Short-term transcutaneous vagus nerve stimulation increases pupil size but does not affect EEG alpha power: A replication of Sharon et al. (2021, Journal of Neuroscience)☆

Beth Lloyd [a,b,*,1], Franz Wurm [a,b,1], Roy de Kleijn [a,b], Sander Nieuwenhuis [a,b]

[a] Institute of Psychology, Leiden University, Leiden, the Netherlands
[b] Leiden Institute for Brain and Cognition, Leiden University, Leiden, the Netherlands

ABSTRACT

Background: Transcutaneous auricular vagus nerve stimulation (taVNS) has been tested as a potential treatment for pharmaco-resistant epilepsy and depression. Its clinical efficacy is thought to depend on taVNS-induced activation of the locus coeruleus and other neuromodulator systems. However, unlike for invasive VNS in rodents, there is little evidence for an effect of taVNS on noradrenergic activity.
Objective: We attempted to replicate recently published findings by Sharon et al. (2021), showing that short bursts of taVNS transiently increased pupil size and decreased EEG alpha power, two correlates of central noradrenergic activity.
Methods: Following the original study, we used a single-blind, sham-controlled, randomized cross-over design. Human volunteers (n = 29) received short-term (3.4 s) taVNS at the maximum level below the pain threshold, while we collected resting-state pupil-size and EEG data. To analyze the data, we used scripts provided by Sharon and colleagues.
Results: Consistent with Sharon et al. (2021), pupil dilation was significantly larger during taVNS than during sham stimulation ($p = .009$; Bayes factor supporting the difference = 7.45). However, we failed to replicate the effect of taVNS on EEG alpha power ($p = .37$); the data were four times more likely under the null hypothesis ($BF_{10} = 0.28$).
Conclusion: Our findings support the effectiveness of short-term taVNS in inducing transient pupil dilation, a correlate of phasic noradrenergic activity. However, we failed to replicate the recent finding by Sharon et al. (2021) that taVNS attenuates EEG alpha activity. Overall, this study highlights the need for continued research on the neural mechanisms underlying taVNS efficacy and its potential as a treatment option for pharmaco-resistant conditions. It also highlights the need for direct replications of influential taVNS studies.

1. Introduction

Vagus nerve stimulation is an invasive technique used to electrically stimulate the cervical vagus nerve [1], which has shown promise in treating pharmaco-resistant epilepsy and depression [2,3]. The vagus nerve has major afferent connections to the nucleus of the solitary tract [4], which in turn directly and indirectly modulates several neuro-modulator systems, including the noradrenergic locus coeruleus [5]. Consequently, animal studies have found that invasive VNS (iVNS)

increases the activity of the noradrenergic [6–10] and cholinergic system [11]. Recently, transcutaneous auricular vagus nerve stimulation (taVNS) has gained traction as a non-invasive alternative to iVNS [12]. taVNS is thought to mimic the effects of iVNS by delivering alternating currents to the auricular branch of the vagus nerve via surface skin electrodes on the outer (left) ear. However, it is still uncertain if taVNS enhances noradrenergic activity, and hence the working mechanisms underlying taVNS are poorly understood [13].

Several studies have examined the effects of taVNS on noninvasive

markers of noradrenergic activity. Invasive VNS in human patients has been found to modulate the amplitude of the P300 [14,15], a cognitive evoked potential and purported correlate of phasic noradrenergic activity [16,17]. However, taVNS studies have produced mixed evidence for an effect on P300 amplitude [18–22]. Studies on the effect of taVNS on salivary alpha-amylase, an indirect hormonal marker of noradrenergic activity [23,24], have also yielded mixed evidence, although a recent pooled mega-analysis found a small but significant taVNS-induced increase in salivary alpha-amylase [25]. Pupil size is considered the most reliable noninvasive marker of noradrenergic activity [26], and iVNS has been found to increase pupil size in humans [24] and animals [25,26]. So when studies began reporting null effects of taVNS on event-related pupil dilation and baseline pupil size [13,21, 27–30], this led to serious doubt about the efficacy and working mechanisms of taVNS [13]. However, then an important study appeared [31], which found that short-term bursts of taVNS (3.4 s) evoked immediate but transient modulations of pupil size and EEG alpha power (8–12 Hz), another potential correlate of noradrenergic activity [32].

Several authors have pointed out that Sharon et al. [31] used higher stimulation intensities (individually calibrated, 2.20 ± 0.24 mA) and shorter stimulation epochs (3.4 s) than the studies cited above, most of which used a fixed intensity of 0.5 mA and stimulation epochs on the order of minutes (sometimes with a 30-s on/30-s off schedule, the regime for therapeutic use of the NEMOS® device). Therefore, the findings of Sharon and colleagues may offer important clues regarding the taVNS parameter values that are most effective in experimental research and clinical settings [33]. However, given the many studies with null effects of taVNS, it remains to be determined if the pupil and EEG results of Sharon and colleagues are replicable. To that end, we closely replicated their experiment and systematically examined the simultaneously acquired EEG and pupil-size measurements from a sample of healthy males and females (n = 29).

2. Materials and methods

2.1. Participants

Thirty-three young adults participated in return for 17 euros or course credits. Data from four participants were excluded because of technical issues, resulting in a final sample of 29 participants (21 female [sex assigned at birth]; mean age 22.5 ± 3.8 years). For the EEG analysis, a further five participants were excluded due to a lack of alpha spectrum and topography in the break data. Therefore, 24 participants remained for the EEG analysis (18 female; mean age 22.2 ± 3.7 years). The final sample sizes of 29 (pupil) and 24 (EEG) exceeded the minimum sample sizes needed (23 and 13, respectively), as indicated by a power analysis with $\alpha = 0.05$, power = 0.80, and effect sizes derived from Sharon et al. (2021; Cohen's d for pupil: 0.62, EEG: −0.87). Individuals with neurological or cardiac disorders, individuals who used psychoactive medication, and individuals who received taVNS stimulation in the four weeks prior to the study were excluded from participation. Participants were instructed not to consume alcohol and caffeine in the 24 h and 3 h, respectively, before the start of the 2-h session. The study was approved by the Psychology Research Ethics Committee at Leiden University (S.T.-V1-3624).

2.2. Procedure

The current study was carried out as closely as possible to the study of Sharon et al. [31], using task and analysis scripts kindly provided by

Sharon and colleagues.[2] The study design was single-blind and sham-controlled and the order of conditions (taVNS first: N = 14 vs. sham first: N = 15; Fig. 1A) was counter-balanced across participants. After setting up the