Brain Stimulation 12 (2019) 1001–1009

Contents lists available at ScienceDirect

Brain Stimulation

journal homepage: http://www.journals.elsevier.com/brain-stimulation

Analytical bias accounts for some of the reported effects of tACS on auditory perception

Boateng Asamoah, Ahmad Khatoun, Myles Mc Laughlin*

Exp ORL, Department of Neurosciences, The Leuven Brain Institute, KU Leuven, B-3000, Leuven, Belgium

ARTICLE INFO

Article history:
Received 20 July 2018
Received in revised form
8 March 2019
Accepted 12 March 2019
Available online 15 March 2019

Keywords:
Analytical bias
tACS
Auditory perception

ABSTRACT

Background: Transcranial alternating current stimulation (tACS) has been shown to modulate auditory, visual, cognitive and motor function. However, tACS effects can often be small and difficult to reproduce. Thus, the establishment of robust experimental and analysis procedures is of high importance. We reviewed the analysis used in six studies that investigated if tACS can phase-modulate auditory perception. All studies used analytical methods that introduce bias and could produce false positive results. Four studies corrected for this bias but two did not.

Objective: Our objectives were two-fold: 1) Use simulated null hypothesis datasets, where no tACS effect is present, to determine if uncorrected analytical bias could account for some of the reported effects on auditory perception. 2) Help establish best practices to correct for bias when analyzing tACS phase-effects on perception.

Methods: We simulated null hypothesis datasets (i.e. no tACS effect) by drawing samples for all tACS and sham conditions from the same normal distribution. We then applied the reported analyses to the null hypothesis datasets.

Results: Reported results from studies that did not correct for analytical bias could be reproduced from the null hypothesis datasets. However, results for studies that did correct for analytical bias could not be reproduced from the null hypothesis datasets.

Conclusion: True effects of tACS on auditory perception can be detected if analytical bias is accounted for by using correction procedures. However, to fully establish the effects of tACS on auditory perception a reanalysis of the data for the studies that used biased analysis without correction procedures is needed.

Introduction

Transcranial alternating current stimulation (tACS) is a non-invasive neuromodulation method in which alternating current is passed through scalp electrodes to entrain neural activity in underlying cortices. tACS modulates auditory [1], visual [2,3], memory [4], cognition [5,6] and motor [7,8] function. It is assumed that these behavioral effects result from the weak alternating current directly causing neural entrainment in the associated cortex. However, reported tACS effects are controversial. Recent studies show that electric field strengths generated in the brain may be too weak to cause neural entrainment [9,10], and reproducibility of reported effects is a concern [11,12]. While other studies show that tACS motor system effects may be caused by stimulation of peripheral nerves in the scalp and not by transcranial stimulation of neurons in the cortex [13]. Thus, it is now critical to determine the most robust and repeatable tACS effects, and establish best practices for analysis of tACS datasets.

With this in mind, we reviewed the literature investigating the phase-effects of low-frequency tACS on auditory perception. To observe maximum auditory perception modulation, tACS must be synchronized to the cortical response from the auditory stimulus. However, two potential sources of variability mean that the delay between tACS and the cortical response may differ between individuals. Auditory system delays arise from a combination of mechanical delays in the middle and inner ear and neural or processing delays caused by transmission from the auditory nerve, through the brainstem and thalamus, to the cortex. EEG N100 literature indicates that individual variation in these auditory system delays is on the order of a few 10s of milliseconds [14]. A

* Corresponding author.
E-mail addresses: boateng.asamoah@kuleuven.be (B. Asamoah), ahmad.khatoun@kuleuven.be (A. Khatoun), myles.mclaughlin@kuleuven.be (M. Mc Laughlin).

https://doi.org/10.1016/j.brs.2019.03.011
1935-861X/

second potential source of delay variation arises from the interaction of tACS generated electric field with the cortical anatomy. Anatomical differences between subjects could mean that orientation of the cortical neurons being stimulated, relative to the tACS generated electric field, differs between individuals. This could theoretically delay any neuromodulation effect by up to half a cycle of the tACS frequency.

Thus, the delay (relative to the auditory stimulus) at which tACS could cause maximum modulation of auditory perception is expected to differ between individuals. This delay variance means that when data from different individuals is averaged together (to examine group statistics) any tACS phase effects would average out to zero. Therefore, to account for individual delay differences it is necessary to use a phase or delay alignment procedure. However, this alignment procedure can introduce analytical bias. Specifically, aligning each individual subject's best performing phase condition to exactly the same phase at the group level, will cause the natural variability within the dataset to be skewed at the phase condition chosen for alignment. This natural variability (or noise distribution) will not be skewed at the other non-aligned phase conditions. Only skewing the noise distribution at the aligned phase could theoretically cause a tACS effect to be detected, even if no effect were present.

We identified six studies that used an alignment procedure to investigate phase-effects of tACS on auditory perception [15–20]. All studies presented subjects with an auditory stimulus while applying tACS to the auditory cortex. They all systematically changed the phase or delay between the tACS and auditory stimulus and used a behavioral task to measure the subject's perception of the stimulus for each phase or delay condition. They also measured auditory perception on the same task during a sham condition (i.e. no tACS). All studies reasoned that, because of variance in delays caused by individual anatomical differences, it was necessary to perform a phase or delay alignment procedure before performing a group analysis. The alignment procedures differed slightly but are conceptually similar: Wilsch et al. [15] selected the best performing delay for each subject and compared this to sham. Note, Wilsch et al. [15] did not apply standard sinusoidal tACS but instead used an oscillation extracted from the speech stimulus envelope. Neuling et al. [16] detected a zero crossing in each subject's data and aligned this to 180° phase. The three Riecke et al. studies [17–19] selected the best condition in each subject's data and aligned this to either 0° or 90°. While Zoefel et al. [20] aligned maximum performance at 0°. These alignment procedures introduce analytical bias which, if not corrected for, increases the chance of detecting a false positive – i.e. detecting a tACS phase effect when in fact none is present.

Four of the six studies [17–20] corrected for this analytical bias. After bias correction, one study [20] did not find any effect of tACS on auditory perception, although they did find an effect of tACS on the functional magnetic resonance imaging (fMRI) derived blood oxygen level dependent (BOLD) signal. The other three studies found a small but significant effect of tACS on auditory perception i.e. performance in the behavioral task during one tACS phase was significantly different than performance during the opposite phase [17,18] or tACS performance during one phase was significantly different than sham [19]. The two remaining studies [15,16] did not correct for analytical bias when analyzing data from the behavioral task and reported highly significant tACS phase effects. We hypothesized that the uncorrected analytical bias introduced by the alignment procedure caused these studies to detect a highly significant tACS phase effect, which in fact may not be present. To test this we simulated datasets with no tACS effect (i.e. a null hypothesis dataset) for the five studies reporting a phase effect of tACS on auditory perception [15–19]. We applied the reported analysis to the null hypothesis datasets to test if we would find false positive effects. We found that, due to the analytical bias, similar results as in Refs. [15,16] could be obtained as a false positive. However, we also found that the null-hypothesis datasets could not account for the tACS phase effects reported in studies that did correct for analytical bias. Our analysis supports the notion that, under certain experimental conditions, tACS can cause small but significant phase-effects on auditory perception. However, to reveal these effects analytical bias must be carefully accounted for.

