# How to test for phasic modulation of neural and behavioural responses

**Authors**: Benedikt Zoefel[a,*], Matthew H Davis[a], Giancarlo Valente[b], and Lars Riecke[b]

**Affiliations**: [a] MRC Cognition and Brain Sciences Unit, University of Cambridge, UK

[b] Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, 6229 EV Maastricht, The Netherlands

*Corresponding author: Benedikt Zoefel

MRC Cognition and Brain Sciences Unit

University of Cambridge

15 Chaucer Rd

Cambridge CB2 7EF

United Kingdom

Phone: +44 1223 273691

Email: benedikt.zoefel@mrc-cbu.cam.ac.uk

**Conflict of Interest**: The authors declare no competing financial interests.

Abstract

Research on whether perception or other processes depend on the phase of neural oscillations is rapidly gaining popularity. However, it is unknown which methods are optimally suited to evaluate the hypothesized phase effect. Using a simulation approach, we here test the ability of different methods to detect such an effect on dichotomous (e.g., "hit" vs "miss") and continuous (e.g., scalp potentials) response variables. We manipulated parameters that characterise the phase effect or define the experimental approach to test for this effect. For each parameter combination and response variable, we identified an optimal method. We found that methods regressing single-trial responses on circular (sine and cosine) predictors perform best for all of the simulated parameters, regardless of the nature of the response variable (dichotomous or continuous). In sum, our study lays a foundation for optimized experimental designs and analyses in future studies investigating the role of phase for neural and behavioural responses. We provide MATLAB code for the statistical methods tested.

Keywords
neural oscillations, phase, EEG, MEG, tACS, simulations

## 1. Introduction

Neural oscillations are cyclic variations in the excitability of neuronal ensembles. Oscillatory phase indexes the instantaneous state of excitability (Buzsáki and Draguhn, 2004) and hence correlates with neuronal firing in intracranial recordings (Kayser et al., 2015). The phase of neural oscillations estimated using non-invasive methods such as electro- or magnetoencephalography (EEG/MEG) has been shown to influence human perception and other aspects of cognition (shown schematically in Fig. 1A). For instance, in both the visual and somatosensory systems, the detection of a near-threshold stimulus and the ability to distinguish two rapidly presented stimuli correlate with EEG/MEG phase (Ai and Ro, 2014; Baumgarten et al., 2015; Busch et al., 2009; Mathewson et al., 2009; Milton and Pleydell-Pearce, 2016; Ronconi et al., 2017; but see Ruzzoli et al., 2019). In these studies, stimuli are often presented at an a-priori unknown, random neural phase in each trial; this leads to the possibility of testing how neural or behavioural responses relate to the phase of spontaneous (i.e. stimulus-unrelated) oscillations extracted post-hoc from concurrent EEG/MEG recordings.

An alternative paradigm exploits the observation that neural oscillations align to rhythmic stimulation. In addition to sensory input (Henry and Obleser, 2012; Hickok et al., 2015; Keitel et al., 2017; Mathewson et al., 2012; Spaak et al., 2014), a commonly used stimulus for this purpose is transcranial alternating current stimulation (tACS; Herrmann et al., 2013; Riecke and Zoefel, 2018; Zoefel and Davis, 2017). In such studies, the phase of the electric stimulus serves as a surrogate for neural phase, and the effect of this presumed neural phase on neural or behavioural responses can be assessed. Imposing phase externally in an experimentally controlled manner allows testing for causal effects of phase, but often limits the number of discrete phase values that can be tested.

Irrespective of the exact paradigm used, a question common to all the studies described above is whether and how neural or behavioural responses are influenced by phase. Importantly, there is currently no consensus on how such phase effects can be best revealed statistically (cf. Asamoah et al., 2019). One common statistical approach is to compare responses across trials grouped into phase bins using *parametric tests* (Baumgarten et al., 2015; Busch et al., 2009; Busch and VanRullen, 2010; Chakravarthi and VanRullen, 2012; Gundlach et al., 2016; Madsen et al., 2019; Mathewson et al., 2009; Neuling et al., 2012; Ng et al., 2012; Riecke et al., 2015a, 2015b, 2018; Ruzzoli et al., 2019; Wilsch et al., 2018; Zoefel et al., 2018a; Zoefel and Heil, 2013). However, individuals are often found to differ in their *"preferred" phase* (defined as the phase observed to yield the participant's maximum response; see Fig. 1A and 2A; e.g., Busch et al., 2009; Riecke et al., 2018). Consequently, data is commonly reported from phase bins that are re-aligned relative to the preferred phase for each participant. This

alignment makes it mandatory to exclude the phase bin used for alignment, which therefore reduces potentially important phase-related variance from subsequent analyses of the aligned data.

A second common statistical approach is the use of non-parametric, *permutation-based methods*: Here, the observed data is compared with a surrogate distribution, designed to simulate outcomes due to random fluctuations in the data (i.e. the null distribution). This distribution can be constructed by repeatedly shuffling responses (or phases) associated with single trials in the original dataset and re-running the original analysis on the permuted datasets. Most studies have quantified phase effects by comparing phase distributions of trials that were grouped according to a specific observed behavioural response (e.g., "hit" vs "miss") (Busch et al., 2009; Dugué et al., 2011; Hanslmayr et al., 2013; Harris et al., 2018; Ng et al., 2012; Ronconi et al., 2017; Ruzzoli et al., 2019; Strauß et al., 2015; ten Oever and Sack, 2015; Wutz et al., 2016). Other studies have used regression-based methods (Kayser, 2019; Kayser et al., 2016; McNair et al., 2019) or circular-linear correlations (Busch and VanRullen, 2010; Chakravarthi and VanRullen, 2012) in combination with permutation tests.

In this study, we conducted Monte-Carlo simulations to determine how reliably different statistical methods can detect a phasic modulation of neural or behavioural responses. In principle, these methods can be applied to reveal phase effects in any type of signal, including but not restricted to oscillatory ones. Indeed, some species show phasic modulation of neural activity in the absence of neural oscillations (Eliav et al., 2018). Nevertheless, we here implicitly assumed that the simulated phases stem from an oscillatory (i.e. narrow-band) signal whose phase can be readily interpreted. Methods to detect specifically oscillatory activity have been described elsewhere (e.g., Haller et al., 2018; Watrous et al., 2018; Zoefel et al., 2018b).

We tested how two sets of parameters impact the sensitivity of the tested methods: Parameters that describe the nature of the phase effect and reflect underlying neural processes (Fig. 1), henceforth termed *neural parameters*, and parameters that depend on the specific experimental design, termed *experimental parameters* (see corresponding sections in Material and Methods). We contrasted the two categories of statistical methods described above (*parametric alignment-based* vs *permutation-based*; Fig. 2B,D), and an additional category of parametric regression-based methods requiring no alignment (Fig. 2C). We investigated phase effects on two different types of response variables: dichotomous ("hit" vs "miss") and continuous (e.g., scalp potentials or response times). Together, the results of our simulations serve to guide researchers in choosing the most appropriate method to detect phase effects in future experiments.

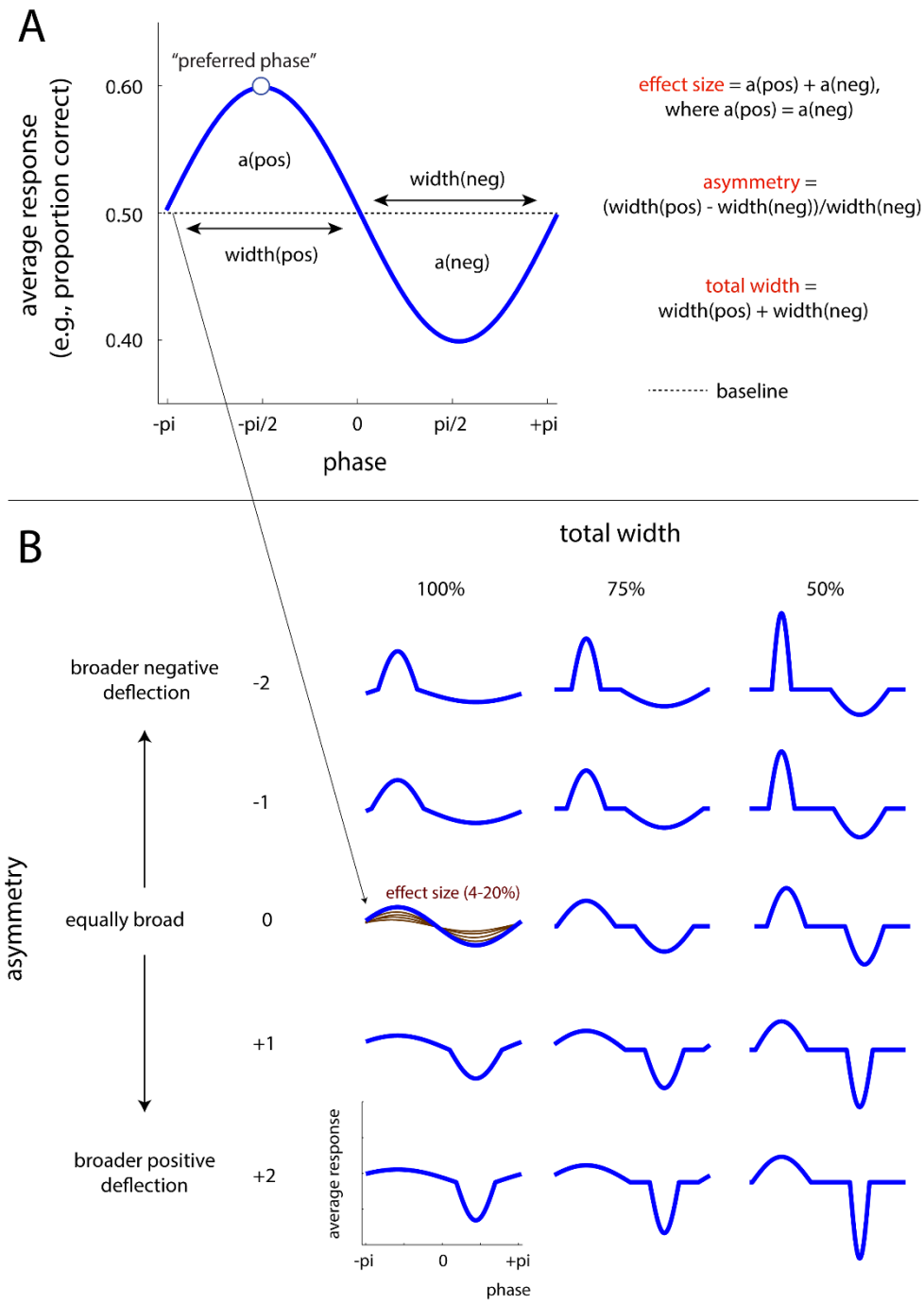



Figure 1. Modelling the phase effect. **A**. Definition of neural parameters. The vertical axis represents the average response (e.g., proportion of hits for a dichotomous response variable, average scalp potential for a continuous response variable) at each simulated phase. This corresponds to *model(phase)* in Section 2.1.1. **B**. Simulated values for neural parameters. These parameters produce different effect shapes that can be sinusoidal (*total width* = 100%, *asymmetry* = 0) or not. Note that *effect size* is identical for all of the blue curves shown. For one exemplary combination of asymmetry and total width, all possible *effect sizes* are shown in brown.

## 2. Material and Methods

Scripts to run the simulations and analyses, as well as the simulated data, are available in the following repository: https://doi.org/10.17863/CAM.41915

### *2.1 Simulating Experimental Data*

We first created a detailed model of the phase effect (Fig. 1), based on various neural parameters (see Section 2.1.2). We then sampled this effect in a simulated experiment, based on various experimental parameters (see Section 2.1.3). For each combination of parameters, we simulated 1000 experiments in which a phasic modulation of the response variable (dichotomous or continuous) is present, and 1000 experiments in which this effect is absent. Each of the simulated experiments consisted of data from 20 virtual participants (Fig. 2A). We then applied various methods (see Section 2.2) to the simulated data and defined, for each of the combination of parameters, a method that is optimally suited (among all of the methods tested) to detect a phase effect at the group level (see Section 2.3).

#### *2.1.1. Simulating neural or behavioural responses*

In each trial of a simulated experiment, a virtual response was recorded at a (e.g., neural or stimulus-specific) phase. This phase was selected in two different ways, depending on *study design* described in Section 2.1.3. A response was then determined for each trial, as follows.

For a dichotomous response variable:

If *rand*(1) >= 1-*model(phase)*, then response = “hit”, otherwise “miss”.

For a continuous response variable:

Response = *normrnd*($\mu$ = *model(phase)*, $\sigma$ = 0.7).

