LETTER Communicated by Rufin VanRullen

A New Foreperiod Effect on Intertrial Phase Coherence. Part I: Existence and Behavioral Relevance

Joaquin Rapela
rapela@ucsd.edu
Swartz Center for Computational Neuroscience, University of California San Diego, La Jolla, CA 92093, and Instituto de Investigación en Luz, Ambiente y Visión, Universidad Nacional de Tucumán–Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina

Marissa Westerfield
mwesterfi@gmail.com
Swartz Center for Computational Neuroscience, University of California, San Diego, La Jolla, CA 92093, U.S.A., Research on Autism and Development Laboratory, University of California, San Diego, La Jolla, CA 92037

Jeanne Townsend
jtownsend@ucsd.edu
Research on Autism and Development Laboratory, University of California, San Diego, La Jolla, CA 92037

This letter makes scientific and methodological contributions. Scientifically, it demonstrates a new and behaviorally relevant effect of temporal expectation on the phase coherence of the electroencephalogram (EEG). Methodologically, it introduces novel methods to characterize EEG recordings at the single-trial level. Expecting events in time can lead to more efficient behavior. A remarkable finding in the study of temporal expectation is the foreperiod effect on reaction time, that is, the influence on reaction time of the delay between a warning signal and a succeeding imperative stimulus to which subjects are instructed to respond as quickly as possible. Here we study a new foreperiod effect in an audiovisual attention-shifting oddball task in which attention-shift cues directed the attention of subjects to impendent deviant stimuli of a given modality and therefore acted as warning signals for these deviants. Standard stimuli, to which subjects did not respond, were interspersed between warning signals and deviants. We hypothesized that foreperiod durations modulated intertrial phase coherence (ITPC, the degree of phase alignment across multiple trials) evoked by behaviorally irrelevant standards and that these modulations are behaviorally meaningful. Using averaged data, we first observed that ITPC evoked by standards closer to the warning signal was significantly different from that evoked by standards

Neural Computation **30, 1–37** (2018)
doi:10.1162/neco_a_01109

further away from it, establishing a new foreperiod effect on ITPC evoked by standards. We call this effect the standard foreperiod (SFP) effect on ITPC. We reasoned that if the SFP influences ITPC evoked by standards, it should be possible to decode the former from the latter an a trial-by-trial basis. We were able to do so showing that this effect can be observed in single trials. We demonstrated the behavioral relevance of the SFP effect on ITPC by showing significant correlations between its strength and subjects' behavioral performance.

1 Introduction

Relating electrical recordings from the brain to psychological phenomena has been a central goal of EEG research since its conception (Lindsley, 1952). The investigation of the neural bases of temporal expectation is a success story along this line. It began at the birth of experimental psychology when Wundt (1874) and Woodrow (1914) discovered effects on reaction times of the duration of the interval between a warning signal and a subsequent imperative stimulus to which subjects were instructed to respond as quickly as possible. This interval is termed the *foreperiod* and these effects the *foreperiod effects on reaction time*. The first electrophysiological marker of temporal expectation in foreperiod experiments was the contingent negative variation (CNV; Walter, Cooper, Aldridge, McCallum, & Winter, 1964), a slow negative shift in voltages recorded with the EEG that develops between a warning signal and the subsequent imperative stimulus. Hillyard and Galambos (1967) highlighted the relevance of the CNV by showing strong correlations between its size and subjects' reaction times.

Recent studies have demonstrated that almost every aspect of perception, motor control and cognition is related to intertrial phase coherence (ITPC). In perception, ITPC has been linked to vision (Busch, Dubois, & VanRullen, 2009; Mathewson, Gratton, Fabiani, Beck, & Ro, 2009), audition (Lakatos, Karmos, Mehta, Ulbert, & Schroeder, 2008; Stefanics et al., 2010), and speech (Luo & Poeppel, 2007; Zoefel & VanRullen, 2016). In addition, ITPC can operate cross-modally, with the alignment of phases in visual brain regions being triggered by auditory stimuli, and vice versa (Thorne, De Vos, Viola, & Debener, 2011; Romei, Gross, & Thut, 2012; Simon & Wallace, 2017). In motor control, ITPC has been connected to eye movements (Hamm, Dyckman, Ethridge, McDowell, & Clementz, 2010; Drewes & VanRullen, 2011). In cognition, it has been linked to attention (Yamagishi, Callan, Anderson, & Kawato, 2008; Gray, Frey, Wilson, & Foxe, 2015), causality judgment (Cravo, Santos, Reyes, Caetano, & Claessens, 2015), stimuli coincidence (Milton & Pleydell-Pearce, 2016), temporal predictions (Samaha, Bauer, Cimaroli, & Postle, 2015) and their monitoring and updating (Barne, Claessens, Reyes, Caetano, & Cravo, 2017), the transmission of prior information to visual cortex (Sherman, Kanai, Seth, & VanRullen,

2016), and to executive functions (e.g., Anguera, Lyman, Zanto, Bollinger, & Gazzaley, 2013).

We analyzed EEG recordings from an audiovisual attention-shifting oddball task where an attention-shift LOOK (HEAR) cue instructed subjects to begin detecting visual (auditory) deviants, and to ignore visual and auditory standards (see Figure 1). Every attention-shifting cue was followed by one or more deviants at variable times. Thus, an attention-shifting cue initiated a period of expectation for an ensuing deviant and therefore acted as warning signals for this deviant in a variable-foreperiod-duration task. The relatively strong EEG due to responses to deviants may occlude potentially weaker foreperiod effects on the ITPC evoked by deviants. However, these effects may be visible on ITPC evoked by standards. Here we study effects of the delay between the presentation of a warning signal and a subsequent standard the standard foreperiod duration (SFPD) on the ITPC evoked by this standard. We call these effects *standard foreperiod* (SFP) effects on ITPC.

Since effects of attention on the EEG can be observed in single trials (e.g., O'Sullivan et al., 2015; Nunez, Vandekerckhove, & Srinivasan, 2016; Horton, Srinivasan, & DZmura, 2014), we anticipated that effects of temporal expectation on the EEG could also be observed in single trials. We reasoned that if the ITPC evoked by a standard depends on the SFPD corresponding to this standard, then it may be possible to reliably decode this SFPD from a single-trial measure of the ITPC evoked by the standard. To measure ITPC in single trials, we used the deviation from the mean phase (DMP; see section 2.5), a measure of the distance between the phase of a single trial and the mean phase of all trials. And to decode an SFPD from the DMP evoked by a standard, we used a multivariate linear-regression model (see section 2.6), estimated using a variational-Bayes technique (see section 2.7).

Relations between the phase of the EEG and behavior have been well documented. Phase has been related to reaction times (Stefanics et al., 2010; Drewes & VanRullen, 2011; Thorne et al., 2011; Zoefel & Heil, 2013) and to the perception of visual (Valera, Toro, John, & Schwartz, 1981; Busch et al., 2009; Mathewson et al., 2009; Mathewson, Fabiani, Gratton, Beck, & Lleras, 2010; Cravo, Rohenkohl, Wyart, & Nobre, 2013; de Graaf et al., 2013; Hanslmayr, Volberg, Wimber, Dalal, & Greenlee, 2013; Spaak, de Lange, & Jensen, 2014; Cravo et al., 2015; Gray et al., 2015) and auditory (Stefanics et al., 2010; Henry & Obleser, 2012; Ng, Schroeder, & Kayser, 2012; Lakatos et al., 2013; Hickok, Farahbod, & Saberi, 2015) stimuli. We thus conjectured that modulations of ITPC by the SFPD could also be related to reaction times or the perception of visual and auditory stimuli.

We hypothesized that in our attention-shifting oddball experiment, the variable delay between the presentation of an attention-shifting cue and an ensuing standard (i.e., the SFPD) modulates in a trial-by-trial basis our single-trial measure of ITPC (i.e., the DMP) and that this modulation is behaviorally relevant.

Table 1: Glossary.

Attended modality	The attended modality (visual or auditory) corresponding to a set of epochs.
DMP (deviation from the mean phase)	Single-trial measure of ITPC.
ECD	Equivalent current dipole for an IC
ERP (event-related potential)	Mean voltage across a group of epochs aligned to an event of interest.
IC (independent component)	Component from an ICA decomposition.
ISI (interstimuli interval)	The interval between two succeding stimuli.
ITPC	Degree of phase alignment across multiple trials. Here we use the DMP and the ITC as measures of ITPC for single trials and groups of trials, respectively. That is, the ITPC is a quantity that we measure with the DMP and the ITC.
SFP (standard foreperiod)	Interval between the presentation of a warning signal and a standard stimulus.
SFPD (standard foreperiod duration)	Duration of an SFP.
Standard modality	The modality of the standard stimuli (visual or auditory) used to align a set of epochs.
Warning signal	Stimulus initiating a period of expectancy for a forthcoming impendent stimulus. The LOOK and HEAR attention-shift cues are the warning signals in this study.

Using a simple method based on trial averaging, we present in section 3.2 direct evidence for the SFP effect on ITPC. In section 3.3, we quantify the strength of the new effect using a single-trial decoding model. These two sections establish the existence of the new SFP effect on ITPC. To demonstrate its behavioral relevance, we show in section 3.4 that its strength is significantly correlated with behavioral measures (stimuli detectability and possibly reaction speed). We close the letter with a discussion in section 4. Abreviated terms are defined in Table 1. A previous more pedantic preprint version of this article appeared in Rapela, Westerfield, Townsend, and Makeig (2016).

2 Methods

2.1 Experiment Information. We analyzed the experimental data first characterized in Ceponiene, Westerfield, Torki, and Townsend (2008). We summarize features of these data relevant to our study; details are given in Ceponiene et al. (2008).

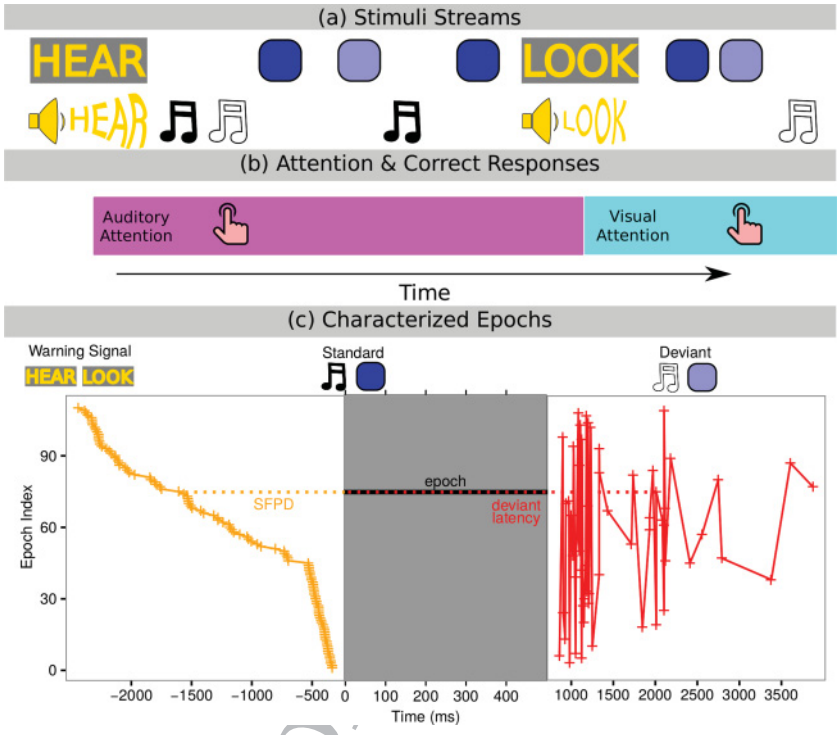


Figure 1: Audiovisual attention-shifting experiment. (a) Audiovisual LOOK and HEAR attention-shift cues, visual standards (dark blue squares) and deviants (light blue squares), and auditory standards (low-pitch sounds, filled notes) and deviants (high-pitch sounds, open notes) were presented sequentially in pseudo-random order. (b) After seeing and hearing a LOOK (HEAR) cue, subjects indicated with a button press when they saw a visual (heard an auditory) deviant. (c) Epochs began with standard visual or auditory stimuli and lasted 500 ms (black solid line in the gray rectangle). To each epoch corresponded an SFPD, the time between the presentation of the previous warning signal and the standard at time zero (orange dotted line). We used only epochs where deviants of the attended modality appeared more than 500 ms after the standard at time zero. SFPDs and deviant latencies correspond to subject av130a and unattended visual standards.

2.1.1 Subjects. We characterized only the younger-adult subpopulation in Ceponiene et al. (2008), comprising 19 subjects (11 females) with a mean age of 25.67 ± 5.94 years.

