bioRxiv preprint doi: https://doi.org/10.1101/517243; this version posted January 11, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

How to test for oscillatory modulation of neural function and behaviour

Authors: Benedikt Zoefel[a,*], Matthew H Davis[a], 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

Number of pages: 22
Number of figures: 7

Keywords: neural oscillations, phase, EEG, MEG, tACS, simulations, shape, behaviour

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.

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

Highlights:

- **The phase of neural oscillations modulates neural responses and behaviour**
- **We test the ability of different methods to detect such a phase effect**
- **We test how this ability changes with nature of effect and experimental parameters**
- **Results represent groundwork for future investigations of oscillatory phase effects**

bioRxiv preprint doi: https://doi.org/10.1101/517243; this version posted January 11, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

The question whether perception or other processes depend on the phase of neural oscillations is a research topic that is rapidly gaining popularity. Importantly 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 and reject an absent one. We reveal optimal methods for different parameters that characterise the phase effect or depend on the experimental design to test for it. We demonstrate that permutation-based methods perform better than parametric ones for all of the simulated parameters. The identity of the optimal parametric method varies with several parameters, in particular with the number of phase bins chosen for the analysis and the width of the phase effect. In contrast, the identity of the best permutation-based method is unaffected by changes in the simulated parameters. In sum, our study lays a foundation for optimized experimental designs and analyses in future studies investigating the role of neural phase for behaviour and neural function. We supply MATLAB code that can be used to identify the ideal method (among all of the methods tested) for a specific scientific hypothesis and set of experimental parameters.

1. Introduction

Neural oscillations are cyclic variations in the excitability of neuronal ensembles. The 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 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; Milton and Pleydell-Pearce, 2016; Ronconi et al., 2017). 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 performance depends on the phase of spontaneous (i.e. stimulus-unrelated) oscillations extracted post-hoc from concurrent EEG/MEG recordings.

An alternative paradigm exploits that neural oscillations align to rhythmically applied input (termed neural entrainment; Lakatos et al., 2008). In such studies, the stimulus phase serves as a surrogate for oscillatory phase. Similarly to the studies described above, changes in perception and other processes are assessed as a function of the presumed neural phase. However, as the phase is imposed externally in each trial, it is under the control of the experimenter, and often only a small number of phase values is tested. Apart from sensory rhythmic stimuli (Henry and Obleser, 2012; Hickok et al., 2015; Keitel et al., 2017; Mathewson et al., 2012; Spaak et al., 2014), a commonly used entrainment stimulus is tACS (Herrmann et al., 2013; Riecke and Zoefel, 2018; Zoefel and Davis, 2017). For instance, it has recently been shown that both the neural responses to speech and speech comprehension depend on the phase of tACS relative to the speech rhythm (Riecke et al., 2018; Wilsch et al., 2018; Zoefel et al., 2018).

Irrespective of the exact paradigm used, a question common to all the studies described above is whether and how perception or other measures are influenced by neural phase. To test this, one common statistical approach is to compare performance across trials grouped into phase bins using *parametric tests* (Baumgarten et al., 2015; Busch et al., 2009; Chakravarthi and Vanrullen, 2012; Neuling et al., 2012; Ng et al., 2012; Riecke et al., 2015a, 2015b, 2018; Wilsch et al., 2018; Zoefel et al., 2018; 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 performance; see Fig. 1A and 2A; e.g., Busch et al., 2009; Riecke et al., 2018). Consequently, simply averaging performance as a function of phase across participants can lead to phase cancellation effects at the group level and reduce the detectability of an effect. A common solution to this problem is to report data from the phase bins that are re-aligned relative to the preferred phase in each participant. However, the performance for the phase bin used for alignment is trivially the maximum at the group level. This bias makes it mandatory to exclude this

bioRxiv preprint doi: https://doi.org/10.1101/517243; this version posted January 11, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

specific phase bin, and therefore reduces important phase-related variance from subsequent analyses of the aligned data. Moreover, whether the participant's maximum performance is optimal for alignment or other approaches are more appropriate is still unclear.

A second common statistical approach is the use of *non-parametric, permutation tests*: 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, for instance, be constructed by repeatedly shuffling performance labels (e.g., "hits" vs "misses") and re-running the original analysis on the permuted datasets (Busch et al., 2009; Dugué et al., 2011; Hanslmayr et al., 2013; Ng et al., 2012; Ronconi et al., 2017; Strauß et al., 2015; ten Oever and Sack, 2015; Wutz et al., 2016).

It is currently unclear which of these statistical approaches is optimal for detecting effects of oscillatory phase on behaviour. In addition, it has become increasingly evident that neural oscillations are not perfect sinusoids (Cole and Voytek, 2017; Jones, 2016). Consequently, certain characteristics of a "genuine" effect of oscillatory phase (as shown in Fig. 1A), and how they affect our ability to detect such a phase effect, remain unknown. Indeed, one important question is whether there is one statistical approach that will be optimal for all forms of data.

In this study, we conducted Monte-Carlo simulations to determine how reliably each of several different statistical methods can detect an oscillatory modulation of performance (and reject an absent one). We tested how two sets of parameters impact the sensitivity of these methods: Parameters that describe the nature of the phase effect and reflect underlying neural processes (effect size, width, and asymmetry; see Fig. 1 and Material and Methods), henceforth termed *neural parameters*, and parameters that depend on the specific experimental design, termed *experimental parameters*. The latter include the number of trials (N_{trials}) and the number of phase bins (N_{bins}) used in the analysis, and a binary distinction between the two alternative paradigms described above (i.e., experiments in which phases are sampled randomly and sorted post-hoc [termed *randomly selected*] or controlled experimentally using sensory or electric stimulation [termed *imposed externally*]). Finally, we contrasted the two categories of statistical methods described above (*parametric* vs *non-parametric, permutation-based* methods). Together, the results of our simulations serve to guide researchers in choosing the most appropriate method to detect oscillatory phase effects in future experiments.

bioRxiv preprint doi: https://doi.org/10.1101/517243; this version posted January 11, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