The terms *rand* and *normrnd* correspond to a random value drawn from a standard uniform distribution and normal distribution, respectively, and *model(phase)* corresponds to the average response (e.g., proportion of hits for a dichotomous response variable) associated with a certain phase in a detailed model of the phase effect (Fig. 1). This model was determined by the simulated neural parameters as described in the next section.

#### *2.1.2 Neural Parameters*

We assigned values to the response variable at various phases (Fig. 1), which spanned a full oscillatory cycle in equidistant steps of $\pi/96$. This resulted in 192 values over a full cycle. How exactly these values varied with phase depended on following parameters:

(1) *Baseline:* For all experiments, the average response (e.g., proportion of hits or scalp potential) across phase was chosen to be 0.5. Using other baselines (e.g., a value of 0) did not change the outcome of these simulations.

(2) *Preferred phase:* For experiments with a phase effect present, the maximum response (most positive deflection from baseline) was assigned to a randomly chosen phase (again, with a resolution of $\pi/96$, i.e. there were 192 possibilities). The minimum response (most negative deflection from baseline) was assigned to the phase bin $\pi$ away from this “preferred” phase. Note that the preferred phase was the only parameter that varied between individual subjects within each experiment.

(3) *Effect size:* Effect size was defined as the area between baseline and the two (positive and negative) deflections. We chose area (rather than, for instance, the distance between positive and negative peaks) as our measure of effect size as it takes into account a larger number of phases (rather than only the most and least preferred ones). We assumed that phasic increases of a response variable (i.e. phase values yielding higher than average responses) have a matched counterpart of phase values leading to

decreased responses (i.e. lower than average), as explained by VanRullen and McLelland (2013). Consequently, phase effects were constructed in such a way that the area under the positive deflection was always identical to that under the negative deflection. We tested effect sizes that correspond to a peak-to-peak modulation of between 4% and 20% (in steps of 4%) for a sinusoidal effect.

(4) *Total width*: The width of a deflection ($width_{pos}$ for the positive deflection, and $width_{neg}$ for the negative deflection) was defined as the number of phase values covered by the deflection relative to the total number of phase values (192). In practice, $width_{pos}$ and $width_{neg}$ were altered by compressing the positive and negative half cycles of a sine wave, respectively, so that they covered the desired amount of phase values (Fig. 1B). The total width of the phase effect was defined as $width_{pos}$+ $width_{neg}$ and it could take the following values in our study: 50%, 75%, and 100%. Note that, assuming a constant effect size (area under the deflection, see *effect size* above), a reduction in total width results in an increase in peak amplitudes, as visible in Fig. 1B.

(5) *Asymmetry:* Asymmetry was defined as

$$\frac{width_{pos} - width_{neg}}{width_{neg}}$$

See Cole and Voytek (2019) and Schaworonkow and Nikulin (2019) for similar definitions. Note that negative or positive values of this parameter reflect a broader negative or positive deflection respectively. We simulated asymmetries between -2 and 2, in steps of 1. Again, note that, assuming a constant effect size, a change in asymmetry results in change in peak amplitudes, as visible in Fig. 1B.

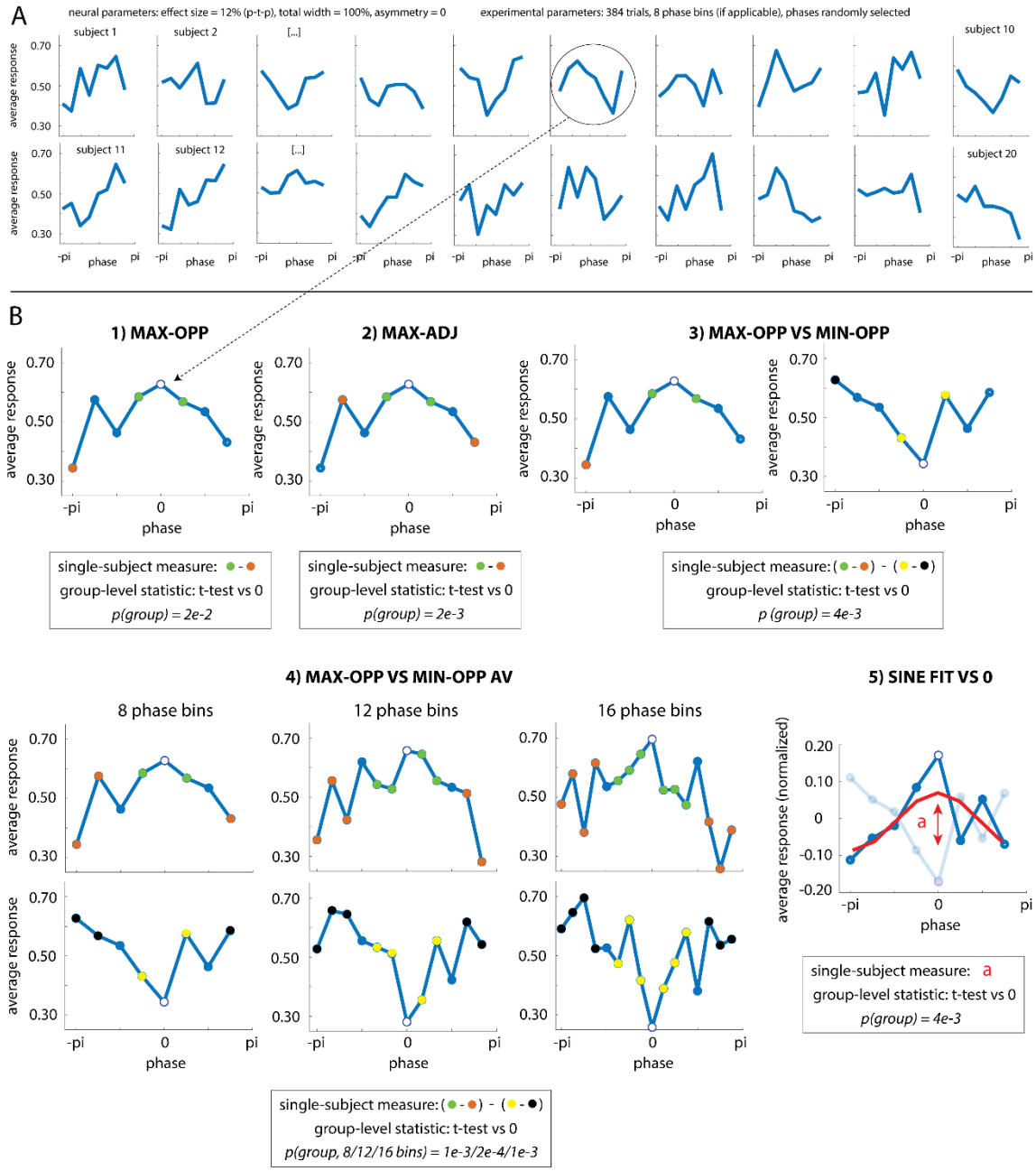
### *2.1.3 Experimental Parameters*

In the context of a typical empirical study, neural parameters are not under the researcher's control because they concern the neural mechanisms by which phase effects are reflected in neural or behavioural responses. However, researchers have control over the experimental protocol by which phase effects are measured. We therefore systematically varied the following experimental parameters:

(1) $N_{trials}$*:* The number of trials in total was either 192, 384, 768, or 1152. Note that the simulated data (Fig. 2A) increasingly approximates the "true" effect (Fig. 1) with increasing number of trials. 1152 trials were not tested for the continuous response variable.

(2) $N_{bins}$*:* For some analyses, data is sorted into phase bins in order to increase the number of trials per bin and, consequently, statistical power. Similarly, a limited number of phase bins are typically tested in studies where the phase is imposed externally (e.g., tACS) because of practical limits on the number of trials that can be tested within an experimental session. The number of phase bins simulated in our study was 4, 6, 8, 12, or 16. 16 phase bins were not tested for the continuous response variable.

(3) *Study design:* We tested two common paradigms for how phases were selected for each virtual trial (see Introduction). In the first paradigm (*randomly selected*), phases were randomly selected in each trial, with 192 possible phase values as described above. In the second paradigm (*imposed externally*), only a small number of equidistant phase values was tested. In this case, the tested phase values corresponded to the centres of the phase bins used for the analysis (see $N_{bins}$).


A
neural parameters: effect size = 12% (p-t-p), total width = 100%, asymmetry = 0
experimental parameters: 384 trials, 8 phase bins (if applicable), phases randomly selected
subject 1
subject 2
[...]
subject 10
subject 11
subject 12
subject 20
average response
phase
B
1) MAX-OPP
2) MAX-ADJ
3) MAX-OPP VS MIN-OPP
single-subject measure: ● - ●
group-level statistic: t-test vs 0
p(group) = 2e-2
p(group) = 2e-3
single-subject measure: (● - ●) - (● - ●)
p (group) = 4e-3
4) MAX-OPP VS MIN-OPP AV
8 phase bins
12 phase bins
16 phase bins
5) SINE FIT VS 0
average response (normalized)
a
p(group, 8/12/16 bins) = 1e-3/2e-4/1e-3
p(group) = 4e-3

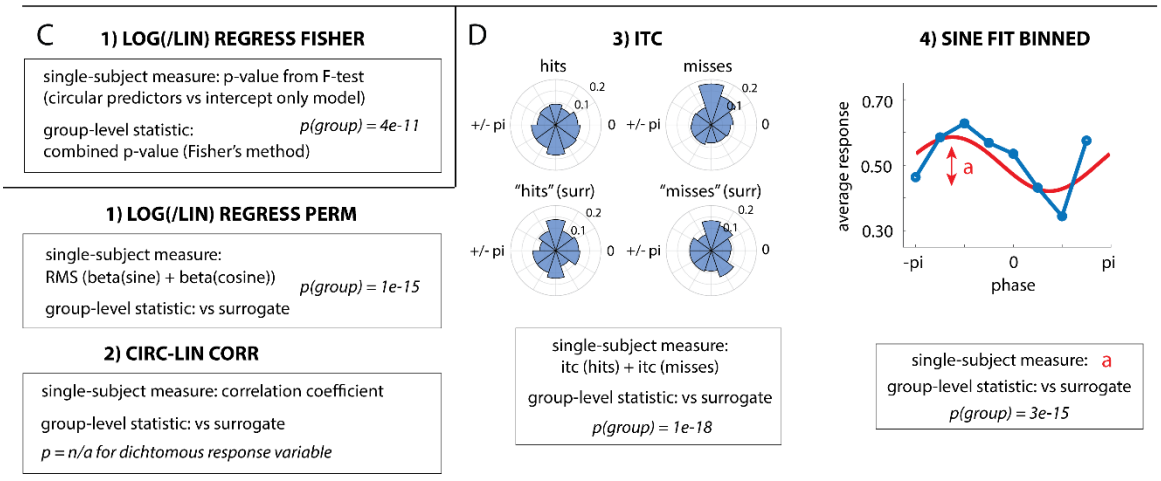

C
1) LOG(/LIN) REGRESS FISHER
single-subject measure: p-value from F-test
(circular predictors vs intercept only model)
group-level statistic:
combined p-value (Fisher's method)
p(group) = 4e-11
1) LOG(/LIN) REGRESS PERM
single-subject measure:
RMS (beta(sine) + beta(cosine))
group-level statistic: vs surrogate
p(group) = 1e-15
2) CIRC-LIN CORR
single-subject measure: correlation coefficient
group-level statistic: vs surrogate
p = n/a for dichtomous response variable
D
3) ITC
hits
misses
"hits" (surr)
"misses" (surr)
+/- pi
single-subject measure:
itc (hits) + itc (misses)
group-level statistic: vs surrogate
p(group) = 1e-18
4) SINE FIT BINNED
average response
phase
a
single-subject measure: a
group-level statistic: vs surrogate
p(group) = 3e-15

Figure 2. Illustration of statistical methods. **A**. Data from 20 virtual participants in one experiment for a given combination of (neural and experimental) parameters and a dichotomous response variable. Note that the preferred phase (the phase yielding the maximum response) differs across participants. **B-D**. Statistical methods that were included in the present study, divided into three categories. For all methods, the single-subject measure of the hypothesized phase effect is visualized and/or described based on data from one exemplary subject. If methods divide data into phase bins, these are shown with circles. For alignment-based methods (B), the bin used for alignment is shown as an open circle. For all methods, p-values shown were obtained for the group level, by applying the respective method to the data shown in A. The panel illustrating method ITC (D, 3) depicts phase distributions for hit and miss trials (observed data and surrogate distribution). See Section 2.2 for further details.

We note that some of the methods we tested do not divide data into phase bins (cf. Fig. 2 and Section 2.2). For these methods, we did not vary $N_{bins}$ if phases were *randomly selected.* In the case of phases *imposed externally*, for these methods, $N_{bins}$ only corresponds to the number of possible phases in each trial.