2.1.2 Stimuli. Stimuli were sequentially presented in visual and auditory streams (see Figure 1a). Auditory stimuli were 100 ms duration, 550 Hz

(deviants), and 500 Hz (standards) sine-wave tones. These tones were played using two loudspeakers located at the sides of a 21 inch computer monitor. Visual stimuli were light blue (deviants) and dark blue (standards) filled squares subtending 3.3 degrees of visual angle, presented on a computer monitor for 100 ms on a light gray background. Interspersed among deviants and standards were attention-shift cues. These cues were presented bimodally for 200 ms by displaying the words HEAR (LOOK) on orange letters on the computer monitor and simultaneously playing the words HEAR (LOOK) on the loudspeakers. At least one deviant followed the presentation of each attention-shifting cue. The interstimulus interval (ISI) between any two consecutive stimuli varied between 100 and 700 ms as random samples from short, medium, and large uniform distributions, with supports [100–300], [301–500], and [501–700] ms, respectively. In each block, the ISIs of 132, 60, and 72 stimuli were drawn from the short, medium, and large ISI distributions, respectively. Each block consisted of 24 cue stimuli (12 visual) and 240 noncue stimuli (120 visual). The 120 noncue stimuli of each modality comprised 24 deviants and 96 standards. The duration of a block was 158 seconds. A video of the experimental stimuli appears at http://sccn.ucsd.edu/∼rapela/avshift/experiment.MP4.

2.1.3 Experimental Design. The experiment had FOCUS-VISION, FOCUS-AUDITION, and SWITCH blocks. In FOCUS-VISION (-AUDITION) blocks, subjects had to detect visual (auditory) deviants and ignore attention-shift cues. In SWITCH blocks, these cues became relevant, and after a LOOK (HEAR) cue, subjects had to detect visual (auditory) deviants. The type of block was told to subjects at the beginning of each block. Subjects pressed a button when they detected deviants. Each subject completed 4 FOCUS-VISION blocks, 4 FOCUS-AUDITION blocks, and 12 SWITCH blocks. Here we characterized only SWITCH blocks. Correct responses in SWITCH blocks are shown in Figure 1b. After a LOOK (HEAR) cue, subjects oriented their attention to the visual (auditory) modality, as indicated by the magenta (cyan) segments in Figure 1b.

2.2 EEG Acquisition. Continuous EEG was recorded from 33 scalp sites (FP1, FPz, FP2, AF3, AF4, F7, F8, F3, Fz, F4, FC1, FC2, FC5, FC6, T7, T8, C3, Cz, C4, CP1, CP2, CP5, CP6, P3, Pz, P4, P7, P8, PO3, PO4, O1, Oz, O2) of the International 10-20 system, using an SA (SA Instruments, San Diego, CA) amplifier and electrode cap, and digitized at 250 Hz. The right mastoid served as reference. EEG data were high- and low-pass filtered with cutoff frequencies of 1 Hz and 50 Hz, respectively.

2.3 ICA Decomposition. EEG-recorded potentials were decomposed into independent components (ICs) using independent component analysis (ICA; see section A.1.1).

2.4 Characterized Epochs. For each IC of each subject, we built four sets of epochs, aligned at time zero to the presentation of attended visual standards, attended auditory standards, unattended visual standards, and unattended auditory standards, with a duration of 500 ms (see Figure 1). We define the standard modality as the modality of the standards used to align a set of epochs (visual or auditory) and the attended modality as the attended modality corresponding to a set of epochs (e.g., the visual (auditory) modality for epochs aligned to the presentation of visual (auditory) standards preceded by a LOOK (HEAR) attention-shifting cue). The mean numbers of epochs aligned to the presentation of attended visual standards, attended auditory standards, unattended visual standards, and unattended auditory standards were 150, 136, 99, and 94, respectively. Note that the mean numbers of epochs aligned to the presentation of attended standards were considerably larger than those aligned to the presentation of unattended standards for both visual and auditory standards. Thus, for testing influences of attention on the SFP effect on ITPC (see section 3.4), we equalized the number of epochs used to fit attended and unattended models. For each attended model, we selected a random subset of the attended epochs of a size equal to the number of epochs used to fit the corresponding unattended model. To avoid possible movement artifacts from responses to deviants, we excluded from further analysis epochs including deviants of the attended modality (epochs including deviants to which subjects were instructed to respond with a mouse click; e.g., red crosses appear after the 500 ms gray box in Figure 1).

2.4.1 Surrogate Epochs to Test That Modulations of ITPC Are due to the Warning Signal and Not to Standards. We built a set of surrogate epochs to test the hypothesis that the warning signal immediately preceding a standard modulated the ITPC allowed significant decodings from models (see section A.2.1). For each epoch in an original data set, we built an epoch in the corresponding surrogate data set. To construct the surrogate epoch, the onset time of the standard on the original data set was shifted by a random number between a minimum and a maximum value. The minimum value was the negative of the SFPD in order to guarantee that the surrogate standard onset occurred after the previous warning signal. The maximum value was the minimum among the latency of the next deviant and the latency of the next warning signal, minus 500 ms. In this way, a surrogate standard occurred before the next deviant and the next warning signal, and neither deviants nor warning signals appeared in the 500 ms window used to decode the SFPD. Since the onset time of standards determines the DMP values used as independent variables in the decoding model (see section 2.6), the surrogate data sets randomized only the independent variables of these models. In this surrogate data set, the distribution of phases at every moment in time was not significantly different from the uniform distribution ($p > 0.01$, Raleigh test). In the main analysis, to decode SFPDs, we used

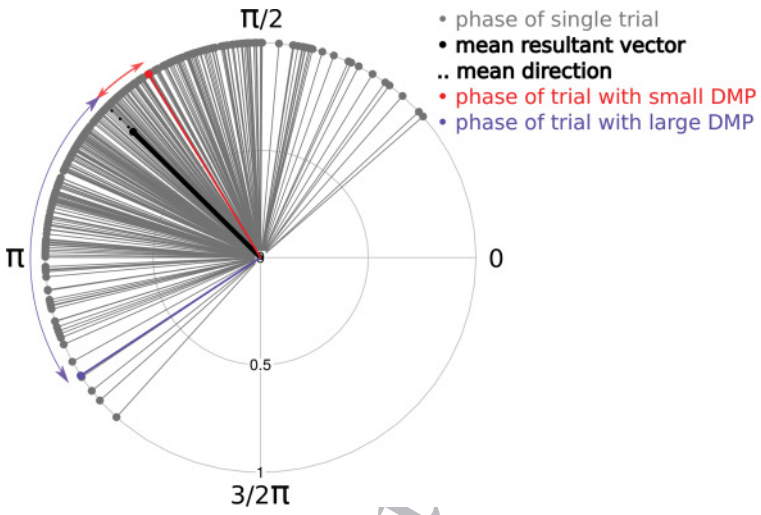


Figure 2: Computation of the DMP. Given a set of phases, represented by gray unit vectors, one first calculates the mean of these unit vectors (the mean resultant vector, black solid vector) and extracts the phase of this mean vector (mean direction; see equation A.6, black dotted unit vector). Then the deviation from the mean phase for a given phase (e.g., the phase corresponding to the red unit vector) is a measure of distance (i.e., circular variance; see equation A.5) between this phase and the mean direction (e.g., red arc). The red and blue unit vectors correspond to phases with small and large DMPs, respectively.

DMPs only at time points at which the distribution of phases was significantly different from the uniform distribution. If we had applied this same criterion with the surrogate epochs, we would have no time points to decode SFPDs. Thus, for epochs in surrogate data sets, we used the same time points as in the original epochs to decode SFP durations.

2.5 Deviation from the Mean Phase. Let θ_0 be the phase of a given trial and $\theta_1, \ldots, \theta_N$ be the phases of trials in a reference set, with all phases measured at the same time-frequency point. Then the deviation from the mean phase is a measure of the distance between the phase of the given trial (i.e., θ_0) and the mean phase of all trials in the reference set (i.e., mean direction, $\bar{\theta}(\theta_1, \ldots, \theta_N)$; see equation A.6), as illustrated in Figure 2. (All cross-references beginning with "A" are in the online appendix.) The circular variance (CV; see equation A.5) is used to measure this distance:

$$DMP(\theta_0|\{\theta_1, \ldots, \theta_N\}) = CV(\theta_0, \bar{\theta}(\theta_1, \ldots, \theta_N)) \tag{2.1}$$

The DMP is zero (one) if the phase of the given trial is equal (opposite) to the mean phase of the trials in the reference set. The DMP has previously been used to investigate interelectrode phase coherence with EEG recordings (Hanslmayr et al., 2007, Figure 2c). It is an appealing measure of single-trial ITPC since, as demonstrated in section A.1.7, when there is large phase concentration, the average DMP approaches intertrial coherence (ITC; Tallon-Baudry, Bertrand, Delpuech, & Pernier, 1996; Delorme & Makeig, 2004), a measure of ITPC averaging information across multiple trials. When there is not a large phase concentration, the mean phase cannot be estimated reliably, and therefore the DMP becomes unsound. Here we computed DMP only when the distribution of phases across trials was significantly different from the uniform distribution ($p < 0.01$, Raleigh test).

2.6 Decoding Model. We used a multivariate linear-regression model to decode the SFPD of trial n from samples of DMP from this trial:

$$\hat{y}[n, \mathbf{w}] = \sum_{k=1}^{K} w[k]x[n, k], \tag{2.2}$$

where $\hat{y}[n, \mathbf{w}]$ is the decoded SFPD, $w[k]$ is a regression coefficient, $x[n, k]$ is the DMP for trial n at sample time k, and K is the number of sample points in the 500 ms long time window following the presentation of standards, for which the distribution of phases across trials was significantly different from the uniform distribution ($p < 0.01$; Rayleigh test).

2.7 Method to Estimate the Coefficients of the Decoding Model. We seek coefficients $\mathbf{w}$ in the linear-regression model (see equation 2.2) such that decodings of the model, $\hat{y}[\cdot, \mathbf{w}]$, are as close as possible to experimental SFPDs, $y[\cdot]$. Mathematically, we seek regression coefficients $\mathbf{w}$ that minimize the least-squares error function,

$$\text{MSE}(\mathbf{w}) = \sum_{n=1}^{N} (y[n] - \hat{y}[n, \mathbf{w}])^2, \tag{2.3}$$

where N is the number of epochs. A difficulty in this estimation is that DMP at neighboring sample points are highly correlated, and correlations increase the variance of ordinary least-squares estimates (Belsley, Kuh, & Welsch, 2004). In addition, in some cases, the number of coefficients, K in equation 2.2, is equal to the number of epochs, N in equation 2.3. This further increases the variance of the ordinary least-squares estimates. To address these problems here, we searched for coefficients that minimized the following ridge-regression error function (Hastie, Tibshirani, & Friedman, 2016, section 3.4.1),

$$\mathrm{RMSE}(\mathbf{w}) = \mathrm{MSE}(\mathbf{w}) + \alpha||\mathbf{w}||^2, \tag{2.4}$$

which adds a penalty constraint to the least-squares error function in equation 2.3, shrinking coefficients estimates toward zero and therefore reducing their variability. In equation 2.4, α determines the strength of the constraint (i.e., for larger α, the penalty constraint more strongly biases the coefficient estimates away from minimizing the least-squares error function in equation 2.3, and toward zero).

To find the optimal $\mathbf{w}$ in equation 2.4, we took a Bayesian approach (Bishop, 2016). We used a gaussian likelihood function,

$$P(\mathbf{y}|\mathbf{w}, \tau, \Phi) = N(\mathbf{y}|\Phi\mathbf{w}, \tau^{-1}I), \tag{2.5}$$

where Φ is the matrix of DMP (i.e., $\Phi[n, k] = x[n, k]$) and τ is the constant precision of $\mathbf{y}$. We chose a normal-gamma prior for $\mathbf{w}$ and τ,

$$P(\mathbf{w}, \tau|\alpha) = N(\mathbf{w}|\mathbf{0}, (\tau\alpha)^{-1}I)\ \mathrm{Gam}(\tau|a_0, b_0), \tag{2.6}$$

and a gamma hyperprior for the hyperparameter α,

$$P(\alpha) = \mathrm{Gam}(\alpha|c_0, d_0), \tag{2.7}$$

and searched for the $\mathbf{w}$ that maximized the log of the posterior distribution:

$$\mathrm{J}(\mathbf{w}) = \log P(\mathbf{w}|\mathbf{y}, \tau, \alpha, \Phi) \tag{2.8}$$

As we prove in proposition 2 in section A.1.8, with this choice of likelihood function and priors, finding the coefficients that maximize the log of the posterior distribution in equation 2.8 is equivalent to finding the coefficients that minimize the ridge-regression error function in equation 2.4, as we set out to do at the beginning of this section. Having rewritten the original ridge-regression optimization problem (see equation 4) as a Bayesian optimization problem (see equation 2.8) we can now use advanced Bayesian inferential methods (Bishop, 2016) to estimate the model coefficients, as we show next.