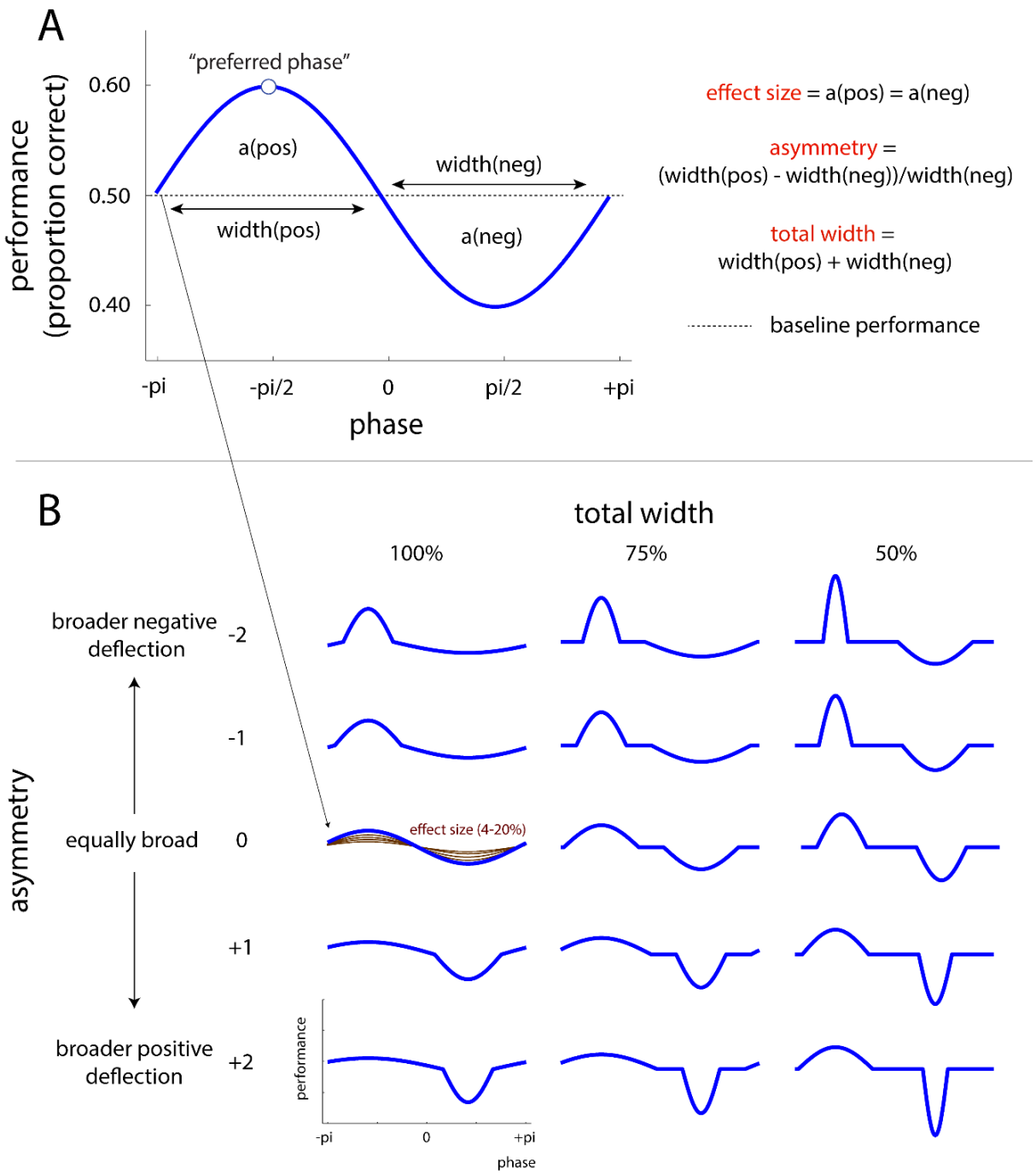

Figure 1. **A**. Definition of neural parameters. **B**. Simulated values for neural parameters. These parameters produce different effect shapes that can be sinusoidal (*asymmetry:* 0, *total width:* 100%) or not. Note that *effect size* is identical for all of the blue curves shown. For one exemplary combination of asymmetry and total width, all possible *effect sizes* are shown in brown.

bioRxiv preprint doi: https://doi.org/10.1101/517243; this version posted January 11, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

2. Material and Methods

Scripts to run the simulations and analyses, as well as the simulated data, are available in the following repository: <link will be provided upon acceptance>.

2.1 Simulating experimental data

We explored how two sets of parameters affect the ability of different methods to detect oscillatory phase effects: Neural parameters, which characterize the nature of the phase effect, and experimental parameters which describe the design to test for the phase effect. We first created a detailed model of the phase effect, based on simulated neural parameters (see Neural Parameters). We then sampled this effect in a simulated experiment, based on simulated experimental parameters (see Experimental Parameters). For each combination of neural and experimental parameters, we simulated 1000 experiments in which a phase-induced modulation of performance is present. In addition, we simulated 1000 experiments in which this effect is absent. Each of the simulated (2 x 1000) experiments consisted of data from 20 virtual participants. We then applied various methods (see Statistical Methods) to the simulated data and tested, for each method, how many of the 1000 experiments involving an effect were correctly identified as such (true positives, or hits) and how many of the 1000 experiments involving no effect were incorrectly labelled as an effect (false positives, or false alarms). For each of the simulated combinations of parameters, we were able to define a method that is optimally suited (among all of the methods tested) to detect a phase effect (see Evaluation of Methods).

2.1.1 Neural parameters

To create a detailed model of phase effects, we defined a performance value (hit probability or similar) for phase values spanning a full oscillatory cycle and spaced equidistantly in steps of 1.875 degrees, i.e. 192 performance values over a full cycle. How performance varied with phase depended on the simulated neural parameters, as described in the following and shown in Fig. 1.

(1) *Baseline performance:* For all experiments, the average hit probability across phase was chosen to be 50% (reflecting, for instance, the probability of detecting a randomly presented near-threshold stimulus). Using other baselines (e.g., a baseline of 0% might be more adequate if performance is quantified as the difference between stimulation and sham conditions in the case of tACS) did not change the outcome of these simulations.

(2) *Preferred phase:* For experiments with a phase effect present, a phase value was chosen randomly as the peak of a positive deflection from baseline performance (again, with a resolution of 1.875 degrees, i.e. there were 192 possibilities). The phase bin π (or 180°) away from the positive peak was chosen as the peak for the negative deflection. 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 performance and either of the two (positive or 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 the modulation of performance in all phase bins (rather than only peak performance). We assumed that a phasic enhancement of performance (i.e. phase values yielding better than average performance) have a matched counterpart of phase values leading to impairment of performance (i.e. worse than average performance), 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

bioRxiv preprint doi: https://doi.org/10.1101/517243; this version posted January 11, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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 cover the desired amount of phase values (Fig. 1B). The total width of the phase effect was defined as $width_{total} = width_{pos} + width_{neg}$. We simulated total widths of 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

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

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 constant effect size, a change in asymmetry results in change in peak amplitudes, as visible in Fig. 1B.