Methods

General approach

Our approach was as follows: 1) Simulate a dataset from the study containing the phase or delay conditions (tACS applied) and the sham or control conditions (no tACS) by randomly drawing samples from a normal distribution. 2) Apply the phase or delay alignment procedure described in each study. 3) Apply the reported analysis and statistical tests to the simulated dataset. 4) Use the simulated dataset to generate the main figures in each study.

Problem illustration

To illustrate how a phase or delay alignment procedure introduces analytical bias we first applied our methodological approach to two generic simulated datasets – a null hypothesis dataset and an alternative (true effect) hypothesis dataset. To simulate a null hypothesis dataset, data for all conditions and subjects were drawn from the same normal distribution. To simulate a true-effect dataset, data for one delay condition was drawn from a normal distribution with a mean different to that of the null hypothesis distribution (Fig. 1, left column). Importantly, to simulate each subject having a different auditory system delay, this data was assigned to a different condition for each subject. The rest of the conditions in the true-effect dataset were then drawn from the same null hypothesis distribution. Note, for clarity of illustration we adopted the simplest approach of showing a tACS true-effect in just one phase bin. A cyclic tACS true-effect could be simulated by drawing each phase condition from a different distribution, but is not necessary to illustrate the problem. We then applied the same generic alignment procedure and statistical analysis to both the null hypothesis and the true-effect datasets.

Assume null hypothesis for each study

For each study, we assumed only the null hypothesis and created one normal distribution using the reported mean and a calculated standard deviation, extracted directly from the relevant figures in each paper. The standard deviation was calculated by multiplying the reported standard error by the square root of the number of subjects. This normal distribution was then used to simulate null hypothesis datasets, to which we applied the alignment procedures and analysis as reported in the study.

If we apply the analysis method as reported in one particular study to a null hypothesis dataset and it produces the same tACS phase effects as reported in that study, we would conclude that no population effect is required to reproduce a similar result. This would indicate that the reported results maybe a false positive. We could then no longer accept the study's conclusions. Alternatively, if we apply the reported analysis to the null hypothesis datasets and cannot produce the reported tACS phase effects, we would conclude that the analysis is not biased towards finding an effect and we can accept the study's conclusions.

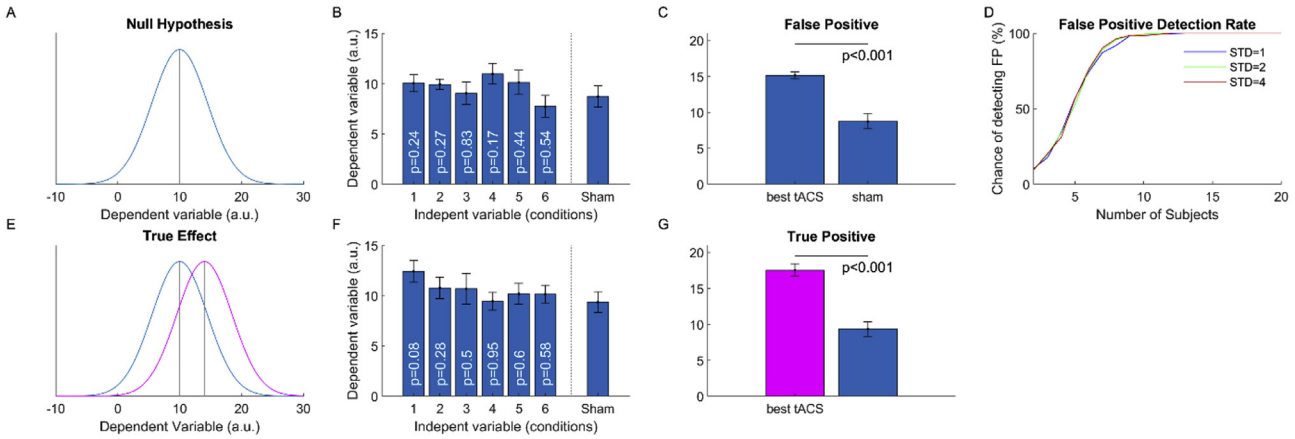

Fig. 1. Problem illustration – analytical bias leads to false positive results. **A**. The null hypothesis – data for all experimental conditions come from the same normal distribution. **B**. For each subject (n = 20), data (dependent variable) for each experimental condition (independent variable) and the sham condition (control) were drawn from the same normal distribution. This is a null hypothesis data set – i.e. there is no effect. Bars show mean and error bars show standard error. **C**. A phase or delay alignment procedure was simulated by selecting the best condition for each subject and putting this into a new condition called 'best tACS'. A paired *t*-test showed a significant difference between the 'best tACS' and the sham condition. The phase or delay alignment procedure introduces bias into the analysis which leads to this false positive result. **D**. The simulation was repeated 500 times for a range of subject numbers and standard deviations (blue, green and red). Increasing the subject number increases the chance of detecting a false positive (FP). **E**. The distribution for a true effect or alternative hypothesis is shown (purple). The mean is different than the null hypothesis (blue). **F**. A true effect dataset is simulated by drawing data for each subject for one randomly chosen condition from the purple distribution; all other conditions and the sham condition are drawn from the blue distribution. **G**. The phase or delay alignment procedure is repeated but this time a true effect is detected. The conceptual problem with this analytical approach is that it is not possible to distinguish between the null and the alternative hypotheses. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Dataset simulation

All data were simulated by randomly drawing samples from a normal distribution. This was done using the *normrnd* function in MATLAB (Mathworks, Natwick, MA). The dataset parameters were matched to those reported in each study. The parameters included: number of subjects, number of phases or delay tACS conditions, number of control or sham conditions, distribution mean and distribution standard deviation.

Implementation of analysis used in studies

We implemented the analysis described in each of the five studies [15–19] as closely as possible. Full descriptions of the analysis can be found in each study. When a part of the analysis description was not clear or we deviated slightly from the described analysis, we state this and give an explanation at the relevant Results section. All analyses were performed in MATLAB.