Due to the large dimensionality of $\mathbf{w}$, evaluating the posterior distribution in equation 2.8 is not feasible, and one needs to resort to approximation schemes (Bishop, 2016). These approximations can be stochastic or deterministic. In principle, stochastic approximations, such as Markov chain Monte Carlo (Metropolis & Ulam, 1949), can give exact evaluations of the posterior distribution, given infinite computational resources, but are often restricted to small-scale problems. Deterministic approximations are based on approximations of the posterior distribution, and although they can never generate exact results, they scale well for large problems. Here

we use the variational Bayes deterministic approximation (Bishop, 2016) and implement it as described in Rapela (2017).

2.8 Details on Model Estimation. If the delay between the warning signal and the following deviant is longer than a threshold that depends on the distribution of deviants, the foreperiod effect on reaction times disappears (Botwinick & Brinley, 1962). Similarly, if we estimate decoding models including a large proportion of standards presented long after the warning signal, decodings become nonsignificant (i.e., the SFP effect on ITPC disappears). To fit decoding models, we used data from standards that were presented before a maximum SFPD after the warning signal. The selection of this maximum is described in section A.1.13. In Figure A.5, we selected the minimum SFPD larger than 1 second, as well as the maximum SFPD, which maximized the decoding power of models. We estimated models only in cases where the number of SFPDs was larger than or equal to the number of sample points between the minimum and maximum SFPD to avoid ill-conditioned regression problems.

For all estimations, we used $a_0 = 1e-2$ and $b_0 = 1e-4$ in equation 2.6 and $c_0 = 1e-2$ and $d_0 = 1e-4$ in equation 2.7. We applied a logarithmic transformation to the dependent variable, $y[n]$, in order to equalize the variance of the residuals (Kutner, Nachtsheim, Neter, & Li, 2005) and standardized the dependent variable and regressors to have zero mean and unit variance (Kutner et al., 2005). Highly influential trials (i.e., trials with a Cook's distance larger than $4/(N-K-1)$, where N and K are the number of trials and regressors) were deleted before estimating model parameters (Kutner et al., 2005).

2.9 Data and Software Sharing. To facilitate the application of the single-trial decoding method described in this letter to other EEG studies, we provide the R (R Core Team, 2012) code implementing these methods, the EEG data from the left parieto-occipital cluster 4, and instructions on how to use this code to analyze these data at https://github.com/joacorapela/singleTrialEEGPredictions.

3 Results

3.1 Behavioral Results. We define the deviant foreperiod duration as the delay between the presentation of a warning signal and the next deviant. We detected significant foreperiod effects on reaction times (i.e., significant correlations between deviant foreperiod durations and subjects' reaction times to deviants) in 10 combinations of subject and attended modality (26% out of a total of 38 combinations of 19 subjects and 2 attended modalities). Five of these combinations corresponded to the visual attended modality. These correlations were all positive for the visual attended modality (indicating that longer deviant foreperiod durations corresponded to longer

reaction times), while for the auditory attended modality, there was a mixture of two positive and three negative significant correlations. Figure A.8a plots deviant foreperiod durations as a function of reaction times for an example subject and attended modality. We also found significant foreperiod effects on detectability (i.e., significant differences between the median deviant foreperiod duration for hits and misses) in 10 other combinations of subject and attended modality (26%). All of these combinations corresponded to the auditory attended modality, and in all of them, subjects more easily detected later deviants (i.e., the median deviant foreperiod was significantly larger for hits than for misses). Figure A.8b plots deviant foreperiod durations for hits and misses for an example subject and attended modality.

Visual deviants were detected more reliably and faster than auditory ones. Mean error rates were 0.22 and 0.09 for detecting auditory and visual deviants, respectively. A repeated-measures ANOVA, with error rate (in the detection of deviants of the attended modality) as dependent variable and attended modality as independent factor, showed a significant main effect of attended modality ($F(1, 17) = 39.35$, $p <$ 1e-4). A post hoc analysis revealed that error rates were smaller when detecting visual than auditory deviants ($z = 6.45$, $p =$ 5.59e-11). Mean response times to auditory and visual attended deviants were 407 and 377 ms, respectively. A repeated-measures ANOVA, with mean response time (to deviants of the attended modality) as dependent variable and attended modality as independent factor, showed a significant main effect of attended modality ($F(1, 18) = 40.29$, $p <$ 1e-4). A post hoc analysis revealed that mean response time was shorter for visual than auditory deviants ($z = 6.52$, $p =$ 3.49e-11).

3.2 Evidence for an SFP Effect on ITPC from a Trial-Averaged Analysis. As a first approximation, to assess if the SFPD modulates ITPC evoked by standards, we used a simple method comparing mean DMP of the 20% trials farther away from the warning signal with that of trials closer to it. The existence of a significant difference in these means would be direct evidence for a SFP effect on ITPC.

Plots in Figure 3 correspond to epochs aligned to the presentation of attended visual standards, and different columns correspond to different example subjects and ICs. The red and blue lines in Figures 3d to 3f plot the mean DMP of the 20% trials farthest from and closest to the warning signal, respectively. Figures 3g and 3h plot the difference between the red and blue lines in the corresponding Figures 3d to 3f. For subject av124a, IC 7, and phase-measured at the peak ITC frequency of 7.8 Hz (left column), the mean DMP was significantly larger for trials with the longest SFPD between 0 and 150 ms after the presentation of standards but smaller between 250 and 320 ms (see Figure 3g). For subject av115a, IC 20, and phase-measured at the peak ITC frequency of 3.9 Hz (central column), the mean DMP was significantly smaller for trials with the longest SFPD between 200 and 320 ms

(see Figure 3h). And for subject av121a, IC 2, and phase-measured at the peak ITC frequency of 4.9 Hz (right column), the mean DMP was similar for trials with the longest and shortest SFPD (see Figure 3i). Figure A.11 shows mean DMPs for all ICs from the left parieto-occipital cluster 4 and attended visual standards.

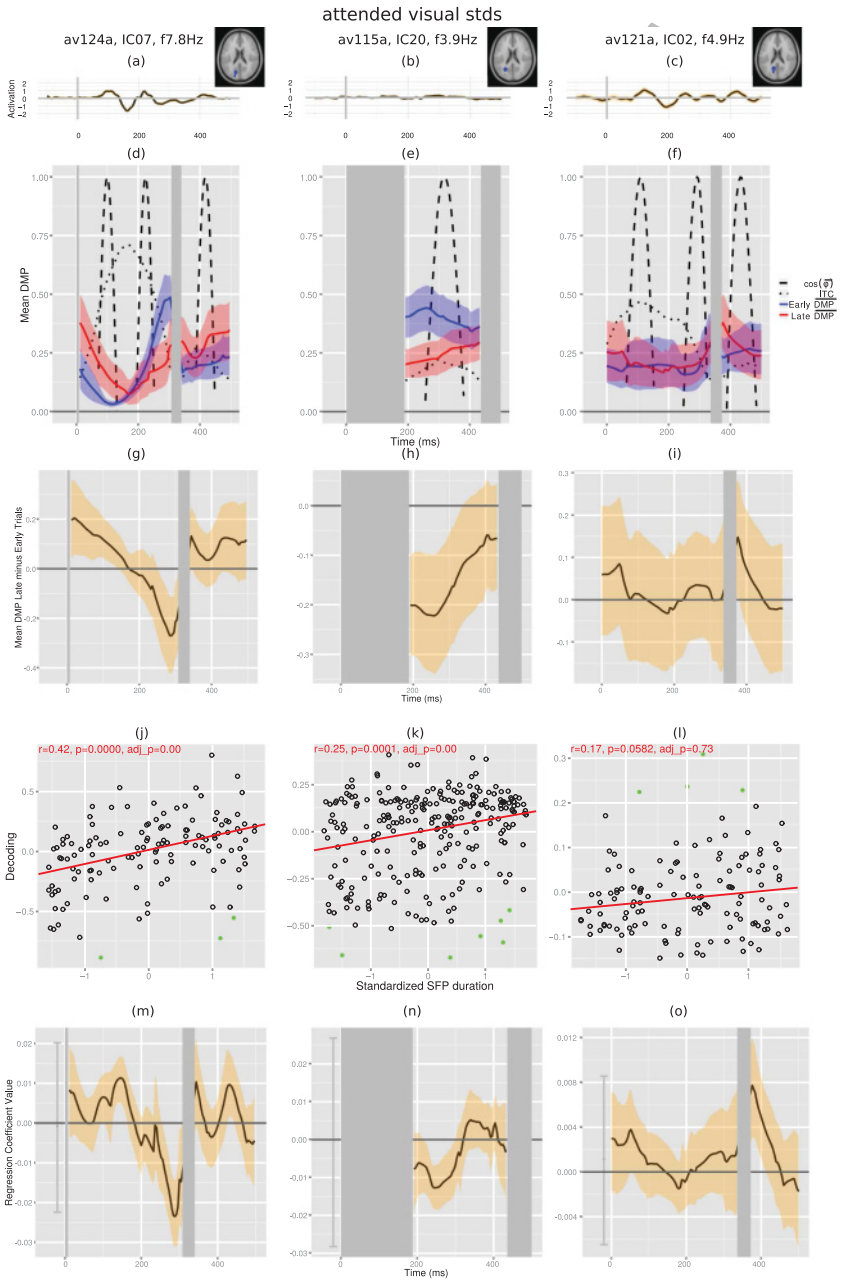
The existence of a foreperiod effect on the ITPC triggered by standards (see Figures 3d to 3i and A.11) is the main finding of this letter. We next quantify the strength of this effect using a single-trial analysis in section 3.3 and argue for its behavioral relevance in section 3.4.

3.3 Quantifying the Strength of the SFP Effect on ITPC Using a Single-Trial Analysis. In the previous section, we presented evidence for the existence of a new foreperiod effect on the ITPC evoked by standards using a simple method operating on averaged data. Here we use a single-trial decoding method to quantify the strength of this effect.

Figures 3j to 3l plot experimental SFPDs versus their leave-one-out cross-validated decodings from DMP. For IC 7 of subject av124a and IC 20 of subject av115a, for which means of DMP of trials closest to and farthest from the warning signal were significantly different (blue versus red curves in Figures 3d and 3e), decodings from models were significantly correlated with experimental SFPDs (adjusted $p < 0.01$; see Figures 3j and 3k). For IC 2 of subject av121a, for which the mean DMP for trials closest to and farthest from the warning signal were not statistically different (blue versus red curves in Figure 3f), decodings from the corresponding model were not significantly correlated with experimental SFPDs (adjusted $p = 0.73$; see Figure 3l). For each IC, standard modality and attended modality, we quantify the strength of the SFP effect on ITPC with the correlation coefficient between the corresponding model's decodings and the experimental SFPDs.

We observed the SFP effect on ITPC in most subjects. Out of a maximum of 32 models estimated per subject, standard modality and attended modality, the number of models with significant correlations between models' decodings and experimental SFPDs (as in Figures 3g to 3i) varied between 0 and 11 (see Figure A.10), and it was different from 0 in 18 of the 19 subjects for attended standards (see Figures A.10a and A.10d) and for 17 of the 19 subjects for unattended standards (see Figures A.10b and A.10c). However, this number of models was not uniformly distributed across subjects. For the auditory attended modality, there was a tendency to obtain more significant correlations for models from subjects achieving smaller error rates (see the right column in Figure A.10).

To visualize how the SFP effect on ITPC distributed across the brain, we grouped all ICs of all subjects using a clustering algorithm (see section A.1.3). Figure 4 plots axial projections of the obtained clusters (Figure A.9 displays axial, sagital, and coronal projections of them, and Table A.1 provides additional information). Some equivalent current dipoles (ECDs)


attended visual stds
av124a, IC07, f7.8Hz
av115a, IC20, f3.9Hz
av121a, IC02, f4.9Hz
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
(j)
(k)
(l)
(m)
(n)
(o)
Activation
Mean DMP
Mean DMP Late minus Early Trials
Decoding
Regression Coefficient Value
Time (ms)
Standardized SFP duration
r=0.42, p=0.0000, adj_p=0.00
r=0.25, p=0.0001, adj_p=0.00
r=0.17, p=0.0582, adj_p=0.73
cos(ϕ)
ITC
Early DMP
Late DMP

of ICs appear to lie outside the brain (e.g., ECDs in cluster 19 appear to lie in the ventricles). However, ICs with ECDs outside the brain were excluded from this study, and the previous appearance is due to the difficult two-dimensional visualization of three-dimensional ECDs, explained in section A.5. For each cluster, standard modality and attended modality, we computed the proportion of models with decodings significantly correlated with experimental SFPDs, as in Figures 3j and 3k. Dots below the image of a cluster in Figure 4 indicate that decodings of more than 40% of the models, estimated from data from that cluster and from the standard modality and attended modality given by the color of the dot, were significantly correlated with experimental SFPDs. For each cluster, standard modality and attended modality, the proportion of models with decodings significantly correlated with experimental SFPDs is given in Table A.2.