### *2.2 Statistical Methods*

Phasic modulations of neural or behavioural responses are traditionally assessed with statistical methods that can be divided into two broad analytical categories: Parametric alignment-based methods and permutation-based methods. We here added a third category, consisting of parametric methods that are based on regression analysis and therefore avoid phase binning and re-alignment. Note that some of the permutation-based methods we investigated are also based on regression.

We conceived and evaluated 24 methods (10 parametric alignment-based, 2 parametric regression-based, and 12 permutation-based methods). Each of these is described in detail below or in the Supplementary Material. In order to keep the paper as concise and clear as possible, here we only report results for methods (1) if they have been used in previous studies or (2) if we identified them as the most sensitive ("winning") method for at least one combination of parameters in the respective method category. Novel methods were only selected if they performed significantly better than an established method for at least one combination of parameters. These criteria led to the selection of 5 parametric alignment-based, 1 parametric regression-based, and 4 permutation-based methods, which are shown in the Results section of this article and are therefore described in detail in the following. These methods are also illustrated in Fig. 2.

#### *2.2.1 Parametric Alignment-based Methods (Fig. 2B)*

All methods in this category divide data into phase bins. They only differ in their single-subject measure of the phase effect, which is calculated from the binned data and described below for each method. At the group level, all methods compare these measures against 0, here using a one-tailed, one-sample t-test. We note that several studies have compared responses across phase bins using ANOVAs (Baumgarten et al., 2015; Busch et al., 2009; Busch and VanRullen, 2010; Ng et al., 2012; Zoefel and Heil, 2013). This approach may be non-optimal since it does not take into account the hypothesized cyclical shape of the phase effect (i.e., it predicts a difference between phase bins, but, in contrast to the methods described below, not the identity of the bins that are maximally different). We therefore decided not to describe it in detail here; however, we confirmed in simulations that its sensitivity is indeed lower than that of the majority of the methods tested (data not shown).

1. MAX-OPP: The maximal response was aligned to the centre bin and the remaining bins phase-wrapped. The response of the bin opposite to the centre bin was subtracted from the average response of the two bins adjacent to the centre bin. This method is a slight modification from one that was applied by, e.g., Riecke et al. (2018).

2. MAX-ADJ: We used the same method as described in (1), but subtracted the average response of the two bins adjacent to the bin opposite to the centre bin from the average response of the two bins adjacent to the centre bin. This method was used in, e.g., Riecke et al. (2018). Note that this method cannot be used for designs with only 4 phase bins.

3. MAX-OPP VS MIN-OPP: We used the same method as described in (1) to compute difference d1. Then these steps were repeated, but after aligning individual data based on the *minimum* (instead of maximum) response, to compute difference d2. If a phase effect were present, we would expect a positive value for d1 and a negative value for d2. We therefore subtracted d2 from d1.

4. MAX-OPP-AV VS MIN-OPP-AV: We included a variation of the MAX-OPP VS MIN-OPP method for designs in which there are 8, 12, or 16 phase bins. In this case, N bins adjacent to the centre bin (e.g., if N = 2, one bin to either side of the centre bin) were used to determine d1, and M bins adjacent to and including the opposite bin (e.g., if M = 3, the one bin to either side of the opposite bin plus the opposite bin) were used to calculate d2, where

N = ($N_{bins}$-2)/2-1

and

M = ($N_{bins}$-2)/2

For instance, in the case of 16 phase bins, 6 phase bins adjacent to the centre bin (3 to either side) were used to calculate d1, and 6 phase bins adjacent to the bin opposite to the centre bin (3 to either side), plus this opposite bin, were used to calculate d2. For 8 phase bins, the method is very similar to its original MAX-OPP VS MIN-OPP method, with M being the only difference between the two (N=2 for both versions, but M=1 for the original version, and M=3 for the variant).

5. SINE FIT VS 0: The average response across phases was first subtracted from all phase bins (resulting in an average of 0). The response that deviated most strongly from 0 (i.e. maximum or minimum) was then aligned to the centre bin and remaining bins phase-wrapped. The sign was flipped (i.e. multiplied by -1) if the bin used for alignment corresponded to a minimum. A sine wave was then fitted to the data, excluding the centre bin and restricting the phase of the fitted sine wave so that its peak corresponded to the position of the omitted centre bin. The amplitude of the fitted sine wave was extracted. Note that this amplitude can assume negative values, corresponding to a flipped sine wave with the trough at the position of the centre bin. A similar method was used in Zoefel et al. (2018a).

### *2.2.2 Parametric Regression-based Methods (Fig. 2C)*

All the parametric regression-based methods that we tested are based on the same single-subject measure of the phase effect: Logistic (for dichotomous response variables) or linear (for continuous response variables) regression was used to test whether phase predicts responses at the single-trial level. Sine- and cosine-transformed phases were included in the regression model, yielding two circular predictors of the participant's response. The full regression model (including phase) was then compared with an intercept-only model, using an F-Test, and yielding a p-value for each participant. The methods investigated differed in how this p-value was combined across participants. One method performed best for all parameter combinations and is therefore the only one reported here. This method, LOG/LIN REGRESS FISHER, combines the obtained p-values according to Fisher's method (Fisher, 1950):

$$T = -2 * \sum_{n=1}^{N} \ln(p_n)$$

where $p_n$ corresponds to the p-value for participant *n*, and T follows, under the null hypothesis, a Chi-square distribution with 2*N degrees of freedom.

A similar method has been used in Tomassini et al. (2017) and Liu and Luo (2019). However, these studies compared the two logistic regression coefficients (sine and cosine), averaged across participants, with 0 (using Hotelling's T-squared statistic), which assumes a consistent preferred phase across participants. As this assumption does not seem to hold in all experimental scenarios (e.g., Busch et al., 2009; Riecke et al., 2018; Zoefel et al., 2018a), we randomly varied preferred phase across individuals and therefore did not include this variant of the method in our simulations.

### *2.2.3 Permutation-based Methods (Fig. 2D)*

All the permutation-based methods that we tested only differ in their single-subject measure of the phase effect, described separately for each method below. At the group level, all methods rely on a comparison of original data and a surrogate distribution. In the following, we use the term "permutation-based methods"; other terms, such as "bootstrap tests" or "resampling statistics", are conceptually equivalent approaches (for an extensive introduction, see Good, 2005).

The common notion underlying permutation-based methods is that a hypothetical effect in the original data (e.g., a correlation between phase and an outcome measure) is abolished by randomly reassigning trial outcomes in single participants. The analysis method that was previously applied to the original dataset can then be applied to the randomized dataset which is known to lack the hypothesized effect (due to the randomization). Typically, the original data is permuted N times, thus yielding N outcomes for a given method, applied to different permutations of the same original data. These N outcomes are combined to form a distribution that is known to be produced by random fluctuations in the data, i.e. under the null hypothesis. If the original outcome is relatively unlikely to stem from the surrogate distribution, the alternative hypothesis is accepted. An advantage of this approach is that, by repeatedly "sampling" from the original data, the surrogate distribution can be used to estimate certain properties of the variable of interest (e.g., its confidence interval, cf. Schaworonkow et al., 2019). Assumptions regarding the data distribution (e.g., normality in the case of parametric methods) can therefore be avoided.

In our simulated data, the surrogate distribution was constructed by assigning single-trial responses to phase bins at random, thereby abolishing any effect of phase, and then applying the statistical methods to the permuted data. This procedure was repeated 100 times, resulting in a surrogate distribution consisting of 100 outcomes that was then compared with the single outcome obtained from the original, non-permuted data. This comparison resulted in statistical (z-) values, according to:

$$z = (d-\mu) / \sigma$$

where z is the z-transformed effect size for the observed data, d is the observed data, and $\mu$ and $\sigma$ are mean and standard deviation of the surrogate distribution, respectively (see VanRullen, 2016). The effect observed was considered reliable if the z-value exceeded a critical value ($z = 1.645$, corresponding to a significance threshold of $\alpha = 0.05$, one-tailed).

For all methods tested, the single-subject measures described below were averaged across participants before the comparison with the surrogate distribution (averaged likewise in each permutation).

1. LOG/LIN REGRESS PERM: Logistic (for the dichotomous response variable) or linear (for the continuous response variable) regression with circular predictors was run as described for parametric regression-based methods. The root-mean square of the two regression coefficients (sine and cosine) was calculated. A similar method was used in Kayser et al. (2016) and McNair et al. (2019).

Note that linear regression with circular predictors is mathematically equivalent to determining the amplitude of a sinusoidal wave at a given frequency with a discrete Fourier transform. Consequently, LIN REGRESS PERM is related to SINE FIT BINNED (described as method 3 below), with the only difference that the former does not require to divide data into phase bins.

2. CIRC-LIN CORR: This method was only used for the continuous response variable. Circular-linear correlation between phase and single-phase responses was calculated. This method has been used previously in Busch and VanRullen (2010) and Chakravarthi and VanRullen (2012). It is important to note that the obtained correlation coefficient is numerically equivalent to the square root of the goodness-of-fit ($R^2$) from linear regression with circular predictors.

3. ITC: For the dichotomous response variable, phases were extracted separately from hit and miss trials. Phase consistency across trials (inter-trial coherence, ITC) was then calculated for each trial category, and the two ITC values summed:

$$ITC_{total} = \left| \frac{1}{N_{hits}} \sum_{n=1}^{N_{hits}} e^{i(\varphi_{hits}(n))} \right| + \left| \frac{1}{N_{misses}} \sum_{n=1}^{N_{misses}} e^{i(\varphi_{misses}(n))} \right|$$

where $\varphi_{hits}(n)$ and $\varphi_{misses}(n)$ is the phase observed in trial n of all hit trials ($N_{hits}$) and miss trials ($N_{misses}$), respectively. For the continuous response variable, single trials were split based on the median response across trials, and $\varphi_{hits}$ and $\varphi_{misses}$ replaced with $\varphi_{outcome\geq=median}$ and $\varphi_{outcome\leq median}$, respectively.

For dichotomous responses, this method corresponds to the Phase Opposition Sum (POS) described in detail in VanRullen (2016) and used in several studies by the same group and others (cited in the original paper). The reasoning is that under the hypothesis that the average phase for hits and misses differs, the phase consistency across trials should be higher if determined for hits and misses separately than for a mix of the two. Other variants of this method exist, all comparing phase distributions for hit and miss trials; as one of the most sensitive variants (VanRullen, 2016), POS is used here as a representative approach in this category.

4. SINE FIT BINNED: A sine wave was fitted to data divided into phase bins and averaged across trials, with phase and amplitude as free (i.e. unconstrained) parameters of this fit. The frequency of the fitted sine wave corresponded to the oscillatory frequency of interest (i.e. a single cycle in a typical experiment). Sine wave fits were achieved by running a discrete Fourier Transform and extracting the amplitude of the first non-DC output component. This method was used in Zoefel and VanRullen (2015) and Zoefel et al. (2018).

*2.3 Evaluation of Methods*

For an experiment with an effect present, if a given method yielded a significant effect at the group level (p-value < α = 0.05), this was counted as a true positive. For an experiment with effect absent, if the method yielded a significant effect, this was counted as a false positive.

The probability of a true positive ($p_{true}$) was computed as the number of true positives observed, divided by the number of experiments with effect present (1000). Analogously, the probability of a false positive ($p_{false}$) was computed as the number of false positives observed, divided by the number of experiments with effect absent (1000). The sensitivity of a method to detect an effect at the group level was quantified using d-prime, the standardized difference between the probabilities of true and false positives:

d' = z($p_{true}$) – z($p_{false}$).

For each combination of parameters, the method yielding the highest d-prime was identified and defined as the "winning" method.

95% Confidence intervals for $p_{true}$ and $p_{false}$were calculated as follows:

$$p \pm 1.96 * SE_p$$

$$SE_p = \sqrt{\frac{p * (1-p)}{n}}$$

where p corresponds to $p_{true}$ or $p_{false}$, respectively, and n corresponds to the number of observations (i.e. experiments).

95% Confidence intervals for d-prime were calculated based on the Gourevitch and Galanter formula as follows (from Jesteadt, 2005):

$$d' \pm 1.96 * SE_{d'}$$

$$SE_{d'} = \sqrt{\frac{p_{true} * (1-p_{true})}{n * \ (ord\ z_{1-p_{true}})^2} + \frac{p_{false} * \left(1-p_{false}\right)}{n * \ (ord\ z_{p_{false}})^2})}$$

where

$$ord\ z_x = \frac{1}{\sqrt{2 * \pi} * e^{\frac{z(x)^2}{2}}}$$

In order to determine whether two methods differed significantly in their performance, we tested whether the confidence interval of their difference included 0. A difference was defined as significant if the following condition was met:

$$x_1 - x_2 > 1.96 * \sqrt{{SE_{x1}}^2 + {SE_{x2}}^2}$$

where $x_1$ and $x_2$ represent the measure of interest ($p_{true}$, $p_{false}$, or d') from two different methods and $x_1 > x_2$.