All the MATLAB scripts to generate the simulated datasets, perform the analysis and make the figures are included in the Supplementary Information Section.

Results

Problem illustration: analytical bias

Fig. 1 provides a generic illustration of analytical bias. The top row shows the null hypothesis and the bottom row shows the alternative hypothesis (or true-effect). Unbiased, analysis and statistical testing should distinguish between the null and alternative hypotheses. If an analysis increases the chances of detecting a false positive (i.e. detecting an effect when none is present), we say that it is biased.

The null hypothesis dataset shown in Fig. 1B was simulated by drawing samples for each condition (1–6, plus sham) from the same normal distribution (Fig. 1A). Thus, no effect is present in the dataset. We now apply a phase or delay alignment analysis procedure. To do this, we select the largest value in conditions 1 through 6 for each subject and add this to a new condition called 'best tACS'. We then use a paired *t*-test to compare the 'best tACS' condition to the sham condition. Selecting the best condition for each subject introduces analytical bias. In this case, the bias is enough to cause detection of a false positive. This analysis would cause us to incorrectly conclude that tACS had an effect, when in fact it had none. To see how often this analysis detects false positives we ran the simulation 500 times for a number of subjects ranging from 2 to 20 and for three distributions with different standard deviations (1, 2, and 4). Increasing the number of subjects increases the chances of detecting a false positive (Fig. 1D). For datasets with more than 10 subjects, there was a 100% chance of detecting a false positive. Changing standard deviation had a negligible effect.

The true-effect dataset is shown on the bottom row. In this case, for each subject data for one randomly selected condition was drawn from the purple distribution (Fig. 1E). All other conditions were drawn from the blue distribution. Applying the same analysis and statistical tests as above, shows that we can detect a true effect (Fig. 1G). The conceptual problem with this analytical approach is that it cannot distinguish between the null hypothesis and a true effect. Therefore, we cannot draw meaningful conclusions from data analyzed in this way.

Wilsch at al 2018 under null hypothesis

Assuming the null hypothesis (i.e. all data points randomly drawn from the same distribution) we now simulated a dataset for Wilsch et al. [15] containing 6 tACS delay conditions and three control conditions (+tDCS, -tDCS and sham). The simulation used a distribution with a mean of −7 and standard deviation of 0.87, approximately matching Wilsch et al. Fig. 2A shows the simulated null hypothesis dataset. We applied the reported delay alignment procedure by selecting the best delay condition for each subject (lowest speech comprehension threshold (SCT)) and adding this to a new condition called 'best tACS'. We compared performance in the best tACS condition to the sham condition using a *t*-test (we used paired two sided test but this is not specified in the study) and

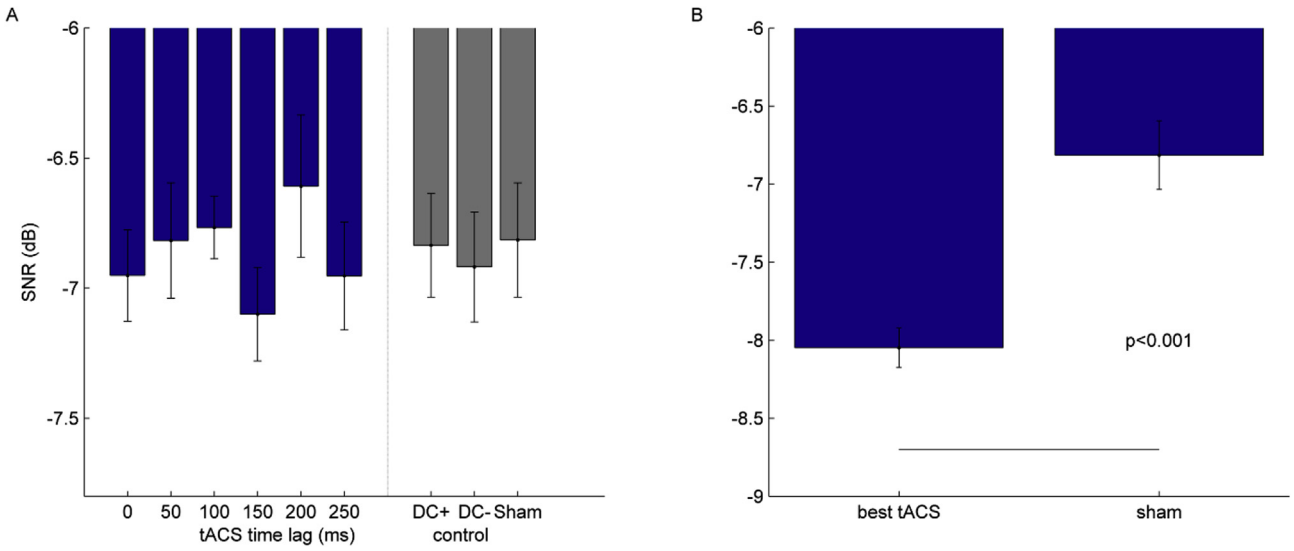

Fig. 2. Applying the Wilsch et al. analysis to a null hypothesis dataset consistently gives false positives. **A**. A null hypothesis dataset was simulated for all the tACS delay conditions and control conditions (sham, anodal and cathodal tDCS) by drawing samples from the same normal distribution. **B**. After application of the described delay alignment procedure a significant difference between the 'best tACS' and sham condition is detected. In the case of our simulated dataset, this is a false positive since no tACS phase effect was present in the simulated null hypothesis dataset. Bars show means and error bars show standard errors.

found that the best tACS condition was significantly greater than sham (Fig. 2B). With the simulated data we know this is a false positive result. Comparing our Fig. 2 with the equivalent figure in Ref. [15] (also Fig. 2) shows very similar effects. We repeated the null hypothesis simulation 100 times and always obtained the same false positive result. Thus, we cannot conclude from the data analysis presented in Wilsch et al. that tACS with speech envelopes modulates speech comprehension.

To support the conclusion that tACS modulates speech comprehension, Wilsch et al. used two separate analysis procedures to show that a derived dependent variable (ΔSCT) was sinusoidally modulated by tACS. ΔSCT was calculated by subtracting sham SCT performance from the SCT for each delay condition for each subject. In the first procedure, they fit a sinusoidal model to the ΔSCT-delay functions for each subject. They also fit the same data with a linear and quadratic function and compare the fits based on the Bayesian Information Criteria (BIC). They found that the data were better fit with a sinusoidal function than with a linear or quadratic function (Fig. 3 in Ref. [15]). They also report that the average frequency of the fitted sinewave was 5.12 Hz. We derived ΔSCT from our simulated dataset and applied the same three curve-fitting methods. Wilsch et al. do not provide the BIC equation. We calculated BIC based on the residual (R) of the fits, using the formula

$$BIC = n \ln\frac{R}{n} + k \ln(n)$$

where n is the number of parameters in the fit and k is the number of conditions [21]. We found that null hypothesis data were also better fit with a sinusoidal function, than with a linear or quadratic function (Fig. 3). We also found that the fitted sinewave had an average frequency of around 5 Hz. We repeated our simulation 100 times and always obtained a similar result (5.1 ± 0.4 Hz). In our case, this result is because noise is better described by a sinusoidal, than a linear or quadratic function. The 5 Hz frequency is a consequence of having 6 conditions each separated by 50 ms.