3.4 The Strength of the SFP Effect on ITPC Is Correlated with Behavior. In the previous section, we quantified the strength of the SFP effect on ITPC. In this section, we show that this strength is significantly correlated with subjects' detection error rates and possibly with mean reaction times.

The ordinate in Figure 5 gives the correlation coefficient between decodings from a model and experimental SFPDs (i.e., the strength of the SFP effect on ITPC; see section 3.3) for all ICs in the central-midline cluster 19 and unattended visual standards. The abscissa provides error rates of the subjects corresponding to the ICs. The significant negative correlation between the strength of the SFP effect on ITPC and subjects' error rates

Figure 3: Standard foreperiod effect on ITPC revealed by a simple method in averaged data and by a decoding method in single-trial data. Each column corresponds to an IC of a subject. Data are from epochs aligned to the presentation of attended visual standards. (a–c) Evoked response potential (ERP) from all trials. (d–f) Mean DMP evoked by the 20% standards closest to (blue curve) and farthest from (red curve) the warning signal. Dashed and dotted lines plot the cosine of the mean phase and the ITC, respectively, from all trials. (g–i) Difference in mean DMP evoked by trials farthest from minus closest to the warning signal (i.e., the red minus the blue curves in panels d to f). (j–l) Model decodings versus standardized SFPDs. (m–o) Standardized coefficients of regression models. Gray boxes cover times without significant phase alignment ($p > 0.01$; Rayleigh test). Colored bands in all panels represent 95% confidence intervals. Significant differences between the red and blue curves in panels d to f Demonstrate that ITPC depends on the SFPD for the subjects and ICs in the left-most two columns but not for the subject in the right-most one. This dependence is corroborated by the significance of correlations in panels j to l. The similarity between panels g to i and m to o shows that the coefficients of a regression model indicate whether standards farther from the warning signal evoke more coherence or decoherent oscillations than standards closer to it (see section 3.5).

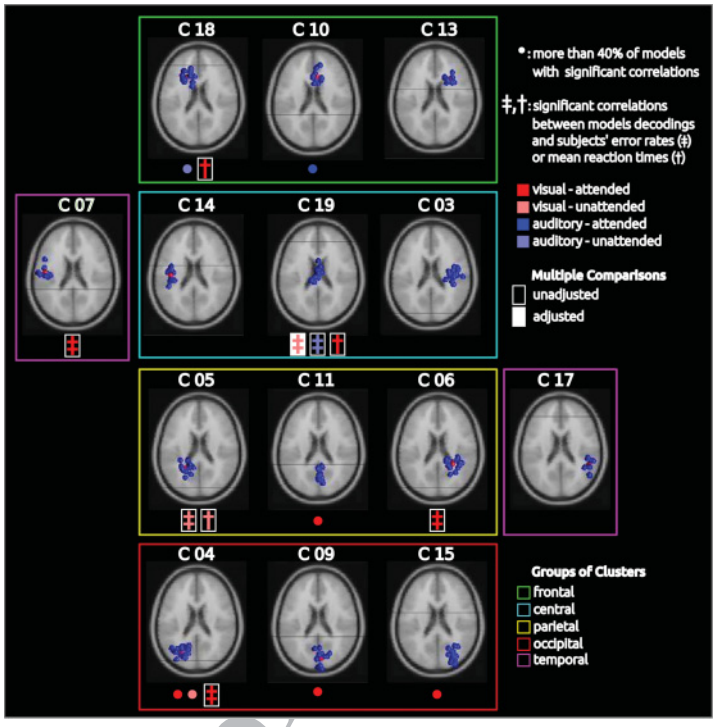


Figure 4: Clusters of ICs, large proportions of significant models, correlations with behavior, and groups of clusters. A blue ball inside a brain slice represents an IC from one subject. Each brain slice displays one cluster of ICs (see section A.1.3). A colored dot below the image of a cluster indicates that in more than 40% of the models estimated from data from that cluster and from the standard modality and the attended modality given by the color of the dot, the correlation coefficient between models' decoding and SFPDs was significantly different from zero (adjusted $p < 0.05$; see section A.1.9). A dagger (double dagger) signals a significant correlation between the strength of the SFP effect on ITPC and subjects' mean reaction times (error rates). Filled (unfilled) rectangles behind daggers indicate that the significance of the corresponding correlation test was corrected (uncorrected) for multiple comparisons. Colored boxes mark groups of clusters used in section 3.6 to study the timing of the SFP effect on ITPC. The clusters with a large proportion of significant models suggest that the SFP effect on ITPC is stronger for visual than auditory standards and for attended than unattended standards. The significant correlations between the strength of the SFP effect on ITPC and subjects' behavioral performance show that the effect is behaviorally relevant.

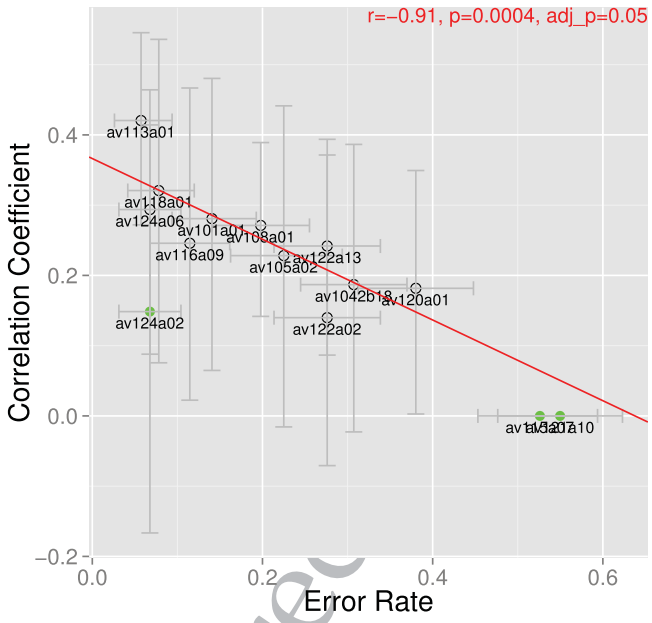


Figure 5: The degree of association between experimental SFPDs and ITPC evoked by standards is related to subjects' error rates in the mid-central cluster 19 and unattended visual standards. Stronger SFP effects on ITPC evoked by standards (i.e., larger correlation coefficients between models' decodings and SFPDs; ordinate) correspond to lower error rates (abscissa). Green points indicate outliers detected in the calculation of robust correlation coefficients (see section A.1.10).

($r = -0.91$, $p = 0.0004$, adjusted $p = 0.05$; p-values were adjusted for multiple comparisons using the procedure described in section A.1.9) shows that the higher is the correlation coefficient (i.e., the association between SFPD and ITPC evoked by standards) in an IC of a subject, the lower is his or her error rate.

We found other unadjusted significant correlations with error rates and mean reaction times that did not pass the multiple-comparison test, as indicated by the daggers and double daggers enclosed in unfilled rectangles in Figure 4 and by some blue entries in Tables A.3 and A.4. Below, we elaborate on these statistics following current good practices in statistics that recommend considering effect sizes and not drawing inferences based on p-values alone (see Greenland et al., 2016).

Note that decoding models were optimized to decode SFPDs and that correlations with behavioral measures resulted without fitting decoding models to behavioral data.

All significant correlations with error rates were negative (see Table A.3), indicating that subjects that reacted the fastest were those for whom it was easiest to decode the SFPD from their ITPC. That all five unadjusted significant correlations with error rates were negative (see Table A.3) is a very rare event under the assumption that all of these correlations occurred by chance (i.e., the probability of finding five negative correlations assuming equal probability for positive and negative correlations is $p = 0.5^5 = 0.03$). Thus, some of these correlations may not be spurious.

The absolute value of correlations with error rates was larger for unattended than attended standards ($p <$ 1e-04; permutation test).

For mean reaction times, the three significant correlations were also negative (see Table A.4), showing that subjects that reacted the fastest were those for which it was easiest to decode the SFPD from ITPC. However, differently from error rates, positive correlation for mean reaction times almost reached significance (e.g., cluster 6 and unattended auditory standards, or cluster 13 and unattended visual standards; see Table A.4).

Correlations (both significant and nonsignificant) were stronger for error rates than for mean reaction times and stronger for the visual than the auditory standard modality. An ANOVA with the absolute value of the correlation coefficient as dependent variable showed significant main effects of behavioral measure type (i.e., error rate or mean reaction time; $F(1, 109) = 7.26$, $p = 0.0082$) and for standard modality ($F(1, 109) = 8.22$, $p = 0.005$). A post hoc analysis indicated that the mean absolute value of the correlation coefficient was larger for error rates than for mean reaction times ($p = 0.005$; Tukey test) and larger for the visual than the auditory standard modality ($p = 0.0082$; Tukey test).

3.5 Interpretation of Decoding Model Coefficients. A positive regression coefficient indicates more decoherence at the corresponding time for trials with longer than shorter SFPDs. This is because the decoding model is a linear model between SFPD and DMP. Then a positive regression coefficient at a given time indicates (ignoring effects of correlation across time in DMP) that the line best fitting SFPDs to DMPs at the given time has a positive slope. This positive slope means that larger SFPDs (i.e., longer SFPDs) are associated with larger DMPs (i.e., more decoherence).

To validate this interpretation and the soundness of the decoding methodology, the bottom row in Figure 3 plots coefficients of the decoding models corresponding to the top rows. Figure 3m shows significantly positive regression coefficients between 100 and 200 ms and significantly negative coefficients between 240 and 300 ms. Therefore, according to the previous interpretation, trials with longer SFPDs should correspond to more decoherent oscillations between 100 ms and 200 ms and to less decoherent oscillations between 240 ms and 300 ms. This is indeed similar to what we observed in Figure 3d. Analogous consistencies hold for the other columns. The normalized cross-correlation (see section A.1.12) between

difference in mean DMP (see Figures 3g to 3i) and corresponding regression coefficients (see Figures 3m to 3o) was 0.80 for Figures 3g to 3m, 0.87 for Figures 3h to 3n, and 0.30 for Figures 3i to 3o. Across all models significantly different from the intercept-only model ($p < 0.01$; likelihood-ratio permutation test; see section A.1.12), the first, second (median), and third quartiles of the normalized cross-correlation distribution were 0.31, 0.69, and 0.85, respectively. These results indicate that on average, regression coefficients were similar to differences in mean DMP, which validates the previous interpretation of regression coefficients, and support the inference that reliable decodings by regression models indicate modulations of ITPC by SFPD.

3.6 Timing of the SFP Effect on ITPC. In the previous section, we advanced an interpretation of the coefficients of the decoding model and used a simple trial-averaging procedure to verify that this interpretation was sound. As shown in Figures 3d to 3f and Figure A.11, modulations of ITPC by SFPD are not constant in time but fluctuate in an oscillatory manner. In the 500 ms window following the presentation of a standard, these coefficients displayed one or more peaks. The time of the largest peak corresponds to the latency after the presentation of standards where modulations of ITPC by the SFPD are strongest. In this section, we use this time to represent the timing of the SFP effect on ITPC and study how this timing varies across brain regions, standard modalities, and attended modalities.

Across all models significantly different from the intercept-only model ($p < 0.01$; likelihood-ratio permutation test; see section A.1.12), the median time of the largest peak of decoding model coefficients was 292 ms, with a 95% confidence interval of [284, 300] ms. We grouped the clusters of ICs into five groups, as indicated in Figure 4, and examined the mean time of the largest peak coefficient of models corresponding to ICs in each group of clusters and attended modality (see Figure 6a) and in each group of clusters and standard modality (see Figure 6b). We found that peaks occurred earlier when attention was oriented to the auditory instead of the visual modality (see Figure 6a). This modulation by attention was strongest at the central group of clusters, where the peak of the models' coefficients occurred earlier for auditory than visual standards (see Figure 6b).