2.1.2 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 behaviour. However, researchers have a range of options available in choosing an experimental protocol by which phase effects on behaviour are measured; we therefore simulated various versions of such a protocol. In each simulated experiment, a virtual stimulus was presented at a specific neural (or electrical) phase, and an outcome (e.g., "hit" or "miss") was determined from the detailed phase effect model for each trial. The phase effect (Fig. 1) was thus sampled from the model using a specific combination of experimental parameters (described below), leading to a performance profile (hit probability) as a function of phase for each virtual participant (Fig. 2A). We simulated our data at the trial level as follows: For each trial, a phase was selected (in two different ways, depending on the parameter *study design* described below), and the outcome ("hit" or "miss") of the trial was determined from detailed phase effect model, that is, based on the hit probability associated with the phase:

if rand(1) >= 1-p(hit, phase), then outcome = "hit"

otherwise: outcome = "miss"

where *rand* corresponds to a random value drawn from a standard uniform distribution and *p(hit, phase)* is the hit probability for a given phase in the model. Note that this hit probability (i.e., the detailed phase effect model) depends on the chosen neural parameters (see (1) to (5) above). For instance, in the example shown in Fig. 1A, a stimulus presented at phase -π/2 would have the probability of 60% of being detected (hit), whereas stimulus presented at phase +π/2 would have a 40% hit probability. From the set of simulated trials, average performance (hit probability) was quantified for each phase bin (shown in Fig. 2A). Note that simulated data for single participants will include random variability (or "noise") depending on the number of trials sampled and other parameters described below. To define and test various experimental protocols, we systematically varied the following experimental parameters:

(6) 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.

(7) N_{bins}*:* The number of phase bins used for the analysis was 4, 6, 8, 12, or 16. Even though electrophysiological data (EEG/MEG) can in theory be analysed using an indefinite number of phase bins (because stimuli are presented at a-priori unknown, random neural phases), this is rarely the case in practice: In typical analyses, data is sorted into phase bins in order to increase the number of trials per bin and, consequently, statistical power (indeed, some of the methods below crucially rely on a

bioRxiv preprint doi: https://doi.org/10.1101/517243; this version posted January 11, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

finite number of phase bins). 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 for each participant.

(8) *Study design:* We tested two common paradigms for how phases were selected for each trial (see Introduction). In the first paradigm (*randomly selected*), phases were randomly selected in each (virtual) trial (with 192 possible phase values as described above). In the second paradigm (*imposed externally*), a small number of phase values was tested. In this case, we assumed that the tested phase values corresponded to the centres of the phase bins used for the analysis (see parameter 7).

Together, we simulated 2 x 1000 experiments (effect present and effect absent) for all combinations of neural parameters (*effect size* x *total width* x *asymmetry*) and experimental parameters (N_{trials} x N_{bins} x *study design*).

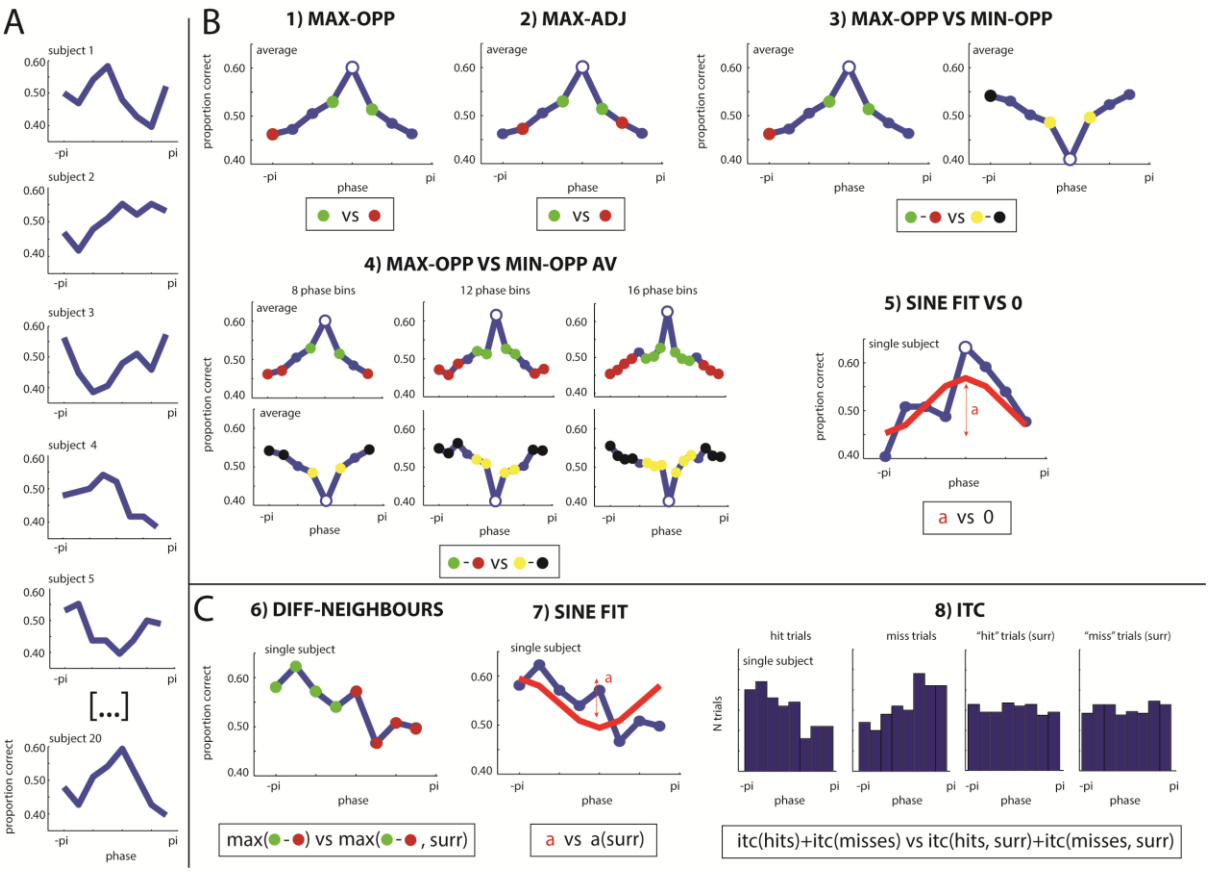

Figure 2. **A**. Six exemplary simulated subjects for a given combination of (neural and experimental) parameters. Note that the preferred phase for performance differs across participants. **B,C**. Parametric (B) and permutation-based (C) methods that were included in the present study, due to their high sensitivity or popularity in the literature. Most parametric methods compare performance across phase bins. In this case, average performance across 20 exemplary participants is shown, with individual preferred (or non-preferred, depending on the method) phases aligned at the centre bin (open circles). An exception is the method SINE FIT VS 0 (5) which involves the fit of a sine wave to individual, phase-aligned data (the corresponding panel shows such an individual fit). For permutation-based methods, alignment was not necessary as results can be compared with a surrogate distribution. For these methods, all panels show data for a single exemplary subject. All panels in B and C show example simulated performance as function of phase, apart from method ITC (8), where observed phase distributions are shown for hit and miss trials (observed data and surrogate distribution). See Statistical Methods for further details.

bioRxiv preprint doi: https://doi.org/10.1101/517243; this version posted January 11, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

2.2 Statistical methods

Phase effects (e.g., a phasic modulation of performance) can be assessed with various statistical methods that can be divided into two broad analytical categories: Parametric methods and non-parametric methods (permutation tests). We conceived and evaluated 20 methods (10 parametric and 10 non-parametric, permutation-based methods) extracting different measures of the phase effect. Each of these is described in detail 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 here as the most sensitive ("winning") method for at least one combination of parameters. These criteria led to the selection of 5 parametric and 3 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. 2B (parametric methods) and 2C (permutation-based methods).