In the second procedure, Wilsch et al. analyzed the power spectra of the ΔSCT-delay functions. They first up-sampled each ΔSCT-delay function by a factor of two, appended 320 zeros and then computed power spectra. They found that the average power spectra peaks at 5 Hz and that the peak is larger than a noise-floor spectrum computed by permuting individual ΔSCT-delay functions (Fig. 4A and B in Ref. [15]). They suggest that this supports the idea that ΔSCT-delay functions are sinusoidally modulated by tACS at a frequency of 5 Hz. We applied the same analysis to our null hypothesis dataset and found that the average power spectra of the up-sampled data also peaked at 5 Hz (Fig. 4B). In our dataset the peak at 5 Hz is caused by noise spread across 6 delay conditions separated by 50 ms. Changing the number of conditions, or delay between them, shifts the spectrum peak. Wilsch et al. did not specify if they computed the noise-floor spectrum by permuting the original or the up-sampled ΔSCT-delay functions. Fig. 4B (dark grey lines) shows the noise-floor computed by permuting the up-sampled data. This noise-floor is below the mean power spectra (blue line) and similar to that shown in Wilsch et al. The correct way to calculate the noise-floor is to permute the original data. We did this using our null hypothesis dataset (light grey line). The power spectra is mostly within the correct noise-floor which indicates, as expected from a null hypothesis dataset, that the peak at 5 Hz is not significant. The slight rise in the blue line at lower frequencies is caused by subtraction of the mean from the permuted data. We repeated this 100 times and always obtained quantitatively similar results.

Neuling at al 2012 under null hypothesis

In Neuling et al. [16] the dependent variable was detection threshold for a tone in noise. They calculated a normalized (detection) threshold for each tACS phase by subtracting a subject's performance during sham from their performance during that particular tACS phase condition. Next they applied a phase alignment procedure by detecting the zero crossing in each subjects threshold-phase function and aligning this to 180°. As analysis of the null hypothesis dataset will demonstrate, this alignment procedure introduces a very similar analytical bias.

Again, we assumed the null hypothesis and simulated a dataset for the Neuling et al. study containing 6 tACS phase conditions and one control condition (sham). The simulation used one distribution with mean 4 and standard deviation 0.11, approximately matching

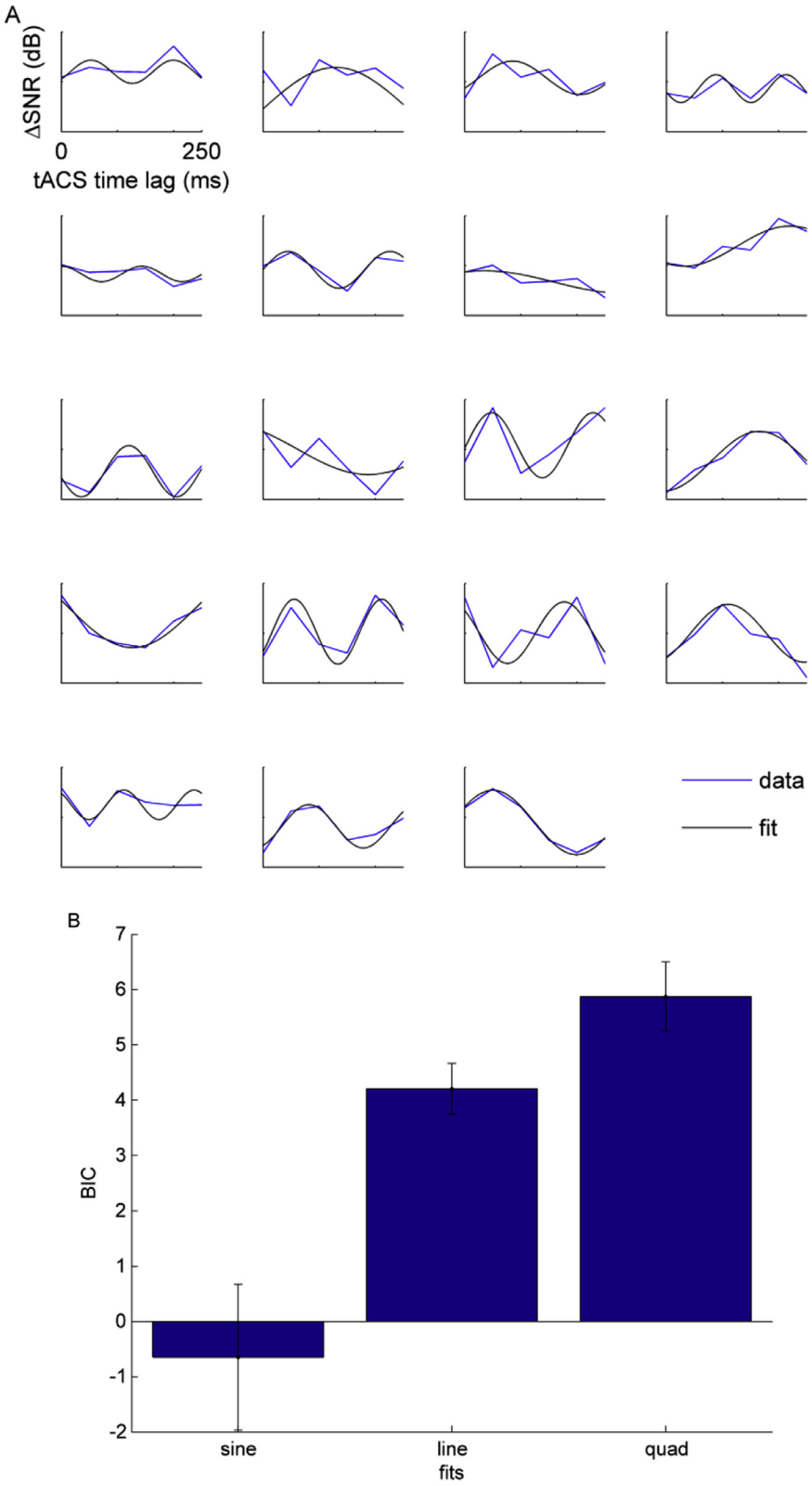

Fig. 3. Curve fitting analysis from Wilsch et al. applied to a null hypothesis dataset. **A**. A derived dependent variable was calculated (ΔSNR) and the null hypothesis data from Fig. 2 was plotted for each individual subject (blue line). The sine curve fitting procedure described in Wilsch et al. gives a reasonable fit to the data (black line). Note, there is no sine modulation present in the simulated null hypothesis data. The sinewave simply captures many of the features present in the noise. **B**. The three curve fitting procedures described in Wilsch et al. were applied to the null hypothesis data. The mean Bayesian Information Criteria (BIC) for each fit is shown (error bars show standard error). The sine function gives a better fit to the null hypothesis data than either a linear or quadratic fit. This follows the same pattern reported by Wilsch et al. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