A first ANOVA, using data from all models significantly different from the intercept-only model ($p < 0.01$; likelihood-ratio permutation test; see section A.1.12), with time of the largest peak of coefficients as dependent variable, found a significant main effect of attended modality ($F(1, 230) = 4.259$, $p = 0.0402$). A post hoc test showed that the peak occurred earlier for the auditory than the visual attended modalities ($z = 1842$, $p = 0.0327$; see the black asterisk to the left of Figure 6a). A second ANOVA restricted to models corresponding to the visual standard modality found a significant main effect of group of clusters ($F(4, 135) = 4.4073$). A post hoc analysis revealed that the peak was earlier for the occipital than the central group of

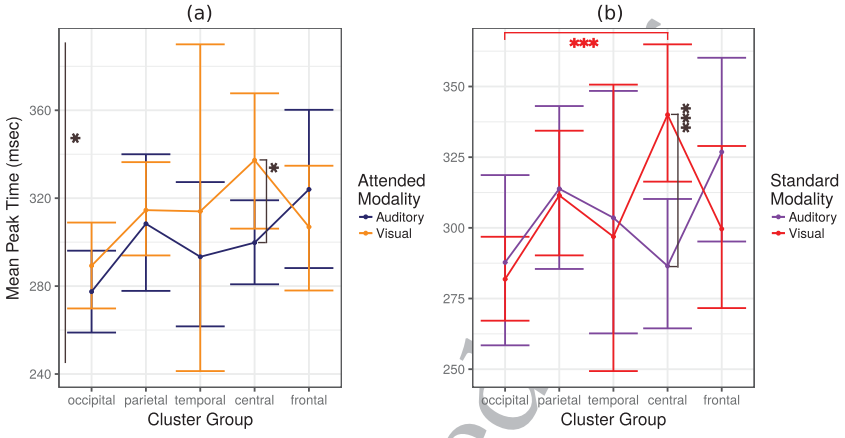


Figure 6: Timing of the SFP effect on ITPC. (a) Mean of coefficients' peak times for models corresponding to auditory (blue points) and visual (orange points) attended modalities, as a function of group of clusters. On average, when attending to audition, the SFP effect on ITPC occurred earlier than when attending to vision. (b) Mean of coefficients' peak times for models corresponding to auditory (violet points) and visual (red points) standard modalities as a function of group of clusters. On average, for visual standards, the SFP effect on ITPC occurred earlier in occipital than in central brain regions. The SFP effect on ITPC was most strongly modulated over the central brain region by both the attended modality (see panel a) and the standard modality (see panel b).

clusters ($z = 4.163$, $p = 3.14\text{e-}05$; see the red asterisks in Figure 6b). A third ANOVA restricted to models corresponding to the central group of clusters found significant main effects of attended modality ($F(1, 43) = 4.5061$, $p = 0.0396$) and of standard modality ($F(1, 43) = 14.2173$, $p = 0.0005$). A post hoc analysis showed that the peak occurred earlier when attention was oriented to the auditory than the visual modality ($z = 2.178$, $p = 0.029360$; black asterisk next to "Central" in Figure 6a) and earlier for auditory than visual standards ($z = 3.406$, $p = 0.000659$; see the black asterisks in Figure 6b).

3.7 Modulations of ITPC Do Not Reflect the Superposition of Evoked Responses by Warning Signals and Subsequent Standards. From the successful decodings of SFPDs from ITPC evoked by standards, we infer that the SFPD is modulating ITPC. Below, we argue that this SFP effect on ITPC is related to temporal expectation (see section 4.1). However, since the amount of evoked response superposition between warning signals and subsequent standards should be larger for standards closer to the warning signal (i.e., for standards with shorter SFPD) than for standards farther away from it (i.e., for standards with larger SFPD), a simpler explanation

of the above results is that ITPC reflects the amount of evoked response superposition. In such a case, higher-order processes, like temporal expectation, may not be needed to explain the previous results. Also, features of the warning signal ERP could be related to subjects' behavioral performance (e.g., the amplitude of the warning signal ERP could be larger for subjects better able to detect deviants). Then, interactions between ERPs of warning signals and following standards should also be related to subjects' behavioral performance. Thus, the simpler evoked-response-superposition explanation could also account for the correlation between the strength of the SFP effect on ITPC and error rates (see Figure 5). In this section, we study this alternative explanation.

We measured ERPs of warning signals and estimated the last significant lag of these ERPs. Then for each original data set (for each subject, component, attended modality, and standard modality), we built a new data set by removing standards that followed the previous warning signal by less than the last significant lag of the warning signal ERP (see section A.1.11). We call the new data set the ERP-superposition-removed data set. We reasoned that if the decoding power of models reflects the amount of evoked response superposition between warning signals and subsequent standards, then the decoding power of models on data sets without such superposition (e.g., on ERP-superposition-removed data sets) should be significantly lower than that on data sets with such superposition (e.g., on original data sets).

Figure 7a plots the decoding power of models estimated from original data sets versus that of models estimated from ERP-superposition-removed data sets, for the midcentral cluster 19 and unattended visual standards. Estimated last significant lags of the warning signal ERP ranged between 200 and 1250 ms (see the inset in Figure 7b). The decoding power of models estimated from original and ERP-superposition-removed data sets was not significantly different (median difference between correlation coefficients for models fitted to original minus that for models fitted to superposition-removed data sets equaled zero with a 95% confidence interval $[-0.05, 0.07]$ that included zero). Figure 7b plots the decoding power of models versus subjects error rates, as in Figure 5, but for models estimated from ERP-superposition-removed data sets. After removing standards whose ERP superimposed with the previous warning signal ERP, there still was a significant, although weaker, correlation between model decodings and subjects' error rates. Figures 7c, 7d, A.12a, and A.12b show a lack of correlation between ERP amplitude, latency, variability, and ITC peak value, respectively, and subjects' error rates, for cluster 19 and unattended visual standards.

That the SFP effect on ITPC can be observed using only standards presented after the last significant lag of the preceding warning signal ERP (see Figure 7a) suggests that the ERP interaction hypothesis does not account for this effect. That the decoding power of models estimated from ERP-superposition-removed data sets is significantly correlated with subjects' error rates (see Figure 7b) shows that the correlations between the strength

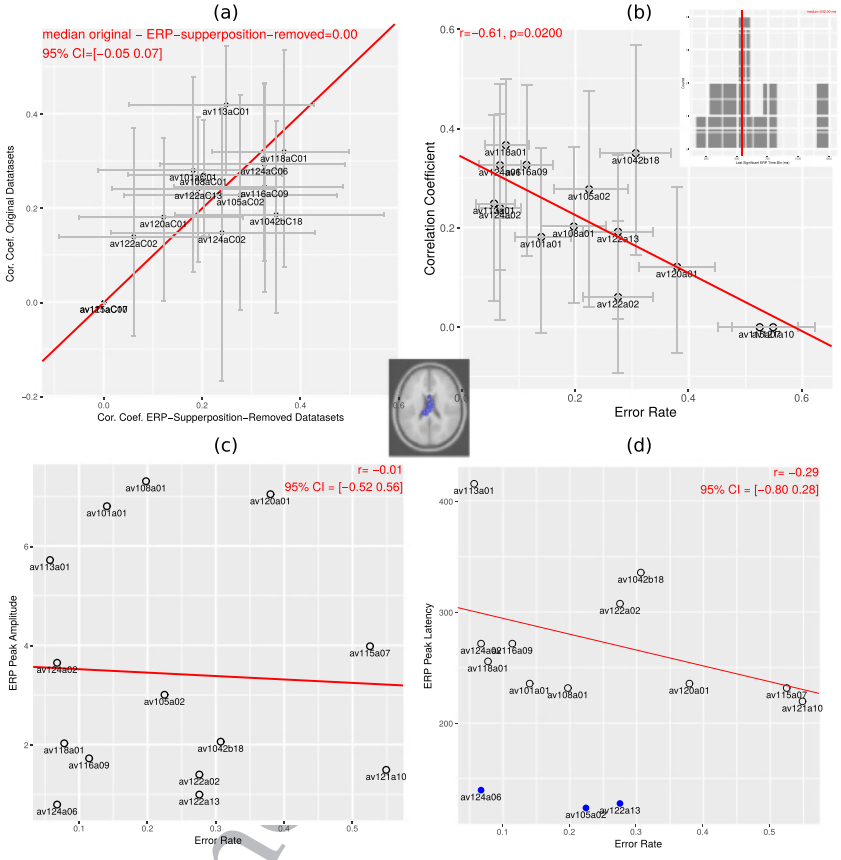


Figure 7: Does the SFP effect on ITPC reflect the superposition of evoked responses from warning signals and subsequent standards? All panels correspond to the mid-central cluster 19 and unattended visual standards. (a) Decoding power of models estimated from original versus ERP-superposition-removed data sets. (b) Decoding power of models estimated from ERP-superposition-removed data sets versus error rates. (c) Peak ERP amplitude versus error rate. (d) Peak ERP latency versus error rates. The inset in panel b is the histogram of the last significant ERP times. The lack of difference in predictive power in panel a, the significant correlation between models' decoding power and subjects' error rates in panel b, and the lack of correlation between ERP features and error rates in panels c and d suggest that the SFP effect on ITPC does not reflect the superposition in evoked responses of warning signals and subsequent standards.

of the SFP effect on ITPC and subjects' error rates shown in Figure 5 are not explained by interactions between ERPs of warning signals and subsequent standards. Additional evidence comes from the observations that ERP peak amplitude, latency, and variability, as well as peak ITC value, are not correlated with subjects' error rates (see Figures 7c, 7d, A.12a, and A.12b). Furthermore, section A.2.3 shows that ERPs from warning signal may hinder the SFP effect on ITPC.

4 Discussion

We have demonstrated the existence of a new foreperiod effect on ITPC and showed that this effect is behaviorally relevant. In an attention-shifting oddball experiment (see Figure 1), we demonstrated that the ITPC evoked by a standard is modulated by the delay between this standard and the preceding attention-shift cue (i.e., warning signal; see Figure 3, clusters with a colored dot in Figure 4, and blue entries in Table A.2). We used a single-trial decoding method to quantify the strength of the new foreperiod effect. We demonstrated that this effect is not an artifact of the decoding method, since it can also be observed using a simple trial averaging method (see Figures 3d to 3i). Both the simple method and the decoding method showed that at some time points, the ITPC of trials closer to the warning signal was significantly larger than that of trial farther away from it, while at some other points, the reverse pattern held. In addition, fluctuations between these two states followed a low frequency ($\sim$1 Hz) sinusoidal pattern (see Figures 3d to 3i, 3m to 3o, and A.11). Importantly, the strength of the SFP effect on ITPC (i.e., the strength of the relation between SFPD and ITPC evoked by standards, or the accuracy in decoding the SFPD from ITPC evoked by a standard) was correlated with subjects' detection performance and possibly with reaction speed (see the examples and summaries of correlations with detection performance in Figure 5, clusters with a double dagger in Figure 4, and blue entries in Table A.3; with reaction speed in clusters with a dagger in Figure 4 and blue entries in Table A.4).

The fact that models accurately decoded the SFP duration from the ITPC evoked by standards does not necessarily mean that these two quantities are correlated with each other. In principle, a sufficiently complex model can decode arbitrarily accurately any experimental variable from any physiological measurements if the complex model overfits the data (Geman, Bienenstock, & Doursat, 1992). Note, however, that the model estimation and evaluation methods used here were designed to avoid overfitting (i.e., we used a regularized error function for parameter estimation in section 2.7 and cross-validation to evaluate the goodness of fit of models in section A.1.12). To further validate that the obtained results were not an artifact of our decoding method, we showed that similar modulations of ITPC were obtained by a simple trial-averaging procedure (see section 3.2). In addition, we developed a control that showed that the SFP effect on ITPC does not reflect

the overlap between the ERPs of the warning signal and the ERPs of ensuing standards. Further evidence for the reliability of our results comes from the significant correlations between the decoding power of the models and subjects' behavioral measures (see section 3.4). We tested the hypothesis that the SFP effect on ITPC is a trivial consequence of interactions between ERPs evoked by warning signals and those evoked by standards, but we did not find support for it (see section 3.7). In this test, we found that four features of the warning signal ERP were not correlated with detectability of deviants. It is thus possible that electrophysiological correlates of detectability of deviants are not present at times of warning signal presentation but that they develop afterward. Our results support this possibility by showing that ITPC evoked by standards presented 1 or more seconds after the warning signal is related to the detectability of forthcoming deviants.