2.2.1 Parametric methods (Fig. 2B)

This section describes the measures of phase effect that were extracted with the different parametric methods. Statistical testing of these measures was done with one-tailed repeated-measures t-tests unless otherwise stated.

1. MAX-OPP: Align maximum performance of individual participants to the centre bin, phase-wrap the remaining bins, then analyse data across participants. Since data from the centre bin is trivially maximal it cannot be used for subsequent analyses. However, performance in the phase bins adjacent to the centre bin should still be larger than that in the bin opposite to (π or 180° away from) the centre bin. Therefore, test whether average performance of the two bins adjacent to the centre bin is higher than performance of the opposite bin. This method is a slight modification from one that was applied by, e.g., Riecke et al. (2018). In Supplementary Materials, we describe a variant of this method in which data is aligned to minimum performance (MIN-OPP).

2. MAX-ADJ: Same as (1), but test whether the average performance of the two bins adjacent to the centre bin is higher than the average performance of the two bins adjacent to the bin opposite (π or 180° away) 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: Same as (1), but subtract performance in the opposite bin from the average of the two bins adjacent to the centre bin to compute difference d1. Then repeat the previous steps, but after aligning individual data based on the *minimal* performance. Again, subtract performance of the opposite bin from the average of the two bins adjacent to the centre bin, now to compute difference d2. If a phase effect were present, we would expect a positive value for d1 and a negative value for d2. However, rather than comparing each of d1 and d2 to zero (as in MAX-OPP), we test whether d1 is greater than d2.

4. MAX-OPP-AV VS MIN-OPP-AV: A variation of the MAX-OPP VS MIN-OPP method used 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) are 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) are used to calculate d2, where

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

And

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

bioRxiv preprint doi: https://doi.org/10.1101/517243; this version posted January 11, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Thus, as an example, in the case of 16 phase bins, 6 phase bins adjacent to the centre bin (3 to either side) are 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, are 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: Align performance that deviates most strongly from average performance (i.e. maximum or minimum performance) at a centre bin. Phase-wrap the remaining bins. Flip the sign (i.e. multiply by -1) if the bin used for alignment corresponds to a minimum. Fit a sine wave to the data, excluding the centre bin and restricting the phase of the fitted sine wave so that its peak corresponds to the position of the omitted centre bin. Extract the amplitude of the fitted sine wave and compare it against 0 in a group-level analysis. Note that the amplitude of the sine wave can assume negative values, corresponding to a flipped sine wave with the trough at the position of the centre bin. An amplitude of 0 is therefore our null hypothesis. A similar method was used in Zoefel et al. (2018).

2.2.2 Permutation-based methods (Fig. 2C)

Unlike the parametric methods described above, the following methods are based on a comparison of original (simulated) data and a surrogate distribution constructed by randomly reassigning trial outcomes in single participants. In all cases, the surrogate distribution was constructed by randomly assigning trial labels ("hit" or "miss") to phase bins, thereby abolishing any effect of phase bin that is present, and then applying the proposed statistical method to the permuted data. This procedure was repeated 100 times, resulting in a surrogate distribution consisting of 100 outcomes that can be 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 μ and σ are mean and standard deviation of the surrogate distribution, respectively. 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).

6. DIFF-NEIGHBOURS: For each subject, compare the average of N neighbouring phase bins with the average of the N opposite phase bins (π or 180° away). Identify the largest difference between these two groups of phase bins from all possible comparisons and average across subjects. Compare this averaged difference with a surrogate distribution.

7. SINE FIT: Fit a sine wave to the individual subject data, with phase and amplitude as free parameters (i.e. unconstrained). The frequency of the fitted sine wave corresponds to the oscillatory frequency of interest (i.e. a single cycle in a typical experiment). Sine wave fits are achieved by running a discrete Fourier Transform on simulated data (e.g., shown in Fig. 2A) and extracting the first non-DC output component as the amplitude value. Compare the average amplitude of this sine wave with a surrogate distribution. This method was used in Zoefel and VanRullen (2015) and Zoefel et al. (2018).

8. ITC: For each subject, extract the phase from hit and miss trials and calculate phase consistency across trials (inter-trial coherence, ITC) for each trial category, then sum the two ITC values:

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

bioRxiv preprint doi: https://doi.org/10.1101/517243; this version posted January 11, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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. Compare ITC_{total} with a surrogate distribution. This corresponds to the Phase Opposition Sum (POS) method 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.

2.3 Evaluation of methods

Each of the methods described in the previous section was applied to the 2 x 1000 experiments, for each parameter combination. For an experiment with an effect present, if the applied method yielded a significant effect (P-value < $\alpha = 0.05$), this was counted as a correct detection. For an experiment with effect absent, if the applied method yielded a significant effect, this was counted as a false alarm. The hit rate was computed as the number of correct detections observed, divided by the number of experiments with effect present (1000). Analogously, the false alarm rate was computed as the number of false alarms observed, divided by the number of experiments with effect absent (1000). The sensitivity of a method to detect an effect was quantified using d-prime, the standardized difference between hit rate and false alarm rate:

d' = z(hit rate) – z(false alarm rate).

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

2.4 Split-data methods for phase alignment (only in parametric methods)

In all of the parametric 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 parametric methods can be improved if this data loss is avoided. We repeated our simulations, 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 performance) 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. The selected methods were modified as described below to exploit this fact. 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).

Modifications of the methods described in Section 2.2.1:

1a. MAX-OPP: Test whether performance of the centre bin is higher than that of the opposite bin (π or 180° away from the centre bin).

2a. MAX-ADJ: Test whether performance of the centre bin is higher than that of the two bins adjacent to the bin opposite (π or 180° away) to the centre bin.

3a. MAX-OPP VS MIN-OPP: d1 and d2 are calculated by subtracting performance in the opposite bin from the centre bin.

4a. MAX-OPP VS MIN-OPP-AV: The centre bin is included in N as defined above, i.e. N = M.

5a. SINE FIT VS 0: Include the centre bin in the data used for the sine fit.

bioRxiv preprint doi: https://doi.org/10.1101/517243; this version posted January 11, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

3. Results

3.1 Effects of experimental parameters

We first pooled across all neural parameters, and revealed the method with the highest sensitivity for each combination of experimental parameters. In Fig. 3, the identity of these winning methods is shown (as a number in each cell), together with their sensitivity (color-coded). We observed that sensitivity increased monotonically with N_{trials}. Although less critical than N_{trials}, sensitivity in parametric 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*). Permutation-based methods were largely unaffected by N_{bins}. It can also be seen that permutation-based methods generally outperform parametric methods. Study design (*randomly selected* vs *imposed externally*) did not strongly affect the ability to detect phase effects. Together, these outcomes show that the highest possible sensitivity to detect phase effects can be achieved by using (1) permutation-based methods (if appropriate), (2) a relatively high N_{bins} (especially for parametric methods), and most importantly (3) a high N_{trials}.