*2.4 Split-data Approach for Alignment-based Methods*

In all of the parametric alignment-based methods tested, the phase bin used for alignment cannot be included in subsequent analyses as it is trivially an extremum. We tested whether the sensitivity of these methods can be improved if this data loss is avoided. We repeated our simulations (only for dichotomous response variable), but used a certain percentage (25%, 50%, or 75%) of the data to estimate the preferred (or "non-preferred", for the alignment based on minimum responses) phase of each participant. This phase was then used for alignment in the remaining part of the data. Only the latter data including all phase bins was used to test for phase effects. Critically, as the phase for alignment was determined in an independent dataset, no phase bin needed to be excluded in this case. On the other hand, this procedure can reduce the number of trials used to test for phase effects (depending on the percentage of the data used to estimate preferred phase).

To apply this split-data approach, some of the methods described in Section 2.2.1 needed to be slightly modified for as described in the Supplementary Materials.

## 3. Results

We observed a very similar pattern of results for the two types of response variables. We therefore focus on results obtained for dichotomous responses, followed by a brief overview of results for continuous ones.

*3.1 Effects of Experimental Parameters*

We first pooled across all neural parameters and revealed the method with the highest sensitivity (i.e. d-prime) for each combination of experimental parameters. In Fig. 3, these highest possible sensitivities

are shown for each method category. Confidence intervals for the sensitivities shown range between d'±0.12 and d'±0.16.

We observed that parametric regression-based methods and permutation-based methods performed very similar (average d' 3.79 vs 3.74) and generally outperformed parametric alignment-based methods (average d' 2.89). Study design (phases *randomly selected* vs *imposed externally*) did not strongly affect the ability to detect phase effects (average d' 3.46 vs 3.49). For all methods, sensitivity increased monotonically with $N_{trials}$. Although less critical than $N_{trials}$, sensitivity of parametric alignment-based methods was modulated by $N_{bins}$, with lowest sensitivity for 4 bins, and highest sensitivity for 12 or 16 bins (depending on $N_{trials}$ and *study design*). Most methods in the other categories do not divide data into phase bins. These methods were not affected by the number of possible phases in each trial, reflected by $N_{bins}$ for phases *imposed externally* (see Materials and Methods). Together, these findings show that the highest possible sensitivity to detect phase effects can be achieved by using (1) regression- or permutation-based methods and (2) a high $N_{trials}$.

Supplementary Figs. 1 and 2 show the probabilities of false and true positives, respectively, which together determine the sensitivity values shown in Fig. 3. False positive probability fluctuated around 5%, reflecting the significance threshold α applied in each simulated experiment. Consequently, none of the methods was biased towards false positives (cf. Asamoah et al., 2019), and differences in sensitivity observed were mostly driven by changes in true positive probability.

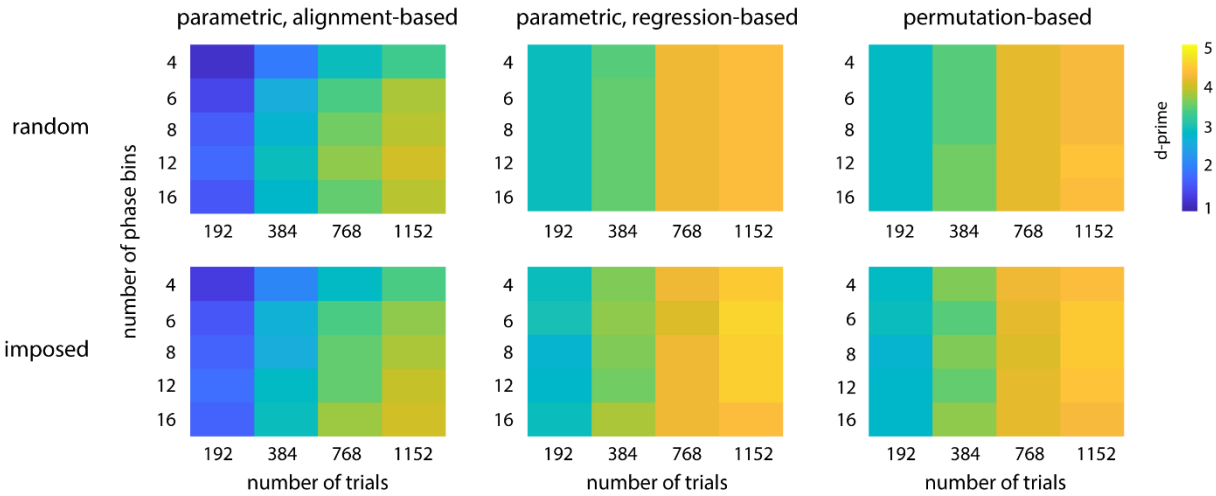



Figure 3. Highest possible sensitivity to detect phase effects (i.e. highest d-prime among all methods tested; color-coded) separately for all experimental parameters tested and for the three method categories. Values were averaged across all neural parameters before maximal sensitivity was determined. The confidence intervals of all sensitivities shown are in the range d'±0.12 to d'±0.16 (not shown).

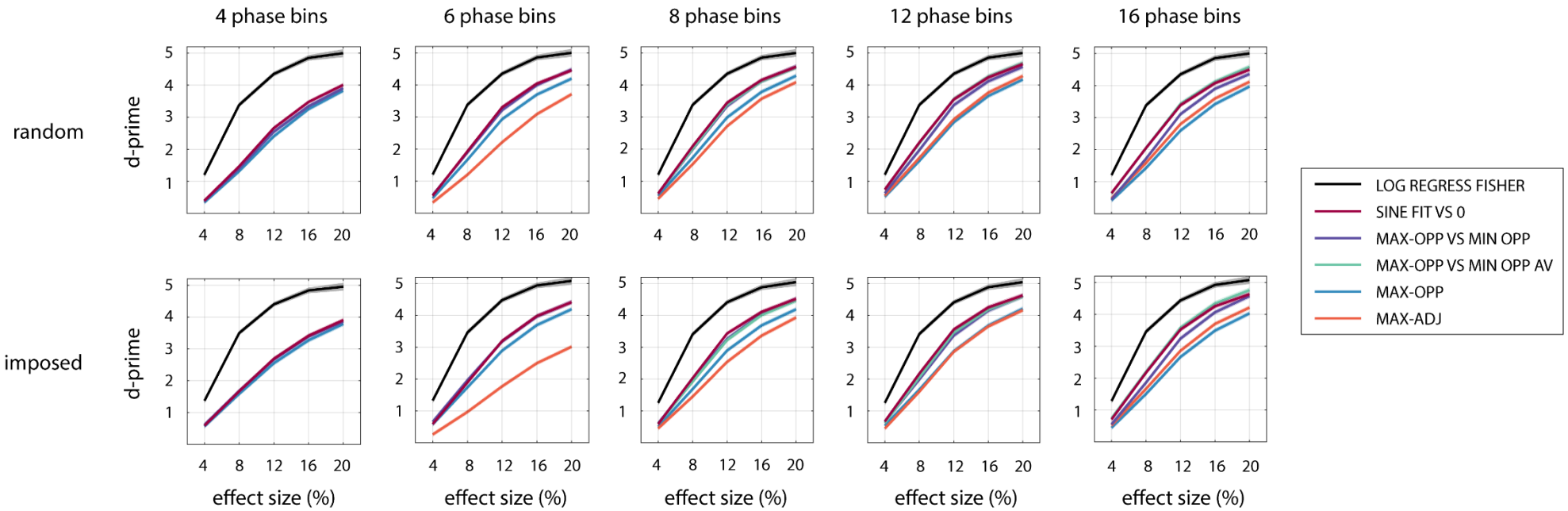



Figure 4. Sensitivity of the selected parametric alignment-based and parametric regression-based methods to detect phase effects, for various combinations of $N_{bins}$, *study design*, and *effect size*. Results were averaged across $N_{trials}$ as this parameter did not affect the identity of the winning methods. *Effect size* corresponds to a peak-to-peak modulation of performance for a sinusoidal shape (*total width* = 100%, *asymmetry* = 0). Note that MAX-ADJ is not defined for 4 phase bins, and MAX-OPP VS MIN OPP AV is not defined for 4 and 6 phase bins. Confidence intervals are shown by shaded areas – note that these are often too narrow to be visible.

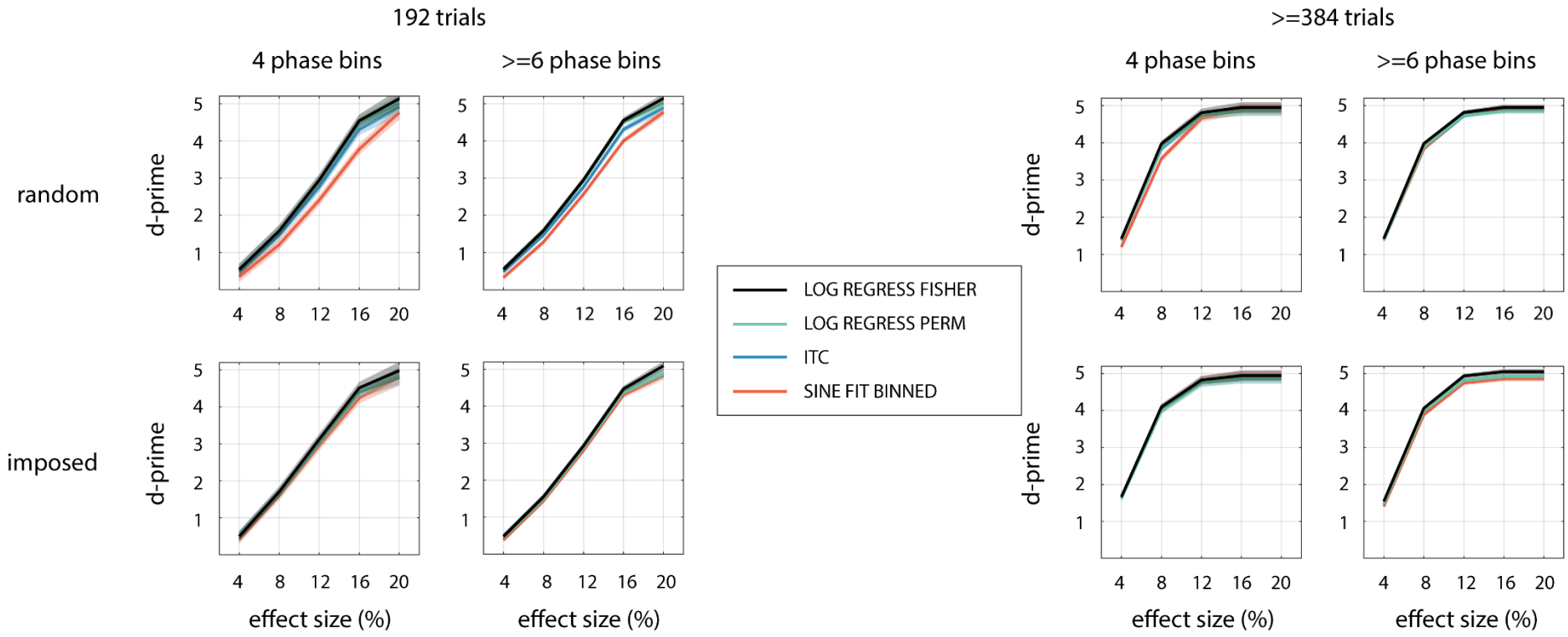



Figure 5. Sensitivity of the selected permutation-based and parametric regression-based methods to detect phase effects. Other conventions as in Fig. 4.

In Fig. 3, for each combination of parameters, we only show the sensitivity of a single "winning" method (that with the numerically highest value). However, in most cases, several methods performed similarly well. In Figs. 4 and 5, we therefore show sensitivity separately for different methods, together with their confidence interval. Results are shown for those experimental parameters which mostly strongly affected the identity of the best method in one or more method categories. Although not an experimental parameter, we also show results as a function of *effect size*, in order to demonstrate its positive effect on sensitivity.

Fig. 4 shows that SINE FIT VS 0 is a parametric alignment-based method with a relatively high sensitivity in all scenarios. MAX-OPP VS MIN-OPP (for lower $N_{bins}$) or MAX-OPP-AV VS MIN-OPP-AV (for higher $N_{bins}$) often perform equally well. The selected parametric regression-based method (LOG REGRESS FISHER; black line) is more sensitive than all alignment-based methods and for all tested combinations of parameters.