4.1 Visual Temporal Expectation May Generate the SFP Effect on ITPC. Since every warning signal was followed by at least one deviant and SFPs were of variable duration, as time after the presentation of a warning signal advanced without a deviant being presented, the probability of a deviant presentation increased. Subjects probably learned this probability function (i.e., hazard function) and adjusted their temporal expectation for a deviant according to this function. We suggest that temporal expectation plays an important role in the generation of the SFP effect on ITPC, although further investigations are needed to validate this suggestion, as discussed in section 4.6. First, the strongly periodic stimulation in our experiment should have induced temporal expectation in subjects. This claim is supported by the observation of reliable foreperiod effects on reaction times in 26% of the subjects and attended modalities (see section 3.1) and by previous arguments relating the foreperiod effect on reaction times to temporal expectation (Walter et al., 1964; Niemi & Näätänen, 1981; Los, 2010). Second, temporal expectation contributes to faster reaction times and improved perception (Correa, 2010), while the strength of the SFP effect on ITPC was larger in subjects achieving larger detection rates and apparently in subjects reacting the fastest (see Figure 5 and Tables A.4 and A.3). Third, as we discuss in section 4.3, the mid-central cluster 19, where we observed significant correlations between the strength of the SFP effect on ITPC and error rates, has been implicated in temporal expectation.

An unresolved question in the field of temporal expectation is whether it affects motor (i.e., rate of motor response), premotor (i.e., response preparation), perceptual (i.e., buildup of information about a stimulus), or executive (i.e., decision mechanism) stages. The majority of the evidence suggests a motor influence (Sanders, 2013; Brunia & Boelhouwer, 1988; Coull & Nobre, 1998), premotor effects have also been reported (Bausenhart, Rolke, Hackley, & Ulrich, 2006; Hackley & Valle-Inclán, 1999; Müller-Gethmann, Ulrich, & Rinkenauer, 2003), and more recent studies have shown influences on perceptual (Correa, Lupiáñez, & Tudela, 2005; Mento, Tarantino, Sarlo, &

Bisiacchi, 2013; Lange, 2009; Rolke, 2008) and executive (Naccache, Blandin, & Dehaene, 2002; Correa, Cappucci, Nobre, & Lupiáñez, 2010) functions. Results from our study suggest that the SFP effect on ITPC is related to late visual processing stages. That the strength of the new foreperiod effect is significantly correlated with stimulus detectability (see Figure 5) and that this effect is calculated from EEG evoked by standards, to which subjects did not respond, indicates that the SFP effect on ITPC is not directly related to motor stages. Also, the long latencies after the presentation of standards at which modulations of the SFP were largest (the median time of the largest peak of the regression coefficients was 292 ms; see section 3.6) argues against relations with early sensory stages. Finally, that the new foreperiod effect was substantially stronger for the visual than for the auditory standard modality suggests that it may be specific to visual processing.

Three pieces of evidence indicate that the new foreperiod effect is stronger for the visual than the auditory standard modality. First, as shown in section 3.3, Figure 4, and the colored cells in Table A.4, the number of clusters with a large proportion of models with decodings significantly correlated with SFPDs was larger for the visual standard modality. Second, as indicated in section 3.4, Figure 4, and Tables A.4 and A.3, the number of clusters showing significant correlations between models' decodings and behavioral measures, as well as the strength of these correlations, was larger for the visual than the auditory standard modality. And third, for the visual standard modality only, clusters where the decoding of more than 40% of models was significantly correlated with SFPDs are localized in visual brain regions (i.e., parieto-occipital cortex). An interesting behavioral correlate of this larger strength of the SFP effect on ITPC for the visual modality is the superior detectability and reaction speed of subjects to visual than auditory deviants (see section 3.1). Note that the primary auditory cortex (Brodmann area 41 on the Heschl gyrus) is buried inside the lateral sulcus, farther away from the scalp than the primary visual cortex (Brodmann area 17). Thus, to identify deeper auditory cortical sources, one needs higher-density EEG recordings than to identify visual cortical sources. Hence, the weakness of the SFP effect on ITPC on the auditory standard modality could be due to the fact that our 32-channel EEG recording system may not have covered auditory cortex with sufficient density to identify its deeper sources.

4.2 Differences with Previous Studies on Temporal Expectation. The new foreperiod effect on ITPC is different in three main ways from influences of temporal expectation on ITPC reported in previous studies (Busch, Dubois, & VanRullen, 2009; Mathewson et al., 2009; Stefanics et al., 2010; Besle et al., 2011; Cravo, Rohenkohl, Wyart, & Nobre, 2011; Mathewson et al., 2012; Cravo et al., 2013; Wilsch, Henry, Herrmann, Maess, & Obleser, 2015; Ten Oever, van Atteveldt, & Sack, 2015; van Diepen, Cohen, Denys, & Mazaheri, 2015).

First, it is measured on a single-trial measure of ITPC, the DMP. Most studies relating ITPC to expectation are built around the ITC. Other measures of ITPC have been used (Stefanics et al., 2010; Cravo et al., 2013; Mathewson et al., 2009), but they also average information across trials. Furthermore, some of these analyses averaged data across subjects (Busch et al., 2009; Mathewson et al., 2009, 2012; Stefanics et al., 2010; Besle et al., 2011; Cravo et al., 2011, 2013; Wilsch et al., 2015; Ten Oever et al., 2015; van Diepen et al., 2015). A potential danger is that these averages may miss critical information at the single-trial or single-subject level.

Second, the new foreperiod effect is observed on ITPC evoked by standards. Recent studies have investigated how temporal expectation influences ITPC on spontaneous oscillations (Busch et al., 2009; Mathewson et al., 2009; Stefanics et al., 2010; Besle et al., 2011; Cravo et al., 2011, 2013; Wilsch et al., 2015; Ten Oever et al., 2015; van Diepen et al., 2015) and on oscillations induced by preceding rhythmic stimulation (Cravo et al., 2013; Mathewson et al., 2012). To our knowledge, this is the first study on influences of temporal expectation on evoked ITPC.

Third, previous investigations studied how expectation influences ITPC at a single time point, for example, immediately before the presentation of a target (Busch et al., 2009; Mathewson et al., 2009, 2012; Stefanics et al., 2010; Cravo et al., 2011, 2013; Ten Oever et al., 2015; van Diepen et al., 2015) or standard (Lakatos et al., 2008; Besle et al., 2011; Wilsch et al., 2015) stimuli. Overall, these studies found that ITPC increases prior to the presentation of expected stimuli so that the phase of oscillations across trials is concentrated around a value of maximal excitability. Our multivariate decoding method allowed us to examine how expectation influences ITPC across a 500 ms period following the presentation of standards scattered between warning signals and the first following deviant. We found that the influence of expectation on ITPC is not constant or monotonic over time, but varies sinusoidally at a low frequency around 1 Hz (see Figures 3m to 3o). These sinusoidal modulations are not an artifact of our decoding method, since almost identical oscillations were observed using the simple averaging method (see Figures 3g to 3i and A.11) and may be related to fluctuations in subjects' detectability of visual stimuli (Fiebelkorn et al., 2011).

4.3 Similarities with Previous Studies on Temporal Expectation. Electrophysiological correlates of temporal expectation have mainly been investigated using the CNV (Walter et al., 1964). It is typically recorded in experiments with a fixed foreperiod of length between 1 and 1.5 seconds. With longer fixed foreperiod lengths, 8 and 6 seconds, Weerts and Lang (1973) and Loveless and Sanford (1975) reported that the CNV is composed of an orienting wave, or O wave, that succeeds the warning signal by approximately 1 second, and the expectancy wave, or E wave, which rises in anticipation to the imperative stimulus. Studies demonstrating that the CNV amplitude is maximal at the vertex (Walter, 1967), that the

supplementary motor area (SMA) and the anterior cingulate cortex (ACC) are the sources of the CNV's O wave (Cui et al., 2000; Zappoli et al., 2000; Gomez et al., 2001), and that the premotor cortex appears to give rise to the CNV's E wave (Hultin et al., 1996) show that the central brain region associated with cluster 19 (see Figures 4 and A.9) is related to the CNV, and therefore to temporal expectation (Walter et al., 1964). Coincidentally, the SFP effect on ITPC was most strongly correlated with behavior at cluster 19 (see section 3.4), and its timing was most strongly influenced by attention and by the standard modality (see section 3.6) at this cluster.

4.4 Merits of the Decoding over the Simple Trial-Averaging Method to Detect the SFP Effect on ITPC. We used two different methods to demonstrate the existence of the SFP effect on ITPC. In Section 3.2, we applied a simple method comparing the mean DMP between trials closer to and farther away from the warning signal, and in section 3.3 we used a decoding method. Here we comment on some benefits of the latter method.

The main advantage of the decoding method is that it yields a sound number that quantifies the strength of the SFP effect on ITPC for a subject, standard modality, and attended modality (i.e., the correlation coefficient between decodings and experimental SFPDs; see the red annotations in Figures 3j to 3l). We can then compare this number with behavioral measures for this subject and attended modality, such as error rates and mean reaction times, to assess the behavioral relevance of the new foreperiod effect (as we did in Figure 5).

The interpretability of decoding model coefficients (see section 3.5) is especially relevant when combined with the previous quantification advantage of the decoding method. By both interpreting decoding models coefficients in Figures 3m to 3o and by observing averages of DMP from the simple method in Figures 3d to 3i and A.11, we learned that at some time points after the presentation of standards evoked, ITPC was larger for standards closer to than farther away from the warning signal, that at some other time points the reverse pattern held, and that the alternation between these two states followed a low-frequency ($\sim$1 Hz) sinusoidal pattern, independent of the frequency at which ITPC was measured. Next, the strong correlations between the strength of the SFP effect on ITPC and subjects' behaviors (see Figure 5 and Tables A.3 and A.4) indicated that this alternation was behaviorally relevant. Note that this behavioral relevance cannot be obtained using the simple method.

Another advantage is that the decoding method is statistically more powerful than the simple method because the former method uses all trials in a statistically optimal way, and not just extreme trials in an ad hoc manner.

4.5 A Head-to-Head Comparison with a Previous Study Highlights Unique Aspects of the Current Analysis. An investigation by Busch et al.

(2009) showed that visual stimuli detectability is related to prestimulus phase. Below we highlight a few differences between this study and ours. The purpose of this comparison is not to belittle the research by Busch et al. (2009), which we believe was excellent, but to emphasize a few novel characteristics of our analysis. First, ITC and the phase coherence index used in the main text of Busch et al. (2009) are phase coherence measures for groups of trials, while the DMP used in this letter is a phase coherence measure for single trials. In supplementary Figure 2, Busch et al. (2009) reported results of using a classifier to predict in single trials if a stimulus will be detected based on the phase at a single time-frequency point. It is a pity that after estimating these single-trial classifiers, Busch et al. (2009) averaged their outputs across subjects and did not attempt to relate the behavior of subjects with the predictive power of the corresponding models, as we did in Figure 5. The fact that after averaging predictions of classifiers of different subjects, they found a correlation between these averaged predictions and detection performance of subjects is remarkable. However, the analysis method is limited by using only one time-frequency point to predict the detectability of a stimulus. The models we used in this letter used a 500 ms window of DMP to decode the SFP duration. That is, the second difference between the study by Busch et al. (2009) and ours is that they used univariate models, while we used multivariate ones. Note that we have optimized models to decode SFP durations and then found significant correlations between the decoding power of these models and behavior without fitting model parameters to behavioral data. That a model optimized to predict behavior from a stimulus feature (e.g., ITPC) achieves reliable predictions does not necessarily mean that the stimulus feature is relevant to the behavior, since the model can overfit the data (Geman et al., 1992). Although there are methods to minimize the risks of overfitting, such as using separate pieces of data to estimate parameters of a model and to evaluate its predictive power (the method used in Busch et al., 2009), this risk never disappears when optimizing a model to predict behavior. Thus, the third difference is that we did not fit parameters to behavioral data while Busch et al. (2009) did. The fourth difference is that the study by Busch et al. (2009) averaged data across subjects, while we analyzed different subjects separately. This single-subject analysis allowed us to observe that subjects whose ITPC was more strongly modulated by the SFPD were those who performed better. Fifth, Busch et al. (2009) reported the analysis of a single electrode (Fz), while our study described the analysis of ICs over the whole brain. It is important to report effects across the whole brain to demonstrate the generality or specificity of a hypothesis tested at a single brain region. Sixth, Busch et al. (2009) analyzed data from EEG channels, while we analyzed data from ICs. Applied to EEG, ICA finds maximally independent sources generating recorded potentials. Scalp representations of these sources resemble fields generated by current dipoles inside the brain, and

biological arguments suggest that these maximally independent sources reflect the synchronized activations of neurons in compact cortical regions (Delorme, Palmer, Onton, Oostenveld, & Makeig, 2012). Thus, our analysis most probably characterized cortical sources, while that of Busch et al. (2009) described mixtures of these sources. An important advantage of the study by Busch et al. (2009) over ours is that for each subject separately, they calibrated visual stimuli so that the detection rate of subjects was close to 50%. No such individualized calibration was performed in our experiment, and the averaged detection rate across subjects was 86%. Having a similar number of trials where subjects perceived and failed to perceive stimuli is expedient to find neural correlates of perception. However, the fact that the methods we used could find reliable phase correlates of stimulus detection with subthreshold stimulation shows that these methods could be applied to a larger number of more naturalistic EEG experiments where stimulation was not optimized for each subject. Another advantage of the study by Busch et al. (2009) over ours is that they investigated modulations by temporal expectation on both amplitude and phase of EEG oscillations, while we studied only modulations on phase. A promising line for future research is to use predictive models with both phase and amplitude regressors to simultaneously evaluate the relative importance of amplitude, phase and their interaction to decode SFPDs.