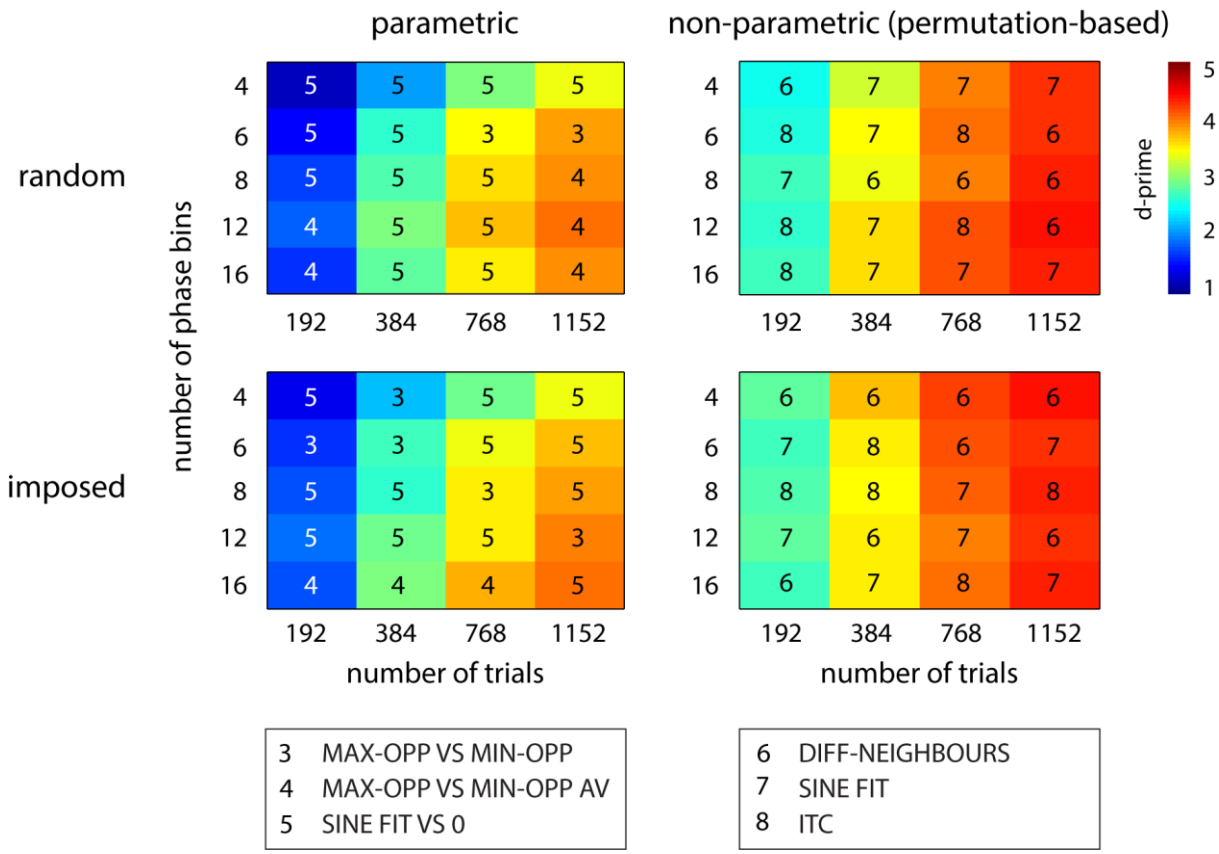

Figure 3. Highest possible sensitivity (i.e. highest d-prime among all methods tested; color-coded) and the corresponding optimal method (number in each cell) to detect phase effects, separately for all experimental parameters tested and for parametric vs permutation-based methods. Values were averaged across all neural parameters before maximal sensitivity was determined.

bioRxiv preprint doi: https://doi.org/10.1101/517243; this version posted January 11, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The identity of the optimal method to detect phase effects for specific combinations of experimental parameters is already indicated in Fig. 3. In Fig. 4, we show sensitivity separately for different methods, and illustrate how the identity of the winner method depends on experimental parameters. Results are shown only for the experimental parameters that most strongly affected the identity of the optimal parametric method (N_{bins} and *study design*). The three permutation-based methods (Fig. 2C) did not differ in their sensitivity, irrespective of the simulated (experimental or neural) parameters. For Fig. 4, sensitivity was therefore averaged across these methods. 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 the parametric method that performs well in all scenarios. For some combinations of experimental parameters, MAX-OPP VS MIN-OPP (for lower N_{bins}) or MAX-OPP-AV VS MIN-OPP-AV (for higher N_{bins}) can be superior (these exceptions can be extracted from Fig. 3). The superiority of permutation-based methods (black lines) is again visible, in particular for medium *effect sizes* and lower N_{bins}.

3.2 Effects of neural parameters

Next, we investigated the effect of neural parameters and their interaction with experimental parameters. In Fig. 5, we again show sensitivities of winning methods and their identities, but for different combinations of neural parameters. As already observed in Fig. 4, sensitivity to detect phase effects was strongly modulated by *effect size*. Sensitivity was highest for *symmetric* shapes, but this effect was more pronounced for parametric methods. Finally, sensitivity was modulated by *total width*, with highest sensitivity for narrow effects. 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 high peak amplitudes are more important to detect phase effects than broad effects.