We found similar sensitivity for LOG REGRESS FISHER and the three selected permutation-based methods (Fig. 5). Some exceptions can be seen for *randomly selected* phases, where SINE FIT BINNED is less sensitive than other methods for low $N_{trials}$ and low $N_{bins}$, and ITC shows a slightly reduced performance for low $N_{trials}$. The highest sensitivity was consistently found for LOG REGRESS FISHER and LOG REGRESS PERM, which performed equally well.

Supplementary Figs. 3 and 4 show corresponding probabilities of false and true positives, respectively.

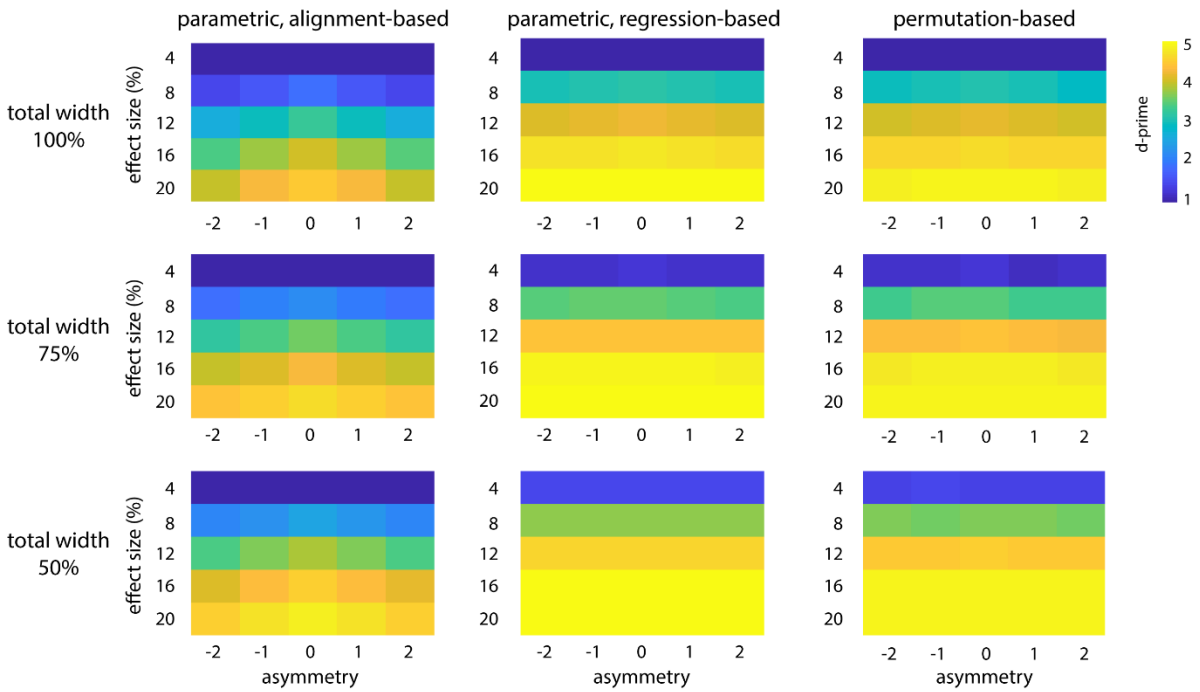



Figure 6. Highest possible sensitivity to detect phase effects, separately for all neural parameters and the three method categories tested. The confidence intervals of all sensitivities shown are in the range d'±0.02 (for low *effect sizes*) to d'±0.12 (for high *effect sizes*).

*3.2 Effects of Neural Parameters*

Next, we investigated the effect of neural parameters and their interaction with experimental parameters. In Fig. 6, we show highest possible sensitivities for different combinations of neural parameters. Confidence intervals range from d'±0.02 (for low *effect sizes*) to d'±0.12 (for high *effect sizes*).

As already illustrated above, sensitivity to detect phase effects increased strongly with *effect size*. Sensitivity was highest for *symmetric* shapes, and this effect was most pronounced for parametric alignment-based methods (average d' of 3.08 vs 2.73 for symmetric vs most asymmetric shape). Other method categories were affected only little by changes in asymmetry (average d' of 3.81 vs 3.77 for parametric regression-based and 3.74 vs 3.68 for permutation-based methods, respectively). Finally, in all categories, sensitivity was modulated by *total width*, with highest sensitivity for narrow effects (increase in d' of ~0.5 for 50% vs 100% *total width*). However, note that *effect size* was kept constant for different widths, resulting in higher peak amplitudes for narrower effects (due to *effect size* being defined based on area rather than peak amplitude; see Materials and Methods). Our results therefore suggest that the methods tested here are more effective in detecting stronger, but transient, phasic modulations than weaker modulations covering the whole oscillatory cycle.

Fig. 7 again illustrates how parametric alignment-based methods perform relative to each other. Results are shown for combinations of $N_{bins}$ and *total width*, as the two parameters that most strongly affected the identity of the optimal method in this category. For comparison, the average sensitivity of LOG REGRESS FISHER and LOG REGRESS PERM is shown, as the two best-performing methods in the other two categories. These did not differ in their sensitivity for any of the parameter combinations shown.

Again, we observed that SINE FIT VS 0 is a parametric alignment-based method that is relatively sensitive in all scenarios. For narrow effects (50% *total width*), MAX-OPP VS MIN-OPP performed better for 6 and 8 phase bins. For broader effects, the sensitivity of MAX-OPP-AV VS MIN-OPP-AV increased with $N_{bins}$, eventually becoming the winning alignment-based method for 12 or more phase bins. Regression- and permutation-based methods again outperformed alignment-based methods for all combinations of parameters.

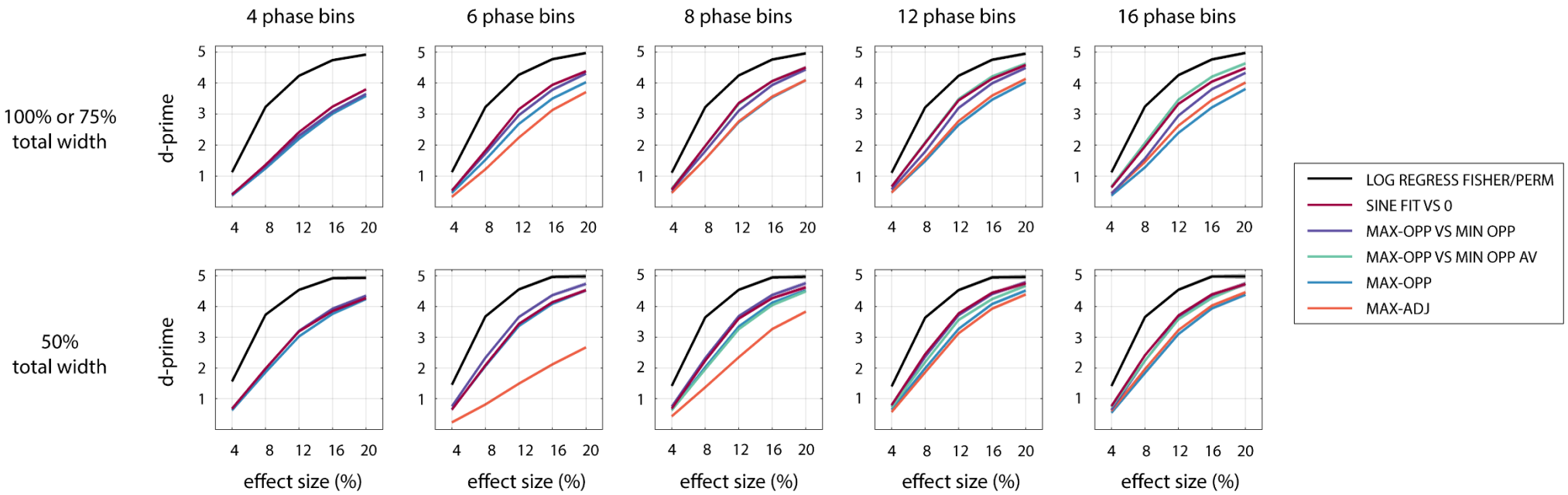



Figure 7. Sensitivity of the selected parametric alignment-based methods to detect phase effects, for combinations of $N_{bins}$, *total width*, and *effect size*. The black line shows the average sensitivity for LOG REGRESS FISHER and LOG REGRESS PERM which performed equally well. For other conventions, see caption of Fig. 4.

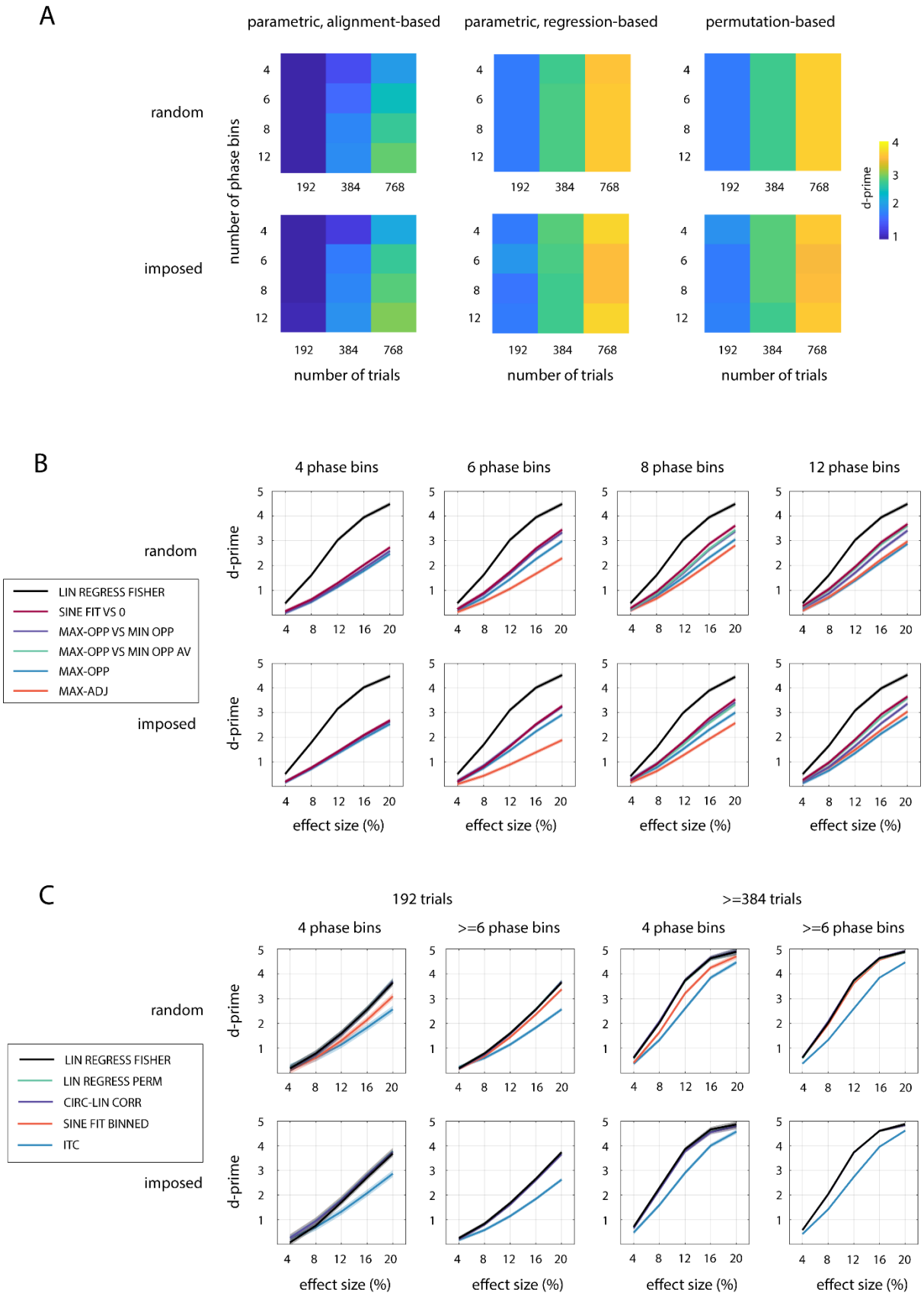



Figure 8. **A**. Highest possible sensitivity to detect the phasic modulation of a continuous response variable, separately for different experimental parameters. Note that overall sensitivity (across all parameters) is not comparable with that for the dichotomous response variable (Fig. 3), as it depends on the parameters that characterize the distribution from which single-trial responses are chosen (see Materials and Methods). **B,C.** Sensitivities separately for the selected parametric alignment-

based (B), parametric regression-based (B,C) and permutation-based methods (C) for various simulated parameters (cf. Figs. 4,5). Note that LIN REGRESS FISHER, LIN REGRESS PERM, and CIRC-LIN CORR overlap for all parameter combinations shown in panel C. For all other conventions, see caption of Fig. 4.