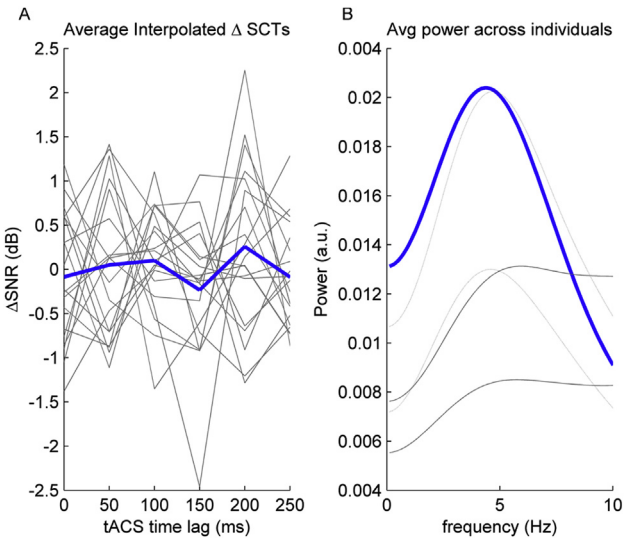

Fig. 4. Spectral analysis from Wilsch et al. applied to a null hypothesis dataset. **A**. The null hypothesis data for each subject was up-sampled and is shown in grey. The average is shown in blue. **B**. The power spectral density of the up-sampled null hypothesis data for each subject was calculated. The average power spectral density (blue) of the null hypothesis simulated data peaks at around 5 Hz. This is similar to the measured data reported in Wilsch et al. A noise floor was calculated by permuting the up-sampled data – dark grey lines. This is similar to the noise floor shown in Wilsch et al. The correct noise floor should be calculated by permuting the original (not up-sampled) data. This correct noise floor is shown in light grey. In our case, the peak at 5 Hz is simply a consequence of having the 6 noisy data points separated by 50 ms and cannot be interpreted as evidence of 5 Hz tACS modulation. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Neuling et al. We calculated the same normalized threshold and then applied the same zero crossing alignment procedure. Neuling et al. did not specify what they did if they did not detect a zero crossing or if they detected two zero crossings. In our analysis, if we did not detect a zero crossing we subtracted the mean from that individuals data, thereby centering the data around zero. If we detected two zero crossings we aligned the one associated with the largest difference in normalized threshold to 180°. The phase aligned normalized thresholds from our null hypothesis dataset are shown in Fig. 5A (compare with Fig. 3 in Ref. [16]). We analyzed the phase aligned null hypothesis dataset using an ANOVA and found a significant main effect of phase. We repeated this 100 times and found a significant main effect 89 times. Data from phase bin in positive half of the sinewave showing the largest effect were compared to data from the phase bin in the negative half showing the largest effect (Fig. 5B). A two sided paired *t*-test showed that these two groups were significantly different. We repeated this 100 times and found a significant effect 94 times. With an alpha value of 0.05 we would expect around 5 false positives out of 100. In our case, both of these statistical effects are false positives, resulting from bias introduced by the phase alignment procedure.

The null hypothesis datasets were fitted with a sinewave for each subject, using nonlinear regression analysis (MATLAB, lsqcurvefit.m) Fig. 5C shows the raw data for each subject (blue lines) and the average fitted sinewave (red line). We selected two exemplary fits by selecting subjects with the lowest residuals from the fitting procedure (Fig. 5D). We still obtain reasonable fits to a sinewave with null hypothesis dataset because of the low sample resolution of 6 phase bins spread across one cycle.

Thus, we cannot conclude from the analysis presented in Neuling et al. that tACS phase modulates detection thresholds of a tone in noise.

Riecke et al., 2015 under null hypothesis

In Riecke et al. [17] the dependent variable was a subject's accuracy at detecting an acoustic click. They measured detection accuracy for 14 subjects across 6 different tACS phase conditions and one sham condition. They applied a phase alignment procedure by detecting each subject's best performing condition and aligning this to 90°. Riecke et al. did correct for the analytical bias by excluding the phase aligned 90° data points from all further analysis. They averaged data from the phase angles of the positive part of the sinewave together (30° and 150°) and averaged data from the phase angles of the negative part of the sinewave together (210°, 270° and 330°). Compared to sham they found no significant difference with the positive or negative sinewave conditions. However, the negative and positive conditions were significantly different. Thus, when analytical bias was corrected, there was a modulatory effect of tACS on auditory perception but neither tACS condition was different from sham.

We assumed the null hypothesis and simulated a dataset containing 6 tACS phase conditions and one sham. The simulation used one distribution with a mean of 70 and standard deviation of 7.4, approximately matching reported values. We applied the reported phase alignment procedure. The average results are shown in Fig. 6A and the six simulated subjects that were best fit with a sinewave are shown in Fig. 6B. We applied the same correction procedure by excluding the 90° condition and averaging together data from the positive and negative sinewave halves. We did not find a significant difference between any of the averaged conditions and sham (Fig. 6C). Nor did we find a significant effect between averaged data from the negative and positive sinewave halves. We repeated this simulation 100 times and only found a significant effect 6 times, approximately in line with an alpha value of 0.05. Thus, the null hypothesis dataset could not account for the reported effect of tACS on acoustic click detection. This shows that by correctly accounting for the analytical bias Riecke et al. were able to detect a true effect of tACS on auditory perception.

To test if the phase effect oscillated at 4 Hz Riecke et al. computed the power spectral density of the time-series data and found that the power at 4 Hz was significantly greater than at 12 Hz. An ANOVA showed a significant effect of frequency. They concluded that this supports the idea that tACS causes an oscillatory modulation of auditory perception which is greatest at the tACS frequency. After running the simulation 100 times a repeated measures ANOVA detected an effect of frequency 37 times, while a paired *t*-test corrected for 2 comparisons found differences between the 4 and 12-Hz 22 times. These are certainly elevated compared to the expected 5% error rate, indicating that the analysis is biased towards detecting an oscillation at 4 Hz.