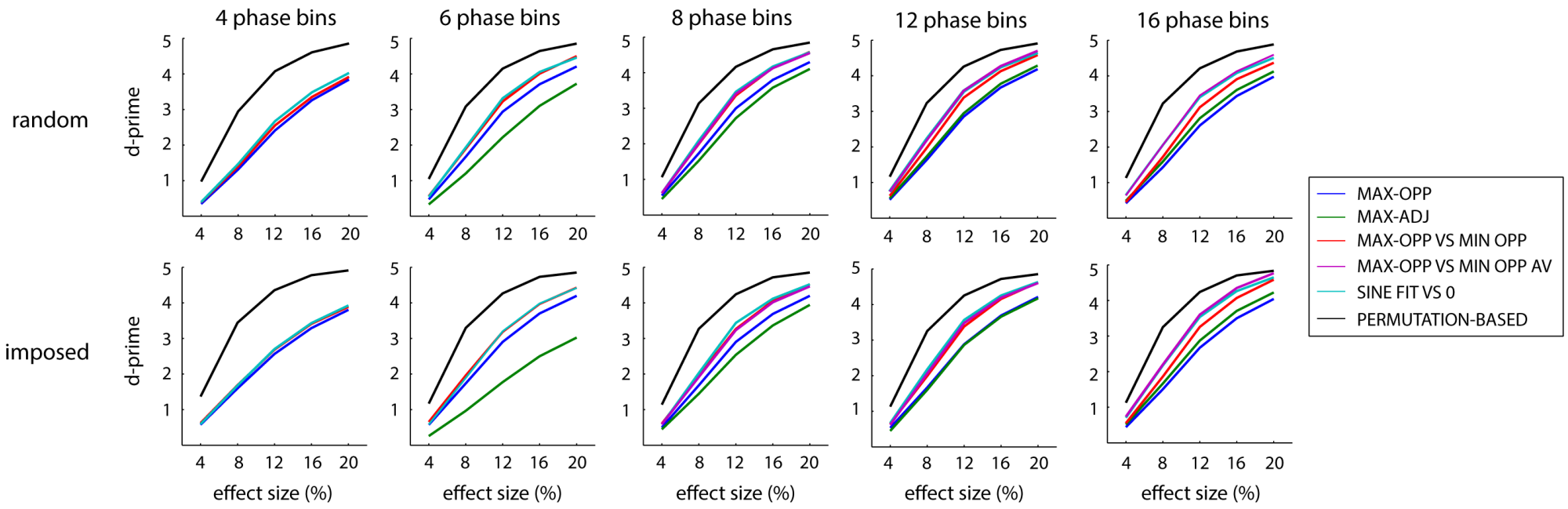

Figure 4. Sensitivity of different methods to detect phase effects, for various combinations of N_{bins}, *study design*, and *effect size*. *Effect size* corresponds to a peak-to-peak modulation of performance for a sinusoidal shape (100% *total width*, *asymmetry* 0). Note that MAX-ADJ is not defined for 4 phase bins, and MAX-OPP VS MIN OPP AV is not defined for 4 and 6 phase bins. The black line shows the average sensitivity across the three selected permutation-based methods (DIFF-NEIGHBOURS, SINE FIT, ITC), as these performed virtually identically.

bioRxiv preprint doi: https://doi.org/10.1101/517243; this version posted January 11, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

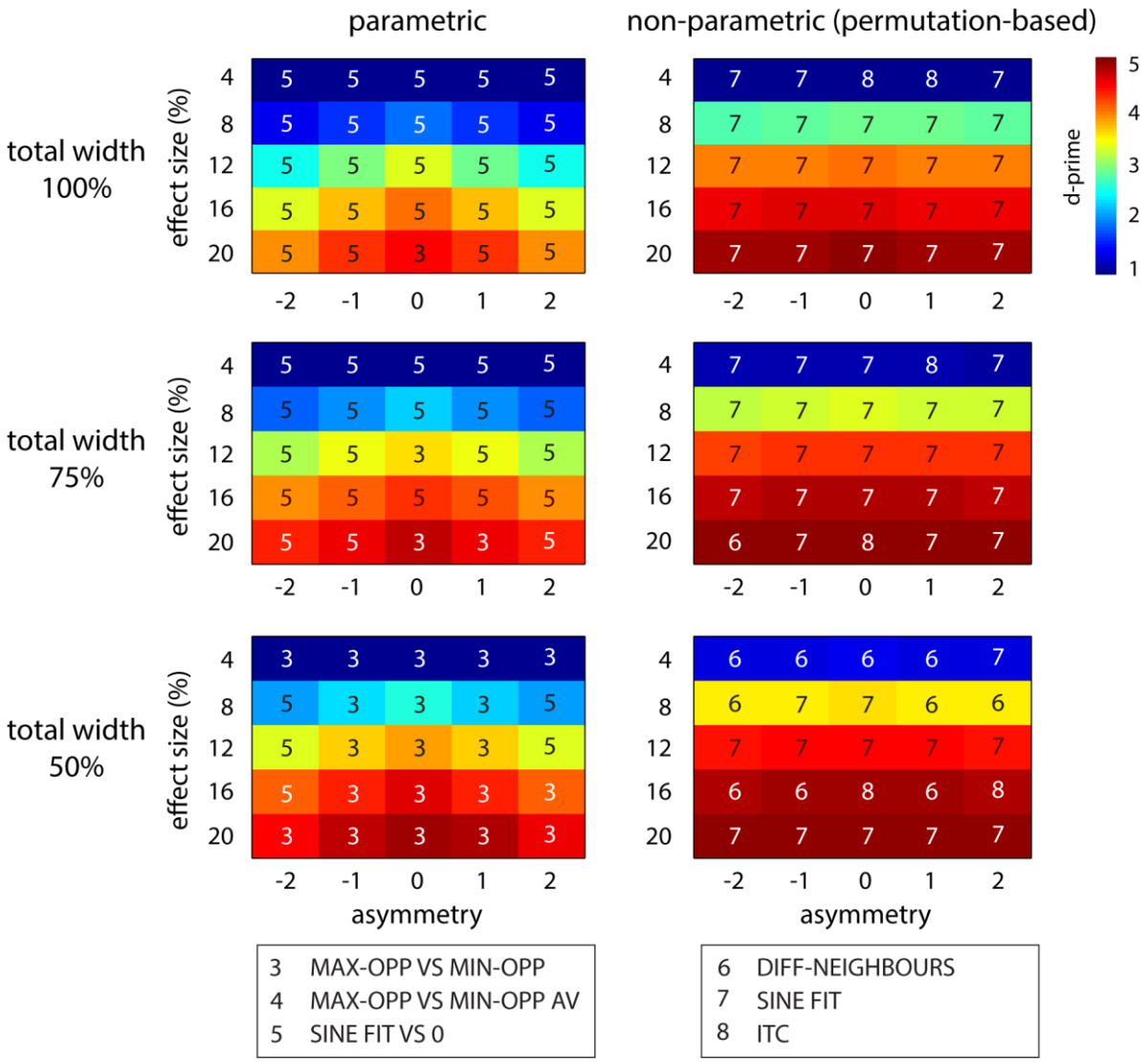

Figure 5. Highest possible sensitivity and corresponding optimal method to detect phase effects, separately for all neural parameters tested and for parametric vs permutation-based methods.

Fig. 6 again shows sensitivity separately for the different parametric methods, and for the average of the three permutation-based methods. In addition to *effect size*, results are shown for combinations of the two parameters that most strongly affected the identity of the optimal method: N_{bins} and *total width*. Sensitivities were averaged across relatively broad effects (75% and 100%) as results were very similar.

Again, we observed that SINE FIT VS 0 is the parametric method that performs well in all scenarios. A noticeable difference are narrow effects (50% *total width*) for which MAX-OPP VS MIN-OPP is the optimal method (in particular for 6 and 8 phase bins). For broader effects, the sensitivity of MAX-OPP-AV VS MIN-OPP-AV seems to increase with N_{bins}, eventually becoming the winning method for 12 or more phase bins. Permutation-based methods again outperformed parametric methods for all combinations of parameters.

In addition to results described here, we provide MATLAB code which determines, for a given combination of neural and experimental parameters, the optimal method (from our pool of methods) to detect phase effects, along with the corresponding sensitivity (<link will be provided upon acceptance>).