### *3.3. Continuous Response Variable: Overview of Results*

We found that the results described above can generalized to a phasic modulation of a continuous response variable. In short, (1) the best parametric regression-based (LIN REGRESS FISHER) and best permutation-based methods performed similarly and outperformed parametric alignment-based ones. (2) LIN REGRESS PERM and CIRC-LIN CORR were the permutation-based methods with the highest sensitivity for all combinations of parameters. Sensitivity of the ITC method decreased strongly, presumably because it relies on a dichotomous trial classification ("hit" vs "miss"; see Materials and Methods). (3) SINE FIT VS 0 was the most sensitive parametric alignment-based method for most parameter combinations. (4) $N_{trials}$, *effect size*, and *effect width* (for all method categories) as well as $N_{bins}$ and *asymmetry* (for alignment-based methods only) had a general modulatory effect on sensitivity. Fig. 8 illustrates these findings.

### *3.4 Effectiveness of Split-data Methods for Phase Alignment*

We also tested whether the sensitivity of parametric alignment-based methods can be improved by estimating individual preferred phases from a subset of data that is used exclusively for this purpose; these estimates were then used to align and test for phase effects in the remaining part of the data without the need to exclude the aligned phase bin (see Materials and Methods). Fig. 9 compares the ability of these split-data methods to detect phase effects ("split-data") with the most common approaches described in the previous sections. We found that the split-data approach did not improve the sensitivity of alignment-based methods. That is, the availability of all trials for all but one phase bin outweighs the possibility to analyse performance in all phase bins based on a possibly smaller overall $N_{trials}$.

In some studies the preferred phase of individual participants may be known beforehand (e.g., from other experiments) so that phase effects can be tested based on a complete recorded dataset (without sacrificing parts of it to estimate this preferred phase). Indeed, when we assumed that this phase was known, enabling phase alignment without the necessity of excluding any data, sensitivity was improved as compared to our standard alignment-based approach. However, sensitivity was still lower than that observed for other method categories.

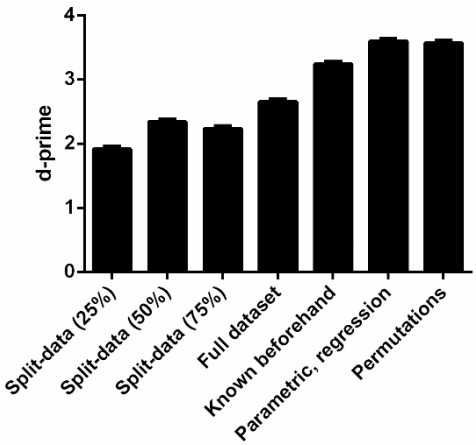



Figure 9. Highest possible sensitivity (i.e. highest d-prime among all methods) for split-data methods and other approaches. Results are shown for simulated experiments of 384 trials each (using other $N_{trials}$ did not change results) and the dichotomous response variable. Sensitivity was averaged across all other parameter combinations before maximal sensitivity was determined. "Split-data": The dataset was split, with one part used to identify the preferred phase (the percentage of data used for this purpose is indicated), the other to test for actual phase effects across all phase bins. "Full dataset": The preferred phase was estimated in the same dataset that is used to test for phase effects, but the phase bin used for alignment was excluded from

the test (i.e. the standard alignment-based approach in this paper). "Known beforehand": The preferred phase was known beforehand (in our case, estimated in another independent dataset of 384 trials) and used to test for phase effects across all phase bins without the necessity of data exclusion. The remaining two bars show equivalent sensitivity of the winning methods in the other two categories. Error bars show the upper limits of 95% confidence intervals.

## 4. Discussion

There is a growing body of evidence for a rhythmic component of human perception and behaviour supported by neural oscillations (Benedetto et al., 2016; Fiebelkorn et al., 2013; Landau and Fries, 2012; VanRullen et al., 2011; Zoefel and VanRullen, 2017). Correspondingly, there is an increased interest in investigating how certain responses (from the detection of simple stimuli such as pure tones, to more complex tasks such as speech comprehension) vary with the phase of neural oscillations (e.g., Baumgarten et al., 2015; Busch et al., 2009; Henry et al., 2014; Henry and Obleser, 2012; Kayser et al., 2016; Mathewson et al., 2009; Ng et al., 2012; Riecke et al., 2018, 2015a, 2015b; Ruzzoli et al., 2019; Strauß et al., 2015; Zoefel et al., 2018a; Zoefel and Heil, 2013; Zoefel and VanRullen, 2015). However, to date, it has been unclear which methods are the most effective for detecting effects of phase on neural or behavioural responses (with the exception of a comprehensive evaluation of methods involving the comparison of two phase distributions, in VanRullen, 2016), and studies have often selected various analysis procedures on a seemingly arbitrary basis. Our goal in this work is to provide criteria to guide readers in choosing statistical procedures to be used for future studies.

### *4.1 Comparing Statistical Approaches*

In all our simulations, we found that, despite being very popular in the published literature, parametric alignment-based methods have a relatively low sensitivity to detect phase effects. In contrast, we observed highest sensitivity for methods which regress single-trial responses in single participants against circular predictor variables and either combine p-values from single participants using Fisher's method (i.e. LOG/LIN REGRESS FISHER), or compare the average regression coefficients or circular-linear correlation coefficients with a surrogate distribution (i.e. LOG/LIN REGRESS PERM or CIRC-LIN CORR). Whereas the latter approach has been used in the past (Busch and VanRullen, 2010; Chakravarthi and Vanrullen, 2012; Kayser, 2019; Kayser et al., 2016; McNair et al., 2019), as far as we know the former method has not been applied in a published manuscript (though, see Liu and Luo, 2019; Tomassini et al., 2017 for a variant of this method).

One of the most popular permutation-methods in the literature is based on testing for differences in phase consistency between hit and miss trials when analysed separately, compared to when analysed as a whole (ITC; used in, e.g., Busch et al., 2009; Dugué et al., 2011; Hanslmayr et al., 2013; Ng et al., 2012; variants of this method are summarised in VanRullen, 2016). The simulations reported here show this to be indeed among the most sensitive methods for dichotomous response variables. However, this method is not intended for, and is therefore less sensitive, to detect a phasic modulation of continuous response variables, for which we found regression-based methods to be optimal.

Although permutation-based methods performed equally well as parametric regression-based methods, they are often more computationally demanding. In addition, randomly re-assigning trials to conditions (as required for the construction of a surrogate distribution) might be problematic where this violates counterbalancing of stimulus materials. Shuffled data can include additional, item-wise variation (which is controlled in the original counter-balanced experimental data; cf. Raaijmakers et al., 1999) and which might thereby lead to invalid results. Consequently, we recommend the use of LOG/LIN REGRESS FISHER as a sensitive and efficient approach to reveal phase effects in all experiments.

### *4.2 Assumptions and Limitations*

The recommendations we give here may be limited to the parameters and parameter ranges tested, and therefore may depend on the assumptions made in our study. For instance, it seems logical that enhanced responses at certain phases (relative to a baseline) have a counterpart in other, opposite phases that lead

to impaired responses relative to the same baseline (VanRullen and McLelland, 2013). Based on this idea, we assumed that positive and negative effect sizes were of equal magnitude. However, the phase profile of positive and negative effects might still differ. Our neural parameters (e.g., *asymmetry*) were based on assuming that neural excitability (and therefore the shape of observed neural oscillations) directly translates into behaviour (Buzsáki and Draguhn, 2004; Lakatos et al., 2005; Peelle and Davis, 2012).

The non-sinusoidal shape of neural oscillations has increasingly been a focus of researchers' attention (Cole and Voytek, 2017; Jones, 2016; Lozano-Soldevilla et al., 2016; Schaworonkow and Nikulin, 2019). To model the effects of non-sinusoidal waveforms on the ability to detect phase effects, we varied both asymmetry of the phase effect and the proportion of the cycle covered by the phase effect *(total width*). In this way, we simulated waveforms that closely resemble those observed in typical electrophysiological recordings. Cole and Voytek (2017) describe the human sensorimotor mu-rhythm with "one extremum (e.g., its peak) consistently sharper than the other (e.g., its trough)". For effects covering the whole cycle (100% *total width*), our asymmetric shapes approximated this description (compare our Fig. 1B, left column, with their Fig. 1A,D). Reductions in *total width* were mostly included to address the possibility that the interaction between neural processes and externally imposed current (e.g., tACS) lead to other, potentially more complex, effect shapes. For instance, it is possible that tACS only affects performance when the applied current is at an extremum (i.e. a peak or trough).

We found that optimal parametric regression-based and permutation-based methods perform in a relatively stable manner across almost all the neural parameters tested here. Consequently, the use of these methods does not require researchers to make a-priori assumptions about the specific nature or shape of the phase effect. Nonetheless, future simulation studies should test how our preferred regression and permutation methods perform on other (e.g., "sawtooth") waveforms that were outside the parameter set used for our simulations.

Finally, there are some plausible experimental scenarios which were not modelled in this study, for example:

(1) We only simulated designs in which phases are selected randomly or controlled externally with similar or equal numbers of trials for each phase, respectively. It remains unclear how methods will perform if certain phases are tested much more often than others. This is potentially an important change as regression-based methods rely on a uniform sampling of phases.

(2) We only simulated a continuous outcome variable with Gaussian distribution and it remains to be determined which other methods will be required to detect a phasic modulation of response variables with more complex distributions (e.g., with a phasic modulation that affects the shape of a distribution but not the mean).

(3) The methods tested in this study employ different strategies to perform inference at the group level (e.g., comparison against chance, combining individual p-values). One common aspect across all methods applied is that single-subject measures of the phase effect can only be either 0 (if no phase effect is present, hence the null hypothesis) or larger than 0 (if a phase effect is present). This can lead to an increased number of false positives in scenarios where only a few subjects show a strong effect, while the rest has no effect. In this context, some authors have used permutation-based methods, similar to those used in our study (Gilron et al., 2017; Stelzer et al., 2013). Others have proposed to test the prevalence of the effect in the population (Allefeld et al., 2016) which, in future work, might represent an interesting alternative to the methods applied here.

### *4.3 Optimising the Detection of Neural Oscillations and their Phase*

This study was designed to compare different statistical methods in their ability to detect a phasic modulation of perception, *given a reliable estimate of phase*. Nevertheless, we here briefly summarise

different issues that might arise when these phases are estimated. As electrical phases are typically imposed externally and therefore known beforehand, some of these issues only concern the estimation of neural phase in electro-/neurophysiological recordings.

First, these recordings are often noisy. This will lead to uncertainty in the measurement of phase (which was not simulated here), in addition to the random fluctuations in the neural or behavioural response (which were simulated here). However, the most likely consequence of this issue is a reduction in the size of the phase effect. In our simulations, the identity of the best method was independent of effect size; thus, the best method would remain unchanged if simulations also included noise in the electrophysiological recordings. In addition, several techniques have been developed recently to reduce the impact of measurement noise (de Cheveigné and Arzounian, 2015; de Cheveigné and Parra, 2014). We also echo others in recommending the cautious use of digital filters (de Cheveigné and Nelken, 2019; Widmann et al., 2015); such filters are often necessary but can lead to spurious phase effects if applied inappropriately (Zoefel and Heil, 2013).

Second, certain variables can fluctuate over the course of an experiment in a way that can complicate reliable estimation of phase. The instantaneous frequency of alpha oscillations (~8-12 Hz) not only varies on an inter-individual level, but also within individuals (Benwell et al., 2019), and depends on factors such as luminance (Benedetto et al., 2018), cognitive load or effort (Klimesch, 2012). If these fluctuations are confined to a relatively narrow range and do not occur too rapidly, common methods, which restrict the signal to a certain frequency range before the phase is estimated (e.g., band-pass filters), should be able to capture them. More sophisticated signal processing methods exist when changes occur more rapidly, such as estimating instantaneous frequency based on peaks and troughs in the signal (Cole and Voytek, 2019). It is also plausible that the neural or behavioural response (e.g., its average across phases, termed "baseline" in this study) fluctuates over time. If these changes do not occur too rapidly (i.e. do not occur only at some of the tested phases), they should only affect the baseline but not any phasic modulation. As different baselines did not affect our results (not shown), our recommendations would remain unchanged. In addition, adaptation procedures ("staircase") can be used to take into account fatigue or learning effects (Leek, 2001).