Two other studies [18,19] from Riecke et al. used similar analytical methods and bias correction procedures. We implemented the analytical methods used in both papers and tested if a null hypothesis dataset could account for the reported results. This was not the case for either of the two studies (results not shown here, but MATLAB analysis scripts are included in the Supplementary Information). Thus, again by correctly accounting for analytical bias, Riecke et al. were able to detect true effects of tACS on auditory perception.

Discussion

We used simulated data to test the hypothesis that some of the reported phase-effects of low-frequency tACS on auditory perception could be accounted for by analytical bias. We found that the null hypothesis datasets produced similar results to those reported in studies that did not correct for analytical bias [15,16]. We showed

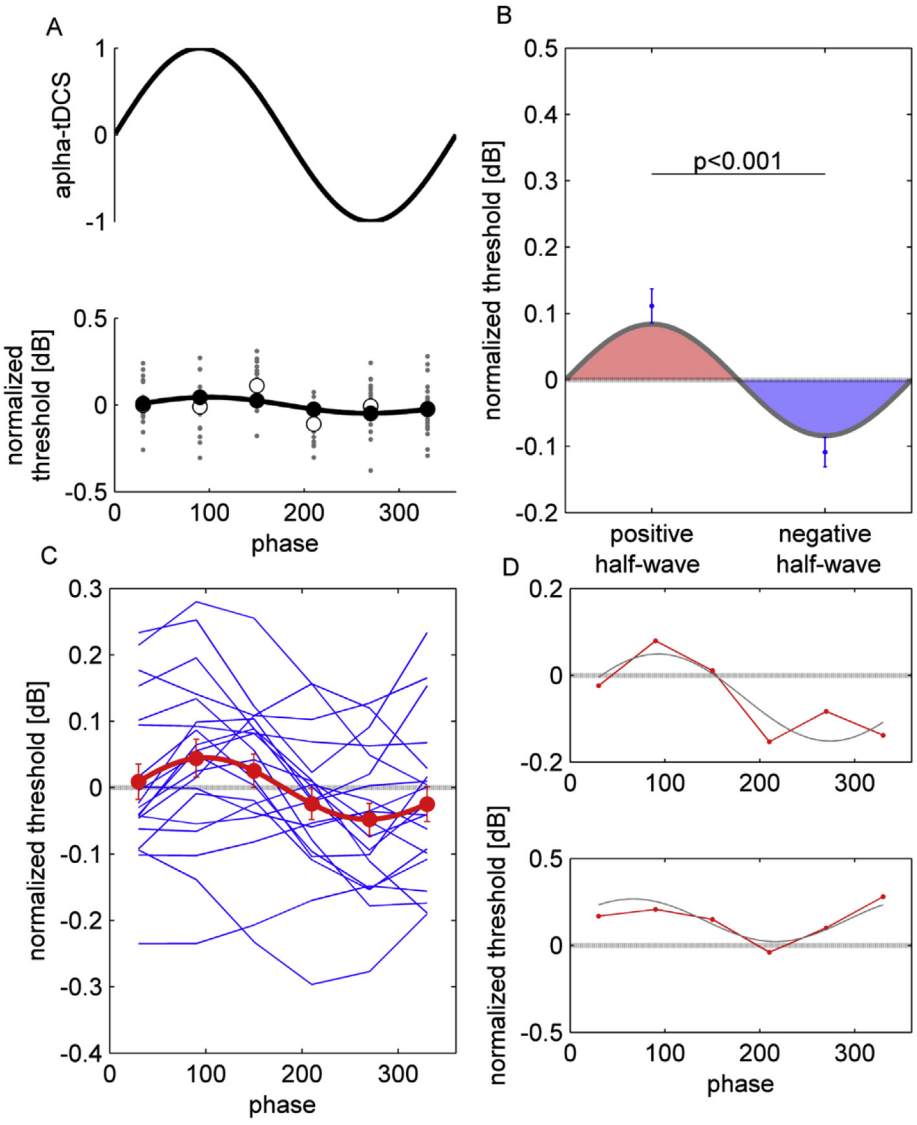

Fig. 5. Applying the Neuling et al. analysis to a null hypothesis dataset consistently gives false positives. **A**. The null hypothesis dataset was simulated by drawing data for each tACS phase condition and the sham condition from the same normal distribution. The null hypothesis data was normalized and a phase alignment procedure was applied as described in Neuling et al. The grey dots show the individual data points. The white points show the average for each phase bin and the black line and dots show the fitted sine function. The null hypothesis dataset contains no sinewave modulation. The apparent modulation is simply a consequence of the phase alignment procedure. **B**. Shows the mean and standard error of the highest phase condition from the positive half of the sinewave and compares this to the lowest phase condition from the negative half of the sinewave. Even for null hypothesis data, these conditions are significantly different. This is a direct result of the phase alignment procedure which introduces bias into the analysis. **C**. Data from each individual simulated subject is shown in blue and the mean sinewave fit is shown in red. Error bars show standard error. **D**. We selected two simulated subjects with the best fit for display. There is no sinewave modulation present in the null hypothesis data. However, a combination of chance effects and low resolution of the phase bins leads to some data being well fit by a sinewave. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

that the analysis used in those studies consistently produces false positive results. This makes it difficult to accept the specific conclusions relating to tACS effects on auditory perception reported in those studies. Fortunately, null hypothesis datasets could not reproduce the results from the studies that did correct for the analytical bias [17–19], indicating that tACS may cause small but significant effects on auditory perception.