4.6 Future Work. Here we suggested that the SFP effect on ITPC is related to temporal expectation, but further research is needed to support this suggestion. We will perform new experiments to test what and how mental processes modulate the SFP effect on ITPC. For this, we will build on thoughtful investigations on mental processes associated with the CNV. Is the SFP effect on ITPC related to expectation (Walter et al., 1964)? Would it disappear if responses to deviants are not required? Would it be attenuated if a proportion of deviants does not require response? Is the SFP effect on ITPC linked to motor preparation (Gaillard, 1977, 1978)? Does the emergence of this effect require motor responses? What is the relation between the SFP effect on ITPC and intention to respond (Low, Borda, Frost, & Kellaway, 1966)? Would the strength of this effect be directly proportional to the anticipated force needed in the motor response? Is the SFP effect on ITPC related to motivation (McCallum & Walter, 1968)? Would the strength of this effect be augmented when subjects are instructed to concentrate hard and respond quickly to deviants? How does this effect relate to arousal (Tecce, 1972)? Would the relation between arousal and this effect be U-shaped, as is the relation between arousal and CNV amplitude (Tecce, 1972, Figure 6b)? Is the SFP effect on ITPC linked to explicit (Macar & Vidal, 2003; Pfeuty, Ragot, & Pouthas, 2003, 2005) or implicit (Praamstra, Kourtis, Kwok, & Oostenveld, 2006) time perception? Would this effect reflect the perceived duration

of a target interval when subjects compare it to the duration of a memorized interval?

5 Summary

Woodrow (1914) demonstrated how context, in the form of the foreperiod of a stimulus, influences reaction times. Here we described a new effect of this same context on intertrial alignment of EEG oscillations evoked by standards. At some times after the presentation of standards, intertrial alignment was larger for trials with shorter than longer foreperiods, while at other times, the reverse held, and fluctuations between these two states followed a low-frequency sinusoidal pattern. Importantly, this effect is behaviorally relevant, since the stronger is the influence of foreperiod on intertrial alignment for a subject (as indicated by a decoding model), the better is his or her behavioral performance (error rates and possibly mean reaction times). We speculated that the SFP effect on ITPC is related to temporal expectation, and future experiments will study this speculation.

Acknowledgments

We thank Steve Hillyard for suggesting Figures 3d to 3f, and to investigate the influence of the maximum SFPD in the decodings of the models (Figure A.2), Eduardo Martínez-Montes and Mateuz Gola for comments on earlier versions of this letter, Cyril Pernet for recommending the use of robust correlation coefficients, and Makoto Miyakoshi for advice on source localization. J.R. thanks Scott Makeig for frequent discussions and the Biblioteca Nacional de Maestros in Buenos Aires, Argentina, for providing a comfortable environment to prepare this letter. We also thank Robert Buffington for computer assistance. We are grateful for the comments from anonymous reviewers that helped us find that the SFP effect on ITPC might be stronger than reported here (see Figure A.5), and that led us to detect a new apparent correlation between the strength of the effect and subjects' error rates (see Figure A.10). Most of the research in this letter was performed with free software. Computations were done with R (R Core Team, 2012), the article was written in LaTeX (Lamport, 1994), figures were prepared with Inkscape (Inkscape Team, 2014), all in a personal computer and in clusters of personal computers running Linux (Torvalds, 2008). The only paid software used in this investigation was Matlab (R 2013), employed solely for initial preprocessing with ICA (Makeig, Bell, Jung, & Sejnowski, 1996) and for plotting ICs' ECDs (e.g., see the inset in Figures 3a to 3c) and clusters of these ECDs (e.g., see Figure 4). This research was partially supported by grant NIH/NIA RO1 AG18030 awarded to J.T. There are no known conflicts of interest associated with this publication.

References

Anguera, J. A., Lyman, K., Zanto, T. P., Bollinger, J., & Gazzaley, A. (2013). Reconciling the influence of task-set switching and motor inhibition processes on stop signal after-effects. *Frontiers in Psychology*, *4*.

Barne, L. C., Claessens, P. M. E., Reyes, M. B., Caetano, M. S., & Cravo, A. M. (2017). Low-frequency cortical oscillations are modulated by temporal prediction and temporal error coding. *NeuroImage*, *146*, 40–46.

Bausenhart, K. M., Rolke, B., Hackley, S. A., & Ulrich, R. (2006). The locus of temporal preparation effects: Evidence from the psychological refractory period paradigm. *Psychonomic Bulletin and Review*, *13*, 536–542.

Belsley, D. A., Kuh, E., & Welsch, R. E. (2004). *Regression diagnostics: Identifying influential data and sources of colinearity*. Hoboken, NJ: Wiley-Interscience.

Besle, J., Schevon, C. A., Mehta, A. D., Lakatos, P., Goodman, R. R., McKhann, G. M., . . . Schroeder, C. E. (2011). Tuning of the human neocortex to the temporal dynamics of attended events. *Journal of Neuroscience*, *31*(9), 3176–3185.

Bishop, C. M. (2016). *Pattern recognition and machine learning*. New York: Springer-Verlag.

Botwinick, J., & Brinley, J. F. (1962). An analysis of set in relation to reaction time. *Journal of Experimental Psychology*, *63*(6), 568–574. doi:10.3389/fpsyg.2013.00649

Brunia, C. H. M., & Boelhouwer, A. J. (1988). Reflexes as a tool: A window in the central nervous system. In P. Ackles, J. Jennigs, & M. Coles (Eds.), *Advances in psychophysiology* (vol. 3, pp. 1–67). Greenwich, CT: JAI Press.

Busch, N. A., Dubois, J., & VanRullen, R. (2009). The phase of ongoing EEG oscillations predicts visual perception. *Journal of Neuroscience*, *29*(24), 7869–7876.

Ceponiene, R., Westerfield, M., Torki, M., & Townsend, J. (2008). Modality-specificity of sensory aging in vision and audition: Evidence from event-related potentials. *Brain Research*, *1215*, 53–68.

Correa, A. (2010). Enhancing behavioural performance by visual temporal orienting. In A. Nobre & J. Coull (Eds.), *Attention and time* (pp. 359–370). Oxford: Oxford University Press.

Correa, Á., Cappucci, P., Nobre, A. C., & Lupiáñez, J. (2010). The two sides of temporal orienting: Facilitating perceptual selection, disrupting response selection. *Experimental Psychology*, *57*(2), 142–148.

Correa, Á., Lupiáñez, J., & Tudela, P. (2005). Attentional preparation based on temporal expectancy modulates processing at the perceptual-level. *Psychonomic Bulletin and Review*, *12*(2), 328–334.

Coull, J. T., & Nobre, A. C. (1998). Where and when to pay attention: The neural systems for directing attention to spatial locations and to time intervals revealed by both PET and fMRI. *Journal of Neuroscience*, *18*(18), 7426–7435.

Cravo, A. M., Rohenkohl, G., Wyart, V., & Nobre, A. C. (2011). Endogenous modulation of low frequency oscillations by temporal expectations. *Journal of Neurophysiology*, *106*(6), 2964–2972.

Cravo, A. M., Rohenkohl, G., Wyart, V., & Nobre, A. C. (2013). Temporal expectation enhances contrast sensitivity by phase entrainment of low-frequency oscillations in visual cortex. *Journal of Neuroscience*, *33*(9), 4002–4010.

Cravo, A. M., Santos, K. M., Reyes, M. B., Caetano, M. S., & Claessens, P. M. (2015). Visual causality judgments correlate with the phase of alpha oscillations. *Journal of Cognitive Neuroscience*, *27*, 1887–1894.

Cui, R., Egkher, A., Huter, D., Lang, W., Lindinger, G., & Deecke, L. (2000). High resolution spatio-temporal analysis of the contingent negative variation in simple or complex motor tasks and a non-motor task. *Clinical Neurophysiology*, *111*, 1847–1859.

de Graaf, T. A., Gross, J., Paterson, G., Rusch, T., Sack, A. T., & Thut, G. (2013). Alpha-band rhythms in visual task performance: Phase-locking by rhythmic sensory stimulation. *PLoS One*, *8*(3), e60035.

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21.

Delorme, A., Palmer, J., Onton, J., Oostenveld, R., & Makeig, S. (2012). Independent EEG sources are dipolar. *PLoS One*, *7*(2).

Drewes, J., & VanRullen, R. (2011). This is the rhythm of your eyes: The phase of ongoing electroencephalogram oscillations modulates saccadic reaction time. *Journal of Neuroscience*, *31*(12), 4698–4708.

Fiebelkorn, I., Foxe, J., Butler, J., Mercier, M., Snyder, A., & Molholm, S. (2011). Ready, set, reset: Stimulus-locked periodicity in behavioral performance demonstrates the consequences of cross-sensory phase reset. *Journal of Neuroscience*, *31*(27), 9971–9981.

Fisher, N. I. (1996). *Statistical analysis of circular data*. Cambridge: Cambridge University Press.

Gaillard, A. W. K. (1977). The late CNV wave: Preparation versus expectancy. *Psychophysiology*, *14*, 563–568.

Gaillard, A. W. K. (1978). *Slow brain potential preceding task performance*. Amsterdam: Academische Press.

Geman, S., Bienenstock, E., & Doursat, R. (1992). Neural networks and the bias/variance dilemma. *Neural Computation*, *4*, 1–58.

Gomez, C., Delinte, A., Vaquero, E., Cardoso, M., Vazquez, M., Crommelinck, M., & Roucoux, A. (2001). Current source density analysis of CNV during temporal gap paradigm. *Brain Topography*, *13*, 149–159.

Gray, M. J., Frey, H.-P., Wilson, T. J., & Foxe, J. J. (2015). Oscillatory recruitment of bilateral visual cortex during spatial attention to competing rhythmic inputs. *Journal of Neuroscience*, *35*(14), 5489–5503.

Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N., & Altman, D. G. (2016). Statistical tests, *p* values, confidence intervals, and power: A guide to misinterpretations. *European Journal of Epidemiology*, *31*(4), 337–350.

Hackley, S. A., & Valle-Inclán, F. (1999). Accessory stimulus effects on response selection: Does arousal speed decision making? *Journal of Cognitive Neuroscience*, *11*, 321–329.

Hamm, J. P., Dyckman, K. A., Ethridge, L. E., McDowell, J. E., & Clementz, B. A. (2010). Preparatory activations across a distributed cortical network determine production of express saccades in humans. *Journal of Neuroscience*, *30*(21), 7350–7357.

Hanslmayr, S., Aslan, A., Staudigl, T., Klimesch, W., Herrmann, C. S., & Bäuml, K.-H. (2007). Prestimulis oscillations predict visual perception performance between and within subjects. *Neuroimage*, *37*, 1465–1473.

Hanslmayr, S., Volberg, G., Wimber, M., Dalal, S. S., & Greenlee, M. W. (2013). Prestimulus oscillatory phase at 7 Hz gates cortical information flow and visual perception. *Current Biology*, *23*(22), 2273–2278.

Hastie, T., Tibshirani, R., & Friedman, J. (2016). *The elements of statistical learning* (2nd ed.). New York: Springer.

Henry, M. J., & Obleser, J. (2012). Frequency modulation entrains slow neural oscillations and optimizes human listening behavior. *Proceedings of the National Academy of Sciences*, *109*(49), 20095–20100.

Hickok, G., Farahbod, H., & Saberi, K. (2015). The rhythm of perception entrainment to acoustic rhythms induces subsequent perceptual oscillation. *Psychological Science*, *26*, 1006–1013.

Hillyard, S. A., & Galambos, R. (1967). Effects of stimulus and response contingencies on a surface negative slow potential shift in man. *Electroencephalography and Clinical Neurophysiology*, *22*, 297–304.

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, *12*(1), 55–67.

Horton, C., Srinivasan, R., & DZmura, M. (2014). Envelope responses in single-trial EEG indicate attended speaker in a cocktail party. *Journal of Neural Engineering*, *11*(4), 046015.