Third, even though the *detection* of phase effects (given a reliable phase estimate) does not seem to be strongly affected by non-sinusoidal waveforms (see previous section), the *estimation* of phase might be. Many conventional methods (e.g., Fourier Transform) assume stationary sinusoidal signals and can be biased if this assumption is not fulfilled. The approach by Cole and Voytek (2019), described above, takes this issue into account and estimates the symmetry of the extracted signal. However, it remains to be shown if this method can be applied reliably if the frequency of interest is not dominant in the recorded signal (and the estimation of peaks and troughs is difficult).

Finally, an independent question is whether the estimated phase reflects true oscillatory activity. In principle, the methods proposed here can be applied to any type of signal. The application of a narrow band-pass filter will reveal seemingly oscillatory signals at the frequency of interest, even if the data does not contain such oscillations (de Cheveigné and Nelken, 2019). Rhythmic input will necessarily lead to rhythmic brain responses, but this might merely reflect the rhythmicity of the stimulus without involving true neural oscillations (as reviewed by Zoefel et al., 2018b). This issue is increasingly recognised, and recently developed methods to identify true oscillatory activity seem promising (Doelling et al., 2019; Haller et al., 2018; Watrous et al., 2018).

### *4.4. Optimising the Detection of Phasic Modulation or Enhancement/Impairment of Behaviour*

The methods presented here serve to detect a phasic *modulation* of neural or behavioural responses. If present, this modulation should reflect both in- and decreased responses relative to the average across all phases (VanRullen and McLelland, 2013). Nonetheless, when neural oscillations are perturbed by rhythmic sensory or electrical stimulation, it can be important to determine whether behaviour is

enhanced, impaired, or both (but at different phases), as compared to an unstimulated baseline condition, e.g. sham stimulation (Riecke and Zoefel, 2018). It is possible that other statistical methods than those explored here will be optimal for this purpose. Future simulation studies could therefore explore statistical methods for comparison of phasic effects with non-phasic control conditions. Answering the question of whether tACS enhances behaviour relative to a control condition is of critical importance for practical applications of non-invasive brain stimulation. Only interventions that enhance desired behaviours or impair non-desired behaviours might find clinical or practical application.

We found that the most crucial experimental parameter for modulating the overall ability to detect phase effects (i.e. without necessarily influencing the identity of the best method) is the number of trials. Given this, it seems advisable to prefer simpler experimental designs with a relatively high number of trials. As evidenced by similar sensitivity for *randomly sampled* and *externally imposed* phases, determining phases post-hoc does not strongly reduce the detectability of phase effects, nor does testing only a limited number of phases (at least for methods which do not divide data into phase bins). Perhaps not surprisingly, we further observed that the most important neural parameter is the size of the hypothesized effect, which also improved sensitivity in a monotonic fashion, especially in the range of effect sizes that is commonly observed (e.g., 5-15% peak-to-peak sinusoidal modulation of behaviour). The low sensitivity of typical experiments for weak phase effects (d-prime of 1 or lower) is particularly relevant for tACS studies in which reported effect sizes are small (Gundlach et al., 2016; Neuling et al., 2012; Riecke et al., 2015a, 2015b, 2018; Riecke and Zoefel, 2018; Wilsch et al., 2018; Zoefel et al., 2018a). Transcranial brain stimulation has recently been subjected to pronounced criticism, challenging its efficacy at manipulating neural responses or behaviour (Lafon et al., 2017; Vöröslakos et al., 2018; but see Krause et al., 2019; Ruhnau et al., 2018). Our results suggest that the inconsistency of results might be partly due to the insensitivity of popular statistical methods to detect weak effects on behaviour. We anticipate two potential solutions to this problem: improving the (1) sensitivity of analysis methods (as proposed in the present paper) and/or (2) improving the efficacy of stimulation protocols (i.e. increasing neural effect sizes). It has recently been suggested that effect sizes might also be increased by adapting stimulation protocols to individual participants, as discussed elsewhere (Romei et al., 2016; Thut et al., 2017). However, these methods intrinsically depend on optimal measurement of stimulation effects on behaviour and hence are probably best built on statistical simulations like those presented here.

The pre-registration of experimental paradigms and analyses is rapidly becoming an essential step in scientific research (Nosek et al., 2018). However, the definition of detailed analysis plans in a young field of research lacking established routine procedures remains challenging (Ledgerwood, 2018). The investigation of oscillatory phase effects on behaviour is one such field in which a proliferation of analysis methods can make it difficult for researchers seeking to pre-register their analysis protocols. We hope that the simulations reported here lay the groundwork for pre-registration of tACS, EEG/MEG and other studies exploring the effect of phase on neural or behavioural responses.

Acknowledgements

This work was supported by a grant from the German Academic Exchange Service (DAAD), the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement number 743482, and the Medical Research Council UK (grant number SUAG/008/RG91365). The authors thank Rufin VanRullen for helpful advice, and Nicholas Bland for providing code for the LIN REGRESS method.

References

Ai, L., Ro, T., 2014. The phase of prestimulus alpha oscillations affects tactile perception. J. Neurophysiol. 111, 1300–1307. https://doi.org/10.1152/jn.00125.2013

Allefeld, C., Görgen, K., Haynes, J.-D., 2016. Valid population inference for information-based imaging: From the second-level t-test to prevalence inference. NeuroImage 141, 378–392. https://doi.org/10.1016/j.neuroimage.2016.07.040
Asamoah, B., Khatoun, A., Mc Laughlin, M., 2019. Analytical bias accounts for some of the reported effects of tACS on auditory perception. Brain Stimulat. 12, 1001–1009. https://doi.org/10.1016/j.brs.2019.03.011
Baumgarten, T.J., Schnitzler, A., Lange, J., 2015. Beta oscillations define discrete perceptual cycles in the somatosensory domain. Proc. Natl. Acad. Sci. U. S. A. 112, 12187–12192. https://doi.org/10.1073/pnas.1501438112
Benedetto, A., Lozano-Soldevilla, D., VanRullen, R., 2018. Different responses of spontaneous and stimulus-related alpha activity to ambient luminance changes. Eur. J. Neurosci. 48, 2599–2608. https://doi.org/10.1111/ejn.13791
Benedetto, A., Spinelli, D., Morrone, M.C., 2016. Rhythmic modulation of visual contrast discrimination triggered by action. Proc. Biol. Sci. 283. https://doi.org/10.1098/rspb.2016.0692
Benwell, C.S.Y., London, R.E., Tagliabue, C.F., Veniero, D., Gross, J., Keitel, C., Thut, G., 2019. Frequency and power of human alpha oscillations drift systematically with time-on-task. NeuroImage 192, 101–114. https://doi.org/10.1016/j.neuroimage.2019.02.067
Busch, N.A., Dubois, J., VanRullen, R., 2009. The phase of ongoing EEG oscillations predicts visual perception. J. Neurosci. 29, 7869–7876. https://doi.org/10.1523/JNEUROSCI.0113-09.2009
Busch, N.A., VanRullen, R., 2010. Spontaneous EEG oscillations reveal periodic sampling of visual attention. Proc. Natl. Acad. Sci. U. S. A. 107, 16048–16053. https://doi.org/10.1073/pnas.1004801107
Buzsáki, G., Draguhn, A., 2004. Neuronal oscillations in cortical networks. Science 304, 1926–1929. https://doi.org/10.1126/science.1099745
Chakravarthi, R., Vanrullen, R., 2012. Conscious updating is a rhythmic process. Proc. Natl. Acad. Sci. U. S. A. 109, 10599–10604. https://doi.org/10.1073/pnas.1121622109
Cole, S., Voytek, B., 2019. Cycle-by-cycle analysis of neural oscillations. J. Neurophysiol. https://doi.org/10.1152/jn.00273.2019
Cole, S.R., Voytek, B., 2017. Brain Oscillations and the Importance of Waveform Shape. Trends Cogn. Sci. 21, 137–149. https://doi.org/10.1016/j.tics.2016.12.008
de Cheveigné, A., Arzounian, D., 2015. Scanning for oscillations. J. Neural Eng. 12, 066020. https://doi.org/10.1088/1741-2560/12/6/066020
de Cheveigné, A., Nelken, I., 2019. Filters: When, Why, and How (Not) to Use Them. Neuron 102, 280–293. https://doi.org/10.1016/j.neuron.2019.02.039
de Cheveigné, A., Parra, L.C., 2014. Joint decorrelation, a versatile tool for multichannel data analysis. NeuroImage 98, 487–505. https://doi.org/10.1016/j.neuroimage.2014.05.068
Doelling, K.B., Assaneo, M.F., Bevilacqua, D., Pesaran, B., Poeppel, D., 2019. An oscillator model better predicts cortical entrainment to music. Proc. Natl. Acad. Sci. 201816414. https://doi.org/10.1073/pnas.1816414116
Dugué, L., Marque, P., VanRullen, R., 2011. The phase of ongoing oscillations mediates the causal relation between brain excitation and visual perception. J. Neurosci. 31, 11889–11893. https://doi.org/10.1523/JNEUROSCI.1161-11.2011
Eliav, T., Geva-Sagiv, M., Yartsev, M.M., Finkelstein, A., Rubin, A., Las, L., Ulanovsky, N., 2018. Nonoscillatory Phase Coding and Synchronization in the Bat Hippocampal Formation. Cell 175, 1119-1130.e15. https://doi.org/10.1016/j.cell.2018.09.017
Fiebelkorn, I.C., Saalmann, Y.B., Kastner, S., 2013. Rhythmic sampling within and between objects despite sustained attention at a cued location. Curr. Biol. 23, 2553–2558. https://doi.org/10.1016/j.cub.2013.10.063
Fisher, R.A., 1950. Statistical Methods for Research Workers, 11th ed. Oliver and Boyd.
Good, P.I., 2005. Permutation, Parametric, and Bootstrap Tests of Hypotheses, 3rd ed, Springer Series in Statistics. Springer, New York.
Gundlach, C., Müller, M.M., Nierhaus, T., Villringer, A., Sehm, B., 2016. Phasic Modulation of Human Somatosensory Perception by Transcranially Applied Oscillating Currents. Brain Stimulat. 9, 712–719. https://doi.org/10.1016/j.brs.2016.04.014
Haller, M., Donoghue, T., Peterson, E., Varma, P., Sebastian, P., Gao, R., Noto, T., Knight, R.T., Shestyuk, A., Voytek, B., 2018. Parameterizing neural power spectra. bioRxiv 299859. https://doi.org/10.1101/299859
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. Curr. Biol. 23, 2273–2278. https://doi.org/10.1016/j.cub.2013.09.020