Assessing phase effects on perception

Here we reviewed the analysis used in studies that investigated tACS phase effects on auditory perception. However, perception or function in non-auditory systems can also show phase specific modulations. Gundlach et al. [23] used tACS to phase modulate somatosensory perception. Because the highest modulation occurred at different phases for each subject, they also employed a phase alignment procedure for the group analysis. The authors noted that this alignment procedure would introduce bias and subsequently used a permutation approach to account for the bias. In Nakazono et al. [22] authors examined the effects of tACS phase and frequency on motor cortex excitability and found an effect for both parameters. However, they did not use a phase alignment procedure, presumably because most subjects showed peak modulations at a similar phase. Therefore, no bias correction was necessary. Outside the context of tACS Busch and van Rullen [24] and Busch et al. [25] looked at performance in a visual attention task as a function of the phase of an ongoing endogenous brain oscillations. In these studies, authors aligned each subject based on their best performing phase. Importantly, they then ignored the peak in performance at the aligned phase (mentioning that this

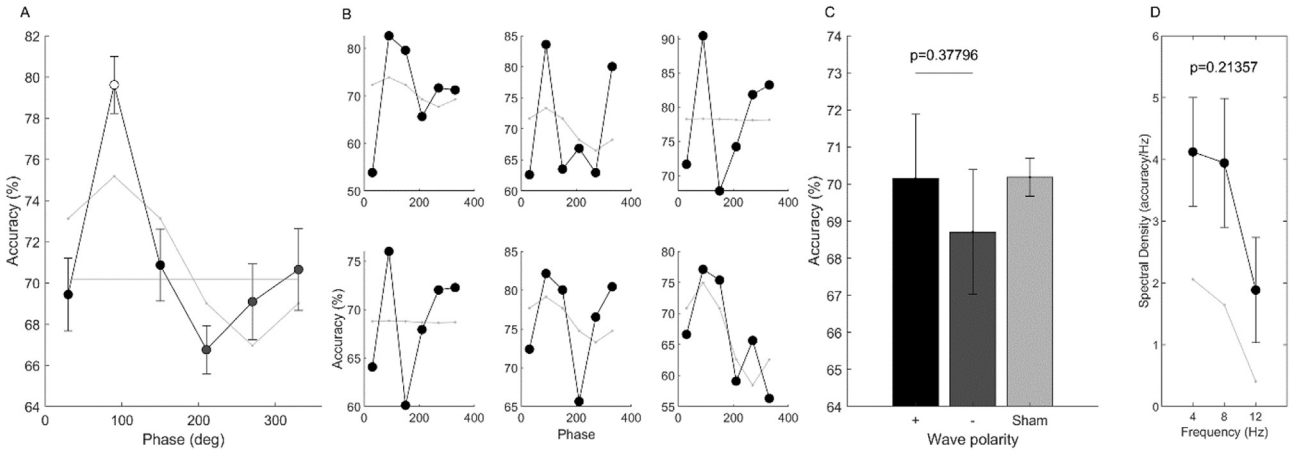

Fig. 6. Analysis of null hypothesis dataset cannot reproduce all the tACS effects reported by Riecke et al., 2015. A. The null hypothesis dataset was simulated by drawing data for each tACS condition and the sham from the same distribution. As described by Riecke et al., 2015, a phase alignment procedure was applied to align the best condition to the 90° phase bin. The large circles and error bars show the mean and standard error for each phase bin. The horizontal grey line shows the mean sham condition. B. The six simulated subjects (black circles and line) showing the best fit to a sinewave (grey line) are shown. C. To compensate for the analytical bias Riecke et al., 2015 removed the 90° phase bin from the analysis. After compensating for analytical bias, they still found a small modulatory effect of tACS on auditory perception. We implemented the same analysis and removed the 90° phase bin and could not find this effect with our null hypothesis dataset. This indicates Riecke et al., 2015 detected a true modulatory effect of tACS on auditory perception. D. The mean power spectra of the time series data for each subject is shown (black circles and lines). Repeating the spectral analysis simulation 100 times showed a somewhat elevated error rate (22 out of 100 repetitions).

peak was trivial) and drew conclusions from the trends in performance at the non-aligned phases. In summary, these studies highlight how different strategies can be used when assessing phase effects on perception or function. As a general principle, when a phase alignment procedure is employed it will create analytical bias at the aligned phase which must be corrected for.

Accounting for analytical bias in tACS phase-effects

Auditory system delays are caused by mechanical delays in the middle and inner ear and neural delays in the auditory pathway. These delays can differ between individuals. An additional source of individual delay variability could be caused by a different orientation of cortical neurons relative to the tACS generated electric field. Therefore, if the aim of tACS experiments is to phase-modulate auditory cortical processing, it is necessary to use a delay tailored to the individual subject. When designing tACS experiment and analysis procedures, two approaches can be used to account for individual differences in delay.

The first, used by all studies reviewed here, is to measure performance in an auditory task under a range of delay conditions and a sham for all subjects. This is a valid approach. However, care must be taken when selecting one delay condition to compare to sham. Aligning the best performing (or zero crossing) condition for each subject and comparing this to sham introduces bias, making it impossible to distinguish between the null hypothesis and a true effect. Our simulations verify that this bias can be corrected for by removing the delay condition used for alignment, as done by Refs. [17–19].

A better approach could be to use a two-stage experiment. In the first stage task, performance for each subject is measured for a range of tACS delays. The best performing delay for each subject is then selected. In the second stage, task performance is measured for each subject using only their individual best delay and this is compared to sham. This design removes the possibility of introducing analytical bias and has the advantage of testing each subject only at their best performing condition (not possible with the single stage approach). This design will increase experimental time by roughly a factor 2.

tACS effects on neural responses

We focused on the reported phase effects of low-frequency tACS on auditory perception. However, two studies also reported effects of low-frequency tACS on the neural response. The neural response analysis differs from the perception analysis methods and do not appear to introduce bias. Neuling et al. [16] reported that low-frequency tACS increased alpha EEG power. Zoefel et al. [20] did not detect an effect of tACS on auditory perception. However, they found a modulatory effect of tACS on fMRI BOLD response to intelligible speech but not to unintelligible speech. Taken together this suggests that tACS may modulate neural responses in the auditory system but that the neural modulation may not always be associated with a measurable effect on perception.

On the other hand, recent studies indicate that the electric field in the brain generated by tACS may be too weak to directly modulate neural responses [9,10,26]. Thus, observed tACS effects on auditory neural and perceptual responses may not be caused by direct modulation of auditory cortex neurons. It has been demonstrated that many of the reported tACS motor system effects can be accounted for by stimulation of peripheral nerves in the skin rather than direct modulation of neurons in the motor cortex [13]. Similar mechanisms, such as the electrical stimulation of the retina or peripheral nerves in the skin providing a (perhaps sub-perceptual) cue should be carefully considered for tACS auditory system effects. Further experiments specifically designed to rule out these possibilities are needed to provide a definitive answer.

Conclusions

Six studies investigating the phase-effects of low-frequency tACS on auditory perception have been published. One study did not detect an effect on auditory perception. In two of the studies, the reported effects on auditory perception can be accounted for by analytical bias. The final three studies used valid approaches to account for analytical bias and could measure effects of tACS on auditory perception. Given these mixed findings, further work is needed to fully establish the extent of the phase-effects of low-frequency tACS on auditory perception.

Conflicts of interest

All authors confirm there are no conflicts of interest that could have influenced the outcome of this work.