Hultin, L., Rossini, P., Romani, G., Högstedt, P., Tecchio, F., & Pizzella, V. (1996). Neuromagnetic localization of the late component of the cognitive negative variation. *Electroencephalography and Clinical Neurophysiology*, *98*, 435–448.

Inkscape Team. (2004). *Inkscape*. www.kernal.org

Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2005). *Applied linear statistical models* (5th ed.). New York: McGraw-Hill/Irwin.

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science*, *320*, 110–113.

Lakatos, P., Musacchia, G., OConnel, M. N., Falchier, A. Y., Javitt, D. C., & Schroeder, C. E. (2013). The spectrotemporal filter mechanism of auditory selective attention. *Neuron*, *77*(4), 750–761.

Lamport, L. (1994). *LaTeX: A document preparation system* (2nd ed.). Reading, MA: Addison-Wesley.

Lancaster, J., Rainey, L., Summerlin, J., Freitas, C., Fox, P., Evans, A., . . . Mazziotta, J. (1997). Automated labeling of the human brain: A preliminary report on the development and evaluation of a forward-transform method. *Human Brain Mapping*, *5*, 238–242.

Lancaster, J. L., Woldorff, M. G., Parsons, L. M., Liotti, M., Freitas, C. S., Rainey, L., Kochunov, . . . Fox, P. T. (2000). Automated Talairach atlas labels for functional brain mapping. *Human Brain Mapping*, *10*, 12–131.

Lange, K. (2009). Brain correlates of early auditory processing are attenuated by expectations for time and pitch. *Brain and Cognition*, *69*(1), 127–137.

Lindsley, D. B. (1952). Psychological phenomena and the electroencephalogram. *Electroencephalography and Clinical Neurophysiology*, *4*(4), 443–456.

Los, S. A. (2010). Foreperiod and sequential effects: Theory and data. *Attention and Time*, *289*(302), 24.

Loveless, N., & Sanford, A. (1975). The impact of warning signal intensity on reaction time and components of the contingent negative variation. *Biological Psychology*, *2*(3), 217–226.

Low, M. D., Borda, R. P., Frost, J. D., & Kellaway, P. (1966). Surface negative slow potential shift associated with conditioning in man. *Neurology*, *16*, 771–782.

Luo, H., & Poeppel, D. (2007). Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron*, *54*(6), 1001–1010.

Macar, F., & Vidal, F. (2003). The CNV peak: An index of decision making and temporal memory. *Psychophysiology*, *40*, 950–954.

Makeig, S., Bell, A. J., Jung, T.-P., & Sejnowski, T. J. (1996). Independent component analysis of electroencephalographic data. In D. Touretzky, M. Mozer, & M. Hasselmo (Eds.), *Advances in neural information processing systems*, *8*. Cambridge MA: MIT Press.

Makeig, S., Debener, S., Onton, J., & Delorme, A. (2004). Mining event-related brain dynamics. *Trends in Cognitive Sciences*, *8*(5), 204–210.

Mathewson, K. E., Fabiani, M., Gratton, G., Beck, D. M., & 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., & Ro, T. (2009). To see or not to see: Prestimulus *alpha* phase predicts visual awareness. *Journal of Neuroscience*, *29*(9), 2725–2732.

Mathewson, K. E., Prudhomme, C., Fabiani, M., Beck, D. M., Lleras, A., & Gratton, G. (2012). Making waves in the stream of consciousness: Entraining oscillations in EEG alpha and fluctuations in visual awareness with rhythmic visual stimulation. *Journal of Cognitive Neuroscience*, *24*(12), 2321–2333.

Matlab. (R 2013). Version 8.1.0.604. Natick, MA: MathWorks.

McCallum, W., & Walter, W. G. (1968). The effects of attention and distraction on the contingent negative variation in nomral and neurotic subjects. *Electroencephalography and Clinical Neurophysiology*, *25*, 319–329.

Mento, G., Tarantino, V., Sarlo, M., & Bisiacchi, P. S. (2013). Automatic temporal expectation: A high-density event-related potential study. *PlOs One*, *8*(5), 1–11.

Metropolis, N., & Ulam, S. (1949). The Monte Carlo method. *Journal of the American Statistical Asociation*, *44*(247), 335–341.

Milton, A., & Pleydell-Pearce, C. W. (2016). The phase of pre-stimulus alpha oscillations influences the visual perception of stimulus timing. *NeuroImage*, *133*, 53–61.

Monto, S., Palva, S., Voipio, J., & Palva, J. M. (2008). Very slow EEG fluctuations predict the dynamics of stimulus detection and oscillation amplitudes in humans. *Journal of Neuroscience*, *28*(33), 8268–8272. doi:10.1523/JNEUROSCI.1910-08.2008

Müller-Gethmann, H., Ulrich, R., & Rinkenauer, G. (2003). Locus of the effect of temporal preparation: Evidence from the lateralized readiness potential. *Psychophysiology*, *40*, 597–611.

Naccache, L., Blandin, E., & Dehaene, S. (2002). Unconscious masked priming depends on temporal attention. *Psychological Science*, *13*(5), 416–424.

Ng, B. S. W., Schroeder, T., & Kayser, C. (2012). A precluding but not ensuring role of entrained low-frequency oscillations for auditory perception. *Journal of Neuroscience*, *32*(35), 12268–12276.

Niemi, P., & Näätänen, R. (1981). Foreperiod and simple reaction time. *Psychological Bulletin*, *89*(1), 133–162.

Nunez, M. D., Vandekerckhove, J., & Srinivasan, R. (2016). How attention influences perceptual decision making: Single-trial EEG correlates of drift-diffusion model parameters. *Journal of Mathematical Psychology*, *76*, 117–130.

O'Sullivan, J. A., Power, A. J., Mesgarani, N., Rajaram, S., Foxe, J. J., Shinn-Cunningham, B. G., . . . Lalor, E. C. (2015). Attentional selection in a cocktail party environment can be decoded from single-trial EEG. *Cerebral Cortex*, *25*(7), 1697–1706.

Palmer, J. A., Kreutz-Delgado, K., Rao, B. D., & Makeig, S. (2007). Modeling and estimation of dependent subspaces with non-radially symmetric and skewed densities. In M. Davies, C. James, A. Abdallah, & M. Plumbey (Eds.), *Lecture Notes in Computer Science: Proceedings of the 7th International Symposium on Independent Component Analysis*. Berlin: Springer.

Pernet, C. R., Wilcox, R. R., & Rousselet, G. A. (2013). Robust correlation analyses: False positive power validation using a new open source Matlab toolbox. *Frontiers in Psychology*, *3*.

Pfeuty, M., Ragot, R., & Pouthas, V. (2003). When time is up: CNV time course differentiates the roles of the hemispheres in the discrimination of short tone durations. *Experimental Brain Research*, *151*, 372–379.

Pfeuty, M., Ragot, R., & Pouthas, V. (2005). Relation between the CNV and timing of an upcoming event. *Neuroscience Letters*, *382*, 106–111.

Praamstra, P., Kourtis, D., Kwok, H. F., & Oostenveld, R. (2006). Neurophysiology of implicit timing in serial choice reaction-time performance. *Journal of Neuroscience*, *26*, 5448–5455.

R Core Team. (2012). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org

Rapela, J. (2017). Derivation of variational Bayes linear regression. http://sccn.ucsd.edu/∼rapela/docs/vblr.pdf

Rapela, J., Westerfield, M., Townsend, J., & Makeig, S. (2016). *A new single-trial foreperiod effect on inter-trial phase coherence. Part I: existence and relevance*. arXiv:1611.00313.

Rolke, B. (2008). Temporal perparation facilitates perceptual identification of letters. *Perception and Psychophysics*, *70*, 1305–1313.

Romei, V., Gross, J., & Thut, G. (2012). Sounds reset rhythms of visual cortex and corresponding human visual perception. *Current Biology*, *22*, 807–813.

Samaha, J., Bauer, P., Cimaroli, S., & Postle, B. R. (2015). Top-down control of the phase of alpha-band oscillations as a mechanism for temporal prediction. *Proceedings of the National Academy of Sciences*, *112*(27), 8439–8444.

Sanders, A. F. (2013). *Elements of human performance: Reaction processes and attention in human skill*. Mahwah, NJ: Erlbaum.

Sherman, M. T., Kanai, R., Seth, A. K., & VanRullen, R. (2016). Rhythmic influence of top-down perceptual priors in the phase of prestimulus occipital alpha oscillations. *Journal of Cognitive Neuroscience*, *28*, 1318–1380.

Simon, D. M., & Wallace, M. T. (2017). Rhythmic modulation of entrained auditory oscillations by visual inputs. *Brain Topography*, *30*(5), 1–14.

Spaak, E., de Lange, F. P., & Jensen, O. (2014). Local entrainment of alpha oscillations by visual stimuli causes cyclic modulation of perception. *J. Neurosci.*, *34*(10), 3536–3544.

Stefanics, G., Hangya, B., Hernádi, I., Winkler, I., Lakatos, P., & Ulbert, I. (2010). Phase entrainment of human delta oscillations can mediate the effects of expectation on reaction speed. *Journal of Neuroscience*, *30*(41), 13578–13585.

Tallon-Baudry, C., Bertrand, O., Delpuech, C., & Pernier, J. (1996). Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. *Journal of Neuroscience*, *16*, 4240–4349.

Tecce, J. J. (1972). Contingent negative variation (CNV) and psychological processes in man. *Psychological Bulletin*, *77*(2), 73–108.

Ten Oever, S., van Atteveldt, N., & Sack, A. T. (2015). Increased stimulus expectancy triggers low-frequency phase reset during restricted vigilance. *Journal of Cognitive Neuroscience*, *27*, 1811–1822.

Thorne, J. D., De Vos, M., Viola, F. C., & Debener, S. (2011). Cross-modal phase reset predicts auditory task performance in humans. *Journal of Neuroscience*, *31*(10), 3853–3861.

Torvalds, L. (2008). *The Linux kernel*. http://kernal.org

Valera, F. J., Toro, A., John, E. R., & Schwartz, E. L. (1981). Perceptual framing and cortical alpha rhythm. *Neuropsychologia*, *19*(5), 675–686.

van Diepen, R. M., Cohen, M. X., Denys, D., & Mazaheri, A. (2015). Attention and temporal expectations modulate power, not phase, of ongoing alpha oscillations. *Journal of Cognitive Neuroscience*, *27*, 1573–1586.

Walter, W. G. (1967). Slow potential changes in the human brain associated with expectancy, decision, and intention. In W. Cobb & C. Morocutti (Eds.), *The evoked potentials*. Amsterdam: Elsevier.

Walter, W. G., Cooper, R., Aldridge, V., McCallum, W., & Winter, A. (1964). Contingent negative variation: An electrical sign of sensorimotor association and expectancy in the human brain. *Nature*, *203*, 380–384.

Weerts, T. C., & Lang, P. J. (1973). The effects of eye fixation and stimulus and response location on the contingent negative variation (CNV). *Biological Psychology*, *1*(1), 1–19.

Westfall, P. H., & Young, S. S. (1993). *Resampling-based multiple testing: Examples and methods for p-value adjustment*. New York: Wiley.

Wilcox, R. (2012). *Introduction to robust estimation and hypothesis testing*. Orlando, FL: Academic Press.

Wilcox, R. (2017). *Statistical functions*. http://dornsife.usc.edu/abs/rwilcox/software/

Wilsch, A., Henry, M. J., Herrmann, B., Maess, B., & Obleser, J. (2015). Slow-delta phase concentration marks improved temporal expectations based on the passage of time. *Psychophysiology*, *52*(7), 910–918.

Woodrow, H. (1914). The measurement of attention. *Psychological Monographs*, *17*(5), i–158.

Wundt, W. M. (1874). *Grunzüge der physiologischen Psychologie*. Leipzig: W. Engelman.

Yamagishi, N., Callan, D. E., Anderson, S. J., & Kawato, M. (2008). Attentional

changes in pre-stimulus oscillatory activity within early visual cortex are predictive of human visual performance. *Brain Research*, *1197*, 115–122.

Zappoli, R., Versari, A., Zappoli, F., Chiaramonti, R., Thyrion, G. D. Z., Arneodo, M. G., & Zerauschek, V. (2000). The effects on auditory neurocognitive evoked responses and cognitive negative variation activity of frontal cortex lesions or ablations in man: Three new case studies. *International Journal of Psychophysiology*, *38*, 109–144.

Zoefel, B., & Heil, P. (2013). Detection of near-threshold sounds is independent of EEG phase in common frequency bands. *Frontiers in Psychology*, *4*.

Zoefel, B., & VanRullen, R. (2016). EEG oscillations entrain their phase to high-level features of speech sound. *NeuroImage*, *124*, 16–23.

Received September 20, 2017; accepted April 3, 2018.