Harris, A.M., Dux, P.E., Mattingley, J.B., 2018. Detecting Unattended Stimuli Depends on the Phase of Prestimulus Neural Oscillations. J. Neurosci. 38, 3092–3101. https://doi.org/10.1523/JNEUROSCI.3006-17.2018
Henry, M.J., Herrmann, B., Obleser, J., 2014. Entrained neural oscillations in multiple frequency bands comodulate behavior. Proc. Natl. Acad. Sci. U. S. A. 111, 14935–14940. https://doi.org/10.1073/pnas.1408741111
Henry, M.J., Obleser, J., 2012. Frequency modulation entrains slow neural oscillations and optimizes human listening behavior. Proc. Natl. Acad. Sci. U. S. A. 109, 20095–20100. https://doi.org/10.1073/pnas.1213390109
Herrmann, C.S., Rach, S., Neuling, T., Strüber, D., 2013. Transcranial alternating current stimulation: a review of the underlying mechanisms and modulation of cognitive processes. Front. Hum. Neurosci. 7, 279. https://doi.org/10.3389/fnhum.2013.00279
Hickok, G., Farahbod, H., Saberi, K., 2015. The Rhythm of Perception: Entrainment to Acoustic Rhythms Induces Subsequent Perceptual Oscillation. Psychol. Sci. 26, 1006–1013. https://doi.org/10.1177/0956797615576533
Jesteadt, W., 2005. The variance of d′ estimates obtained in yes—no and two-interval forced choice procedures. Percept. Psychophys. 67, 72–80. https://doi.org/10.3758/BF03195013
Jones, S.R., 2016. When brain rhythms aren't 'rhythmic': implication for their mechanisms and meaning. Curr. Opin. Neurobiol., Systems neuroscience 40, 72–80. https://doi.org/10.1016/j.conb.2016.06.010
Kayser, C., 2019. Evidence for the rhythmic perceptual sampling of auditory scenes. bioRxiv 618652. https://doi.org/10.1101/618652
Kayser, C., Wilson, C., Safaai, H., Sakata, S., Panzeri, S., 2015. Rhythmic Auditory Cortex Activity at Multiple Timescales Shapes Stimulus–Response Gain and Background Firing. J. Neurosci. 35, 7750–7762. https://doi.org/10.1523/JNEUROSCI.0268-15.2015
Kayser, S.J., McNair, S.W., Kayser, C., 2016. Prestimulus influences on auditory perception from sensory representations and decision processes. Proc. Natl. Acad. Sci. 113, 4842–4847. https://doi.org/10.1073/pnas.1524087113
Keitel, C., Thut, G., Gross, J., 2017. Visual cortex responses reflect temporal structure of continuous quasi-rhythmic sensory stimulation. NeuroImage 146, 58–70. https://doi.org/10.1016/j.neuroimage.2016.11.043
Klimesch, W., 2012. Alpha-band oscillations, attention, and controlled access to stored information. Trends Cogn. Sci. 16, 606–617. https://doi.org/10.1016/j.tics.2012.10.007
Krause, M.R., Vieira, P.G., Csorba, B.A., Pilly, P.K., Pack, C.C., 2019. Transcranial alternating current stimulation entrains single-neuron activity in the primate brain. Proc. Natl. Acad. Sci. 116, 5747–5755. https://doi.org/10.1073/pnas.1815958116
Lafon, B., Henin, S., Huang, Y., Friedman, D., Melloni, L., Thesen, T., Doyle, W., Buzsáki, G., Devinsky, O., Parra, L.C., A Liu, A., 2017. Low frequency transcranial electrical stimulation does not entrain sleep rhythms measured by human intracranial recordings. Nat. Commun. 8, 1199. https://doi.org/10.1038/s41467-017-01045-x
Lakatos, P., Shah, A.S., Knuth, K.H., Ulbert, I., Karmos, G., Schroeder, C.E., 2005. An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. J. Neurophysiol. 94, 1904–1911. https://doi.org/10.1152/jn.00263.2005
Landau, A.N., Fries, P., 2012. Attention samples stimuli rhythmically. Curr. Biol. 22, 1000–1004. https://doi.org/10.1016/j.cub.2012.03.054
Ledgerwood, A., 2018. The preregistration revolution needs to distinguish between predictions and analyses. Proc. Natl. Acad. Sci. 201812592. https://doi.org/10.1073/pnas.1812592115
Leek, M.R., 2001. Adaptive procedures in psychophysical research. Percept. Psychophys. 63, 1279–1292. https://doi.org/10.3758/BF03194543
Liu, L., Luo, H., 2019. Behavioral oscillation in global/local processing: Global alpha oscillations mediate global precedence effect. J. Vis. 19, 12–12. https://doi.org/10.1167/19.5.12
Lozano-Soldevilla, D., ter Huurne, N., Oostenveld, R., 2016. Neuronal Oscillations with Non-sinusoidal Morphology Produce Spurious Phase-to-Amplitude Coupling and Directionality. Front. Comput. Neurosci. 10. https://doi.org/10.3389/fncom.2016.00087
Madsen, K.H., Karabanov, A.N., Krohne, L.G., Safeldt, M.G., Tomasevic, L., Siebner, H.R., 2019. No trace of phase: Corticomotor excitability is not tuned by phase of pericentral mu-rhythm. Brain Stimulat. https://doi.org/10.1016/j.brs.2019.05.005
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. J. Neurosci. 29, 2725–2732. https://doi.org/10.1523/JNEUROSCI.3963-08.2009

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. J. Cogn. Neurosci. 24, 2321–2333. https://doi.org/10.1162/jocn_a_00288
McNair, S.W., Kayser, S.J., Kayser, C., 2019. Consistent pre-stimulus influences on auditory perception across the lifespan. Neuroimage 186, 22–32. https://doi.org/10.1016/j.neuroimage.2018.10.085
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. https://doi.org/10.1016/j.neuroimage.2016.02.065
Neuling, T., Rach, S., Wagner, S., Wolters, C.H., Herrmann, C.S., 2012. Good vibrations: oscillatory phase shapes perception. NeuroImage 63, 771–778. https://doi.org/10.1016/j.neuroimage.2012.07.024
Ng, B.S.W., Schroeder, T., Kayser, C., 2012. A precluding but not ensuring role of entrained low-frequency oscillations for auditory perception. J. Neurosci. 32, 12268–12276. https://doi.org/10.1523/JNEUROSCI.1877-12.2012
Nosek, B.A., Ebersole, C.R., DeHaven, A.C., Mellor, D.T., 2018. The preregistration revolution. Proc. Natl. Acad. Sci. 201708274. https://doi.org/10.1073/pnas.1708274114
Peelle, J.E., Davis, M.H., 2012. Neural Oscillations Carry Speech Rhythm through to Comprehension. Front. Psychol. 3, 320. https://doi.org/10.3389/fpsyg.2012.00320
Raaijmakers, J.G.W., Schrijnemakers, J.M.C., Gremmen, F., 1999. How to Deal with "The Language-as-Fixed-Effect Fallacy": Common Misconceptions and Alternative Solutions. J. Mem. Lang. 41, 416–426. https://doi.org/10.1006/jmla.1999.2650
Riecke, L., Formisano, E., Herrmann, C.S., Sack, A.T., 2015a. 4-Hz Transcranial Alternating Current Stimulation Phase Modulates Hearing. Brain Stimulat. 8, 777–783. https://doi.org/10.1016/j.brs.2015.04.004
Riecke, L., Formisano, E., Sorger, B., Başkent, D., Gaudrain, E., 2018. Neural Entrainment to Speech Modulates Speech Intelligibility. Curr. Biol. 28, 161-169.e5. https://doi.org/10.1016/j.cub.2017.11.033
Riecke, L., Sack, A.T., Schroeder, C.E., 2015b. Endogenous Delta/Theta Sound-Brain Phase Entrainment Accelerates the Buildup of Auditory Streaming. Curr. Biol. 25, 3196–3201. https://doi.org/10.1016/j.cub.2015.10.045
Riecke, L., Zoefel, B., 2018. Conveying Temporal Information to the Auditory System via Transcranial Current Stimulation. Acta Acust. United Acust. 104, 883–886. https://doi.org/10.3813/AAA.919235
Romei, V., Thut, G., Silvanto, J., 2016. Information-Based Approaches of Noninvasive Transcranial Brain Stimulation. Trends Neurosci. 39, 782–795. https://doi.org/10.1016/j.tins.2016.09.001
Ronconi, L., Oosterhof, N.N., Bonmassar, C., Melcher, D., 2017. Multiple oscillatory rhythms determine the temporal organization of perception. Proc. Natl. Acad. Sci. U. S. A. 114, 13435–13440. https://doi.org/10.1073/pnas.1714522114
Ruhnau, P., Rufener, K.S., Heinze, H.-J., Zaehle, T., 2018. Sailing in a sea of disbelief: In vivo measurements of transcranial electric stimulation in human subcortical structures. Brain Stimul. 11, 241–243. https://doi.org/10.1016/j.brs.2017.09.015
Ruzzoli, M., Torralba, M., Fernández, L.M., Soto-Faraco, S., 2019. The relevance of alpha phase in human perception. Cortex. https://doi.org/10.1016/j.cortex.2019.05.012
Schaworonkow, N., Nikulin, V.V., 2019. Spatial neuronal synchronization and the waveform of oscillations: Implications for EEG and MEG. PLOS Comput. Biol. 15, e1007055. https://doi.org/10.1371/journal.pcbi.1007055
Schaworonkow, N., Triesch, J., Ziemann, U., Zrenner, C., 2019. EEG-triggered TMS reveals stronger brain state-dependent modulation of motor evoked potentials at weaker stimulation intensities. Brain Stimulat. 12, 110–118. https://doi.org/10.1016/j.brs.2018.09.009
Spaak, E., de Lange, F.P., Jensen, O., 2014. Local entrainment of α oscillations by visual stimuli causes cyclic modulation of perception. J. Neurosci. 34, 3536–3544. https://doi.org/10.1523/JNEUROSCI.4385-13.2014
Stelzer, J., Chen, Y., Turner, R., 2013. Statistical inference and multiple testing correction in classification-based multi-voxel pattern analysis (MVPA): random permutations and cluster size control. NeuroImage 65, 69–82. https://doi.org/10.1016/j.neuroimage.2012.09.063
Strauß, A., Henry, M.J., Scharinger, M., Obleser, J., 2015. Alpha phase determines successful lexical decision in noise. J. Neurosci. 35, 3256–3262. https://doi.org/10.1523/JNEUROSCI.3357-14.2015
ten Oever, S., Sack, A.T., 2015. Oscillatory phase shapes syllable perception. Proc. Natl. Acad. Sci. U. S. A. 112, 15833–15837. https://doi.org/10.1073/pnas.1517519112
Thut, G., Bergmann, T.O., Fröhlich, F., Soekadar, S.R., Brittain, J.-S., Valero-Cabré, A., Sack, A.T., Miniussi, C., Antal, A., Siebner, H.R., Ziemann, U., Herrmann, C.S., 2017. Guiding transcranial brain stimulation by EEG/MEG to interact with ongoing brain activity and associated functions: A position paper. Clin. Neurophysiol. 128, 843–857. https://doi.org/10.1016/j.clinph.2017.01.003

Tomassini, A., Ambrogioni, L., Medendorp, W.P., Maris, E., 2017. Theta oscillations locked to intended actions rhythmically modulate perception. eLife 6. https://doi.org/10.7554/eLife.25618
VanRullen, R., 2016. How to Evaluate Phase Differences between Trial Groups in Ongoing Electrophysiological Signals. Front. Neurosci. 10, 426. https://doi.org/10.3389/fnins.2016.00426
VanRullen, R., Busch, N.A., Drewes, J., Dubois, J., 2011. Ongoing EEG Phase as a Trial-by-Trial Predictor of Perceptual and Attentional Variability. Front. Psychol. 2. https://doi.org/10.3389/fpsyg.2011.00060
VanRullen, R., McLelland, D., 2013. What goes up must come down: EEG phase modulates auditory perception in both directions. Front. Psychol. 4. https://doi.org/10.3389/fpsyg.2013.00016
Vöröslakos, M., Takeuchi, Y., Brinyiczki, K., Zombori, T., Oliva, A., Fernández-Ruiz, A., Kozák, G., Kincses, Z.T., Iványi, B., Buzsáki, G., Berényi, A., 2018. Direct effects of transcranial electric stimulation on brain circuits in rats and humans. Nat. Commun. 9, 483. https://doi.org/10.1038/s41467-018-02928-3
Watrous, A.J., Miller, J., Qasim, S.E., Fried, I., Jacobs, J., 2018. Phase-tuned neuronal firing encodes human contextual representations for navigational goals. eLife 7, e32554. https://doi.org/10.7554/eLife.32554
Widmann, A., Schröger, E., Maess, B., 2015. Digital filter design for electrophysiological data--a practical approach. J. Neurosci. Methods 250, 34–46. https://doi.org/10.1016/j.jneumeth.2014.08.002
Wilsch, A., Neuling, T., Obleser, J., Herrmann, C.S., 2018. Transcranial alternating current stimulation with speech envelopes modulates speech comprehension. NeuroImage 172, 766–774. https://doi.org/10.1016/j.neuroimage.2018.01.038
Wutz, A., Muschter, E., van Koningsbruggen, M.G., Weisz, N., Melcher, D., 2016. Temporal Integration Windows in Neural Processing and Perception Aligned to Saccadic Eye Movements. Curr. Biol. 26, 1659–1668. https://doi.org/10.1016/j.cub.2016.04.070
Zoefel, B., Archer-Boyd, A., Davis, M.H., 2018a. Phase Entrainment of Brain Oscillations Causally Modulates Neural Responses to Intelligible Speech. Curr. Biol. 28, 401-408.e5. https://doi.org/10.1016/j.cub.2017.11.071
Zoefel, B., Davis, M.H., 2017. Transcranial electric stimulation for the investigation of speech perception and comprehension. Lang. Cogn. Neurosci. 32, 910–923. https://doi.org/10.1080/23273798.2016.1247970
Zoefel, B., Heil, P., 2013. Detection of Near-Threshold Sounds is Independent of EEG Phase in Common Frequency Bands. Front. Psychol. 4, 262. https://doi.org/10.3389/fpsyg.2013.00262
Zoefel, B., ten Oever, S., Sack, A.T., 2018b. The Involvement of Endogenous Neural Oscillations in the Processing of Rhythmic Input: More Than a Regular Repetition of Evoked Neural Responses. Front. Neurosci. 12. https://doi.org/10.3389/fnins.2018.00095
Zoefel, B., VanRullen, R., 2017. Oscillatory Mechanisms of Stimulus Processing and Selection in the Visual and Auditory Systems: State-of-the-Art, Speculations and Suggestions. Front. Neurosci. 11. https://doi.org/10.3389/fnins.2017.00296
Zoefel, B., VanRullen, R., 2015. Selective perceptual phase entrainment to speech rhythm in the absence of spectral energy fluctuations. J. Neurosci.. 35, 1954–1964. https://doi.org/10.1523/JNEUROSCI.3484-14.2015