Funding

This work was supported by KU Leuven Research Funding STG/14/024 and EGM-D2929-C24/17/091 and by an EIT Health Innovation by Ideas, NEURO-WEAR Project. Boateng Asamoah is SB PhD fellow at FWO 501100003130.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.brs.2019.03.011.

References

[1] Rufener KS, Zaehle T, Oechslin MS, Meyer M. 40 Hz-Transcranial alternating current stimulation (tACS) selectively modulates speech perception. Int J Psychophysiol 2016;101:18–24. https://doi.org/10.1016/j.ijpsycho.2016.01.002.

[2] Kanai R, Chaieb L, Antal A, Walsh V, Paulus W. Frequency-dependent electrical stimulation of the visual cortex. Curr Biol 2008;18:1839–43. https://doi.org/10.1016/j.cub.2008.10.027.

[3] Kanai R, Paulus W, Walsh V. Transcranial alternating current stimulation (tACS) modulates cortical excitability as assessed by TMS-induced phosphene thresholds. Clin Neurophysiol 2010;121:1551–4. https://doi.org/10.1016/j.clinph.2010.03.022.

[4] Marshall L, Helgadóttir H, Mölle M, Born J. Boosting slow oscillations during sleep potentiates memory. Nature 2006;444:610–3. https://doi.org/10.1038/nature05278.

[5] Neubauer AC, Wammerl M, Benedek M, Jauk E, Jaušovec N. The influence of transcranial alternating current stimulation (tACS) on fluid intelligence: an fMRI study. Pers Indiv Differ 2017;118:50–5. https://doi.org/10.1016/j.paid.2017.04.016.

[6] Pahor A, Jaušovec N. The effects of theta transcranial alternating current stimulation (tACS) on fluid intelligence. Int J Psychophysiol 2014;93:322–31. https://doi.org/10.1016/j.ijpsycho.2014.06.015.

[7] Joundi RA, Jenkinson N, Brittain JS, Aziz TZ, Brown P. Driving oscillatory activity in the human cortex enhances motor performance. Curr Biol 2012;22:403–7. https://doi.org/10.1016/j.cub.2012.01.024.

[8] Khatoun A, Breukers J, Beeck Op de S, Nica IG, Aerts J-M, Seynaeve L, et al. Using high-amplitude and focused transcranial alternating current stimulation to entrain physiological tremor. Sci Rep 2018;8:4927. https://doi.org/10.1038/s41598-018-23290-w.

[9] Lafon B, Henin S, Huang Y, Friedman D, Melloni L, Thesen T, et al. Low frequency transcranial electrical stimulation does not entrain sleep rhythms measured by human intracranial recordings. Nat Commun 2017;8:1–13. https://doi.org/10.1038/s41467-017-01045-x.

[10] Vöröslakos M, Takeuchi Y, Brinyiczki K, Zombori T, Oliva A, Fernández-Ruiz A, et al. Direct effects of transcranial electric stimulation on brain circuits in rats and humans. Nat Commun 2018;9. https://doi.org/10.1038/s41467-018-02928-3.

[11] Bikson M, Brunoni AR, Charvet LE, Clark VP, Cohen LG, Deng Z-D, et al. Rigor and reproducibility in research with transcranial electrical stimulation: an NIMH-sponsored workshop. Brain Stimul 2017. https://doi.org/10.1016/j.brs.2017.12.008.

[12] Héroux ME, Loo CK, Taylor JL, Gandevia SC. Questionable science and reproducibility in electrical brain stimulation research. PLoS One 2017;12:e0175635. https://doi.org/10.1371/journal.pone.0175635.

[13] Asamoah B, Khatoun A, Mc Laughlin M. tACS motor system effects can be caused by transcutaneous stimulation of peripheral nerves. Nat Commun 2019;10:266. https://doi.org/10.1038/s41467-018-08183-w.

[14] Näätänen R, Picton T. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. Psychophysiology 1987;24:375–425.

[15] Wilsch A, Neuling T, Obleser J, Herrmann CS. Transcranial alternating current stimulation with speech envelopes modulates speech comprehension. Neuroimage 2018;172:766–74. https://doi.org/10.1016/j.neuroimage.2018.01.038.

[16] Neuling T, Rach S, Wagner S, Wolters CH, Herrmann CS. Good vibrations: oscillatory phase shapes perception. Neuroimage 2012;63:771–8. https://doi.org/10.1016/j.neuroimage.2012.07.024.

[17] Riecke L, Formisano E, Herrmann CS, Sack AT. 4-Hz transcranial alternating current stimulation phase modulates hearing. Brain Stimul 2015;8:777–83. https://doi.org/10.1016/j.brs.2015.04.004.

[18] Riecke L, Sack AT, Schroeder CE. Endogenous delta/theta sound-brain phase entrainment accelerates the buildup of auditory streaming. Curr Biol 2015;25:3196–201. https://doi.org/10.1016/j.cub.2015.10.045.

[19] Riecke L, Formisano E, Sorger B, Başkent D, Gaudrain E. Neural entrainment to speech modulates speech intelligibility. Curr Biol 2018;28:161–9. https://doi.org/10.1016/j.cub.2017.11.033. e5.

[20] Zoefel B, Archer-Boyd A, Davis MH. Phase entrainment of brain oscillations causally modulates neural responses to intelligible speech. Curr Biol 2018;28:401–8. https://doi.org/10.1016/j.cub.2017.11.071. e5.

[21] Bayesian information Criteria. Inf. Criteria stat. Model. New York, NY: Springer; 2008. p. 211–37. https://doi.org/10.1007/978-0-387-71887-3_9.

[22] Nakazono H, Ogata K, Kuroda T, Tobimatsu S. Phase and frequency-dependent effects of transcranial alternating current stimulation on motor cortical excitability. PLoS One 2016;11:1–13. https://doi.org/10.1371/journal.pone.0162521.

[23] Gundlach C, Müller MM, Nierhaus T, Villringer A, Sehm B. Phasic modulation of human somatosensory perception by transcranially applied oscillating currents. Brain Stimul 2016;9:712–9. https://doi.org/10.1016/j.brs.2016.04.014.

[24] Busch N, VanRullen R. Pre-stimulus EEG oscillations reveal periodic sampling of visual attention. J Vis 2010;10. https://doi.org/10.1167/10.7.219. 219–219.

[25] Busch NA, Dubois J, VanRullen R. The phase of ongoing EEG oscillations predicts visual perception. J Neurosci 2009;29:7869–76. https://doi.org/10.1523/jneurosci.0113-09.2009.

[26] Huang Y, Liu AA, Lafon B, Friedman D, Dayan M, Wang X, et al. Measurements and models of electric fields in the in vivo human brain during transcranial electric stimulation. Elife 2017;6. https://doi.org/10.7554/eLife.18834.