WHR, Braun C, Arnold M, Witte H, Taub E. Coherence of gamma-band EEG activity as a basis for associative learning. Nature 1999;397:434–6.

Mima T, Oluwatimilehin T, Hiraoka T, Hallet M. Transient interhemispheric neuronal synchrony correlates with object recognition. J Neurosci 2001;21:3942–8.

Mooney CM. Closure with negative after-images under flickering light. Can J Psychol 1956;10:191–9.

Nunez PL, Srinivasan R, Westdorp AF, Wijesinghe RS, Tucker DM, Silberstein RB, Cadusch PJ. EEG coherency I: statistics, reference electrode, volume conduction, laplacians, cortical imaging, and interpretation at multiple scales. Electroenceph Clin Neurophysiol 1997;103:499–515.

Pantev C. Evoked and induced gamma band activity of the human cortex. Brain Topogr 1995;7:321–30.

Pascual-Marqui RD. The spherical spline laplacian does not produce artifactually high coherences: comments on two articles by Biggins, et al. (letter to the editor). Electroenceph Clin Neurophysiol 1993;87:62–4.

Pascual-Marqui RD, Gonzalez-Andino SL, Valdes-Sosa PA, Biscay-Lirio R. Current source density estimation and interpolation based on the spherical harmonic fourier expansion. Int J Neurosci 1988;43: 237–49.

Pernier J, Perrin F, Bertrand O. Scalp current densities: concept and properties. Electroenceph Clin Neurophyiol 1988;69:385–9.

Perrin F. Comments on article by Biggins et al. (letter to the editor). Electroenceph Clin Neurophyiol 1992;83:171–2.

Rappelsberger P. The reference problem and mapping of coherence: a simulation study. Brain Topogr 1990;2:63–72.

Ribary U, Ioannides AA, Singh KD, Hasson R, Bolton JPR, Lado F, Mogilner A, Llinas R. Magnetic field topography of coherent thalmocortical 40-hz oscillations in humans. Proc Natl Acad Sci USA 1991;88:11037–41.

Rodriguez E, George N, Lachaux JP, Martinerie J, Renault B, Varela FJ. Perception's shadow: long-distance synchronization of human brain activity. Nature 1999;397:430–3.

Semlitsch HV, Anderer P, Schuster P, Presslich O. A solution for reliable and valid reduction of ocular artifacts applied to the P300 ERP. Psychophysiology 1986;23:695–703.

Singer W. Neural synchrony: a versatile code for the definition of relations? Neuron 1999;26:49–65.

Singer W, Gray CM. Visual feature integration and the temporal correlation hypothesis. Annu Rev Neurosci 1995;18:555–86.

Summerfield C, Jack AI, Burgess AP. Induced gamma activity is associated with conscious awareness of pattern masked nouns. Int J Neurosci 1999;44:93–100.

Tallon-Baudry C, Bertrand O. Oscillatory gamma activity in humans and its role in object representation. Trends Cogn Sci 1999;3:151–62.

Tallon C, Bertrand O, Bouchet P, Pernier J. Gamma-range activity evoked by coherent visual stimuli in humans. Eur J Neurosci 1995;7: 1285–91.

Tallon-Baudry C, Bertrand O, Delpuech C, Pernier J. Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. J Neurosci 1996;16:4240–9.

Tallon-Baudry C, Bertrand O, Delpuech C, Pernier J. Oscillatory gamma band (30–70 Hz) activity induced by visual search task in human. J Neurosci 1997;17:722–34.

Tallon-Baudry C, Bertrand O, Peronnet F, Pernier J. Induced gamma-band activity during the delay of a visual short term memory task in humans. J Neurosci 1998;18:4244–54.

Tass P, Rosenblum MG, Weule J, Kurths J, Pikovsky A, Volkmann J, Schnitsler A, Freund H-J. Detection of n:m phase locking from noisy data: application to magnetoencephalography. Phys Rev Lett 1998;81:3291–4.

Tiitinen H, Sinkkonen J, Reinikainen K, Alho K, Lavikainen J, Naatanen R. Selective attention enhances the auditory 40-hz transient response in humans. Nature 1993;364:59–60.

Vasey MW, Thayer JF. The continuing problem of false positives in repeated measures ANOVA in psychophysiology: a multivariate solution. Psychophysiology 1987;24:479–86.

Varela FJ. Resonant cell assemblies: a new approach to cognitive function and neural synchrony. Biol Res 1995;28:81–95.

Varela F, Lachaux JP, Rodriguez E, Martinerie J. The brainweb: phase synchronization and large-scale integration. Nature Rev Neurosci 2001; 2:229–39.

von der Malsburg c, Singer W. Principles of cortical network organization. In: Rakic P, Singer W, editors. Neurobiology of the neocortex: proceedings of the Dahlem conference. Chichester: Wiley; 1988. p. 69–99.

Wasserman S, Bockenholt U. Bootstrapping: applications to psychophysiology. Psychophysiology 1989;26:208–21.

Yarbus AL. Eye movements and vision. New York: Plenum Press; 1967.