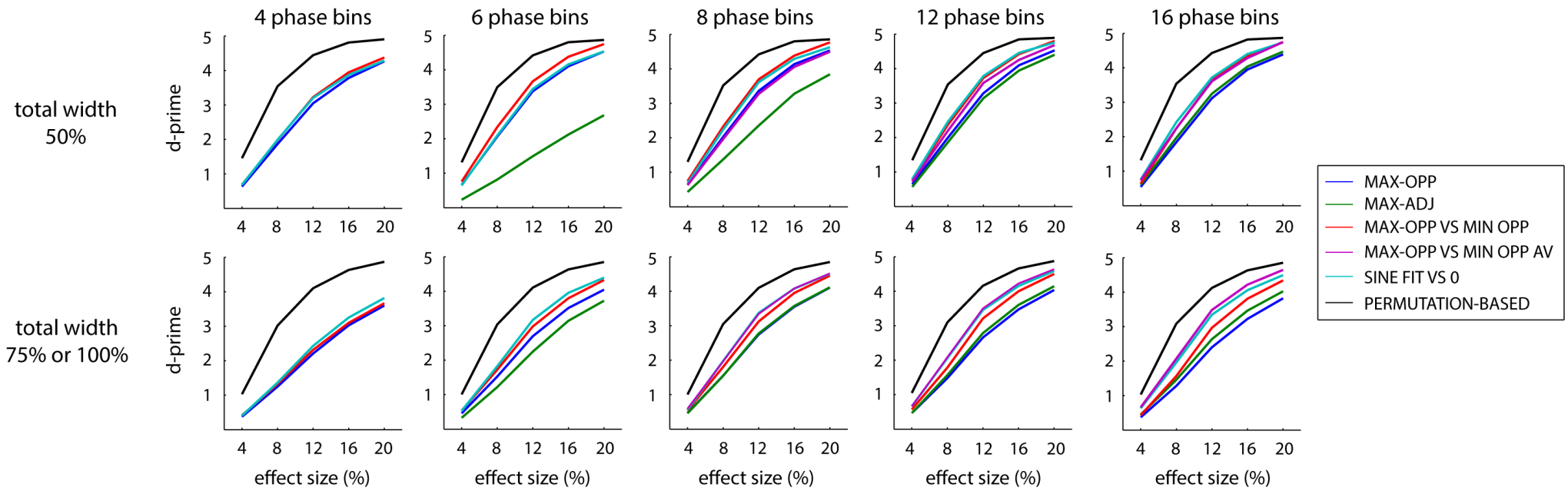

Figure 6. Sensitivity of different methods to detect phase effects, for combinations of N_{bins}, *total width*, and *effect size*. For conventions, see caption of Fig. 4.

bioRxiv preprint doi: https://doi.org/10.1101/517243; this version posted January 11, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

3.3 Effectiveness of split-data methods for phase alignment

We also tested whether the sensitivity of parametric methods can be improved if individual preferred phases are estimated 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. 7 compares the ability of these split-data methods to detect phase effects ("split-data") with other approaches, including the standard ones described here ("full dataset" for parametric methods, "permutations" for permutation-based methods). We found that the split-data approach did not improve the sensitivity of parametric 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}.

It is theoretically possible that the preferred phase of individual participants is 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 data alignment without the necessity of excluding phase bins in the dataset used to test for phase effects, sensitivity was improved as compared to our standard parametric approach. However, sensitivity was still lower than that observed for permutation-based methods.

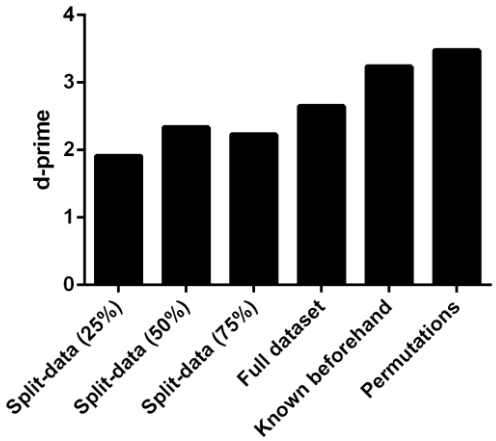

Figure 7. Highest possible sensitivity (i.e. highest d-prime among all methods) for split-data methods and other approaches. Results are shown for simulated experiments of 384 trials each (using other N_{trials} did not change results). Sensitivity was averaged across all other parameter combinations 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 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. "Permutations": Equivalent sensitivity for permutation-based methods (i.e. standard approach). Note that results for "full dataset" and "permutations" correspond to the sensitivities shown for parametric and permutation-based methods in Fig. 3, respectively, averaged across N_{bins} and *study design*, and for 384 trials.

bioRxiv preprint doi: https://doi.org/10.1101/517243; this version posted January 11, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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 measures of performance (from the detection of simple stimuli such as pure tones, to more complex tasks such as speech comprehension) vary with the phase of ongoing neural oscillations (e.g., Baumgarten et al., 2015; Busch et al., 2009; Henry et al., 2014; Henry and Obleser, 2012; Ng et al., 2012; Riecke et al., 2018, 2015a, 2015b; Strauß et al., 2015; Zoefel et al., 2018; 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 neural phase on behaviour (with the exception of a comprehensive evaluation of methods involving the comparison of two phase distributions, in VanRullen, 2016), and studies have used an arbitrary selection of different analysis procedures.

Parametric vs Non-Parametric Statistical Approaches

Here, we show that many of the most popular parametric approaches to analysis of this sort of data are not necessarily the best. In all our simulations, we found optimal sensitivity for non-parametric, permutation-based methods in which the observed data is compared with a surrogate distribution. One of the most popular 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. Another method that we found to perform well used the amplitude of a sine wave fitted to the data as a proxy for the magnitude of phase effects (SINE FIT; used in Zoefel et al., 2018; Zoefel and VanRullen, 2015), whereas a third good method was novel and involved systematic comparison of performance between groups of neighbouring phase bins (DIFF-NEIGHBOURS).

Non-parametric methods not only perform better than parametric methods, they also seem to perform in a relatively stable manner across almost all the neural parameters tested here. Hence, the use of these methods does not require researchers to make a-priori assumptions about the specific nature of the phase effect. On the other hand, non-parametric methods are often more computationally demanding than simpler, parametric analysis methods. The outcome of non-parametric tests can also be relatively abstract (e.g., a single value reflecting the degree of phasic modulation of performance) and they may not so readily allow the visualization of performance as a function of oscillatory phase. Perhaps most critically, randomly re-assigning trials to conditions (as required for the construction of a surrogate distribution) might be problematic where it violates counterbalancing of stimulus materials. Shuffled data can include additional, item-wise variation (which is controlled in the original counter-balanced experimental data) and thereby lead to invalid results. Thus, although we recommend the use of permutation-based methods to analyse phase effects, it needs to be ensured that the fundamental assumption of permutation methods – that trials can be randomly exchanged between conditions – is met for the particular experiment that is being designed.

Perhaps for these reasons, then, many studies analyse their data by aligning the maximum[1] of individual data to a common phase bin, and using repeated measures parametric statistical tests to compare performance in the remaining bins. This can be done by comparing average performance in bins

[1] Note that, in principle, using the minimum (instead of the maximum) for alignment is an equally valid approach. Due to our assumption that a phase effect always consists of two peaks (a maximum and a minimum), methods based on minimum-alignment (e.g., MIN-OPP) performed similarly to equivalent methods based on maximum-alignment (e.g., MAX-OPP). An exception is for asymmetric neural profiles: MAX-OPP was often superior to MIN-OPP if the peak amplitude of the positive deflection was larger than that of the negative one, and vice versa.

bioRxiv preprint doi: https://doi.org/10.1101/517243; this version posted January 11, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

adjacent to the bin used for alignment with that in the opposite bin or in those adjacent to it, using t-tests[2] (MAX-OPP or MAX-ADJ; used in this or similar form in Riecke et al., 2015b, 2015a).

In our simulations, we found that a more sophisticated alignment procedure (MAX-OPP VS MIN-OPP, and its variant MAX-OPP VS MIN-OPP AV; neither of which have been previously used in the literature) as well as a method involving the fitting of a sine wave (SINE FIT VS 0; used in Zoefel et al., 2018) can improve sensitivity to detect phase effects relative to the other parametric methods tested. These methods involve fewer statistical tests, require less computational effort and can offer more intuitive results. However, their ability to detect phase effects was not only reduced as compared to permutation-based tests, but establishing which of these three alternative methods is optimal can be hard to establish in advance. Our simulations show that the choice of optimal method depends on the nature of the underlying phase effect (i.e. neural parameters) and the researcher's methodological decisions (i.e. experimental parameters). While the latter are easy to control (within limits of participant time and tolerance), the former depends on neural and anatomical variables that at this time cannot be specified in advance.

In the future, we might be able to estimate the shape of the "true" phase effect by combining data obtained in many independent studies investigating a similar question with comparable methods. Nevertheless, the shape of neural oscillations seems to depend on many different factors, such as the brain regions involved in a task or the presence of certain neurotransmitters (Cole and Voytek, 2017). It is therefore likely that these factors also influence the shape of the phase effect, and that this shape deviates from the often assumed perfect sinusoidal shape. As long as neural parameters remain unknown, we recommend choosing a parametric method that performs well for all possible scenarios, i.e. the optimal method identified when averaging across all neural parameters (illustrated in Figs. 3 and 4). Finally, we note that the recommendations we give here can only be considered valid for the different parameters and parameter ranges tested, and assumptions made. For instance, it seems logical that enhanced performance at certain phase bins (relative to average performance) has a counterpart in phase bins leading to impaired performance (VanRullen and McLelland, 2013). Some neural parameters (e.g., asymmetry) were inspired by the notion that neural excitability (and therefore also the shape of neural oscillations) directly translates into behaviour (Buzsáki and Draguhn, 2004; Lakatos et al., 2005; Peelle and Davis, 2012), but there is thus-far only limited experimental evidence for this assumption. In particular, the interaction between neural processes and externally imposed current (e.g., tACS) might lead to other, potentially more complex, effect shapes.

Optimising the detection of phasic modulation or enhancement/impairment of behaviour

The methods presented here serve to detect a *modulation* of performance or behaviour by the phase of neural oscillations. If present, this modulation should reflect both an enhancement and impairment of performance relative to the average across all phases (VanRullen and McLelland, 2013). However, if neural oscillations are perturbed either by rhythmic sensory or electrical stimulation, it can be interesting to determine whether behaviour is enhanced, impaired, or both (but at different phases), as compared to an unstimulated, or sham condition (Riecke and Zoefel, 2018). Future studies could explore optimal statistical methods for this purpose. This is important for potential practical applications of non-invasive brain stimulation that seek to enhance or impair behaviour rather than to produce both enhancement and suppression.

---

[2] Other studies simply compare performance across all phase bins (apart from that used for alignment) using ANOVAs (e.g., Baumgarten et al., 2015; Busch et al., 2009; Ng et al., 2012; Zoefel and Heil, 2013). This approach is 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 MAX-OPP and MAX-ADJ, 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 for the majority of the methods tested (data not shown).

bioRxiv preprint doi: https://doi.org/10.1101/517243; this version posted January 11, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

We found that the most crucial experimental parameter for the overall ability to detect phase effects (i.e. without necessarily influencing the identity of the best method) is the number of trials, not the number of phase bins. Given this, it seems advisable to prefer simpler experimental designs with a relatively high number of trials. 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 linear 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 the reported effect sizes are consistently small (Gundlach et al., 2016; Neuling et al., 2012; Riecke et al., 2018, 2015a, 2015b; Riecke and Zoefel, 2018; Wilsch et al., 2018; Zoefel et al., 2018). Transcranial brain stimulation has recently been exposed to pronounced criticism, challenging its efficacy to manipulate neural responses or behaviour (Lafon et al., 2017; Vöröslakos et al., 2018; but see 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 simulation evidence like that presented here.

Individualised stimulation protocols will also need to address the issue of different preferred phases across participants (e.g., Busch et al., 2009; Riecke et al., 2018). The exclusion of the aligned phase bin in parametric methods might explain the observed superiority of permutation-based methods which do not require excluding any data. This could be taken to mean that parametric methods can be further improved if we avoid this data exclusion. However, we found that dedicating a subset of the collected data to estimating preferred phases is not helpful compared to excluding data from the preferred phase in a single analysis (Fig. 7). Nevertheless, we also show that sensitivity of parametric methods can be enhanced if preferred phases for individual participants are known beforehand (i.e. if no trials have to be sacrificed for this purpose). These phases could, for instance, have been determined in similar previous studies in which these or other participants have already participated. Of course, a crucial question that follows from this possibility is whether preferred phases are stable across individuals or studies. The answer to this question is currently unknown; thus, future studies are necessary, and considering the anatomical and functional determinants of preferred phases for behaviour remains an important topic for future investigation.

The pre-registration of experimental paradigms and analyses is rapidly becoming an essential step in scientific research (Nosek et al., 2018). However, in particular the definition of detailed analysis plans is difficult in a young field of research lacking established routine procedures (Ledgerwood, 2018). The investigation of oscillatory effects of neural phase on behaviour is one such field in which a proliferation of analysis methods has challenged researchers seeking to pre-register their analysis protocols. We hope that our study provides important groundwork for pre-registration of tACS, EEG/MEG and other studies exploring the effect of neural phase on behaviour.

bioRxiv preprint doi: https://doi.org/10.1101/517243; this version posted January 11, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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

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., 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

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

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.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

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

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

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

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

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

Jones, S.R., 2016. When brain rhythms aren't 'rhythmic': implication for their mechanisms and meaning. Curr. Opin. Neurobiol. 40, 72–80. https://doi.org/10.1016/j.conb.2016.06.010

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

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

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., 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. https://doi.org/10.1126/science.1154735

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

bioRxiv preprint doi: https://doi.org/10.1101/517243; this version posted January 11, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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

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

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

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

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

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

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

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

bioRxiv preprint doi: https://doi.org/10.1101/517243; this version posted January 11, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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., 2018. 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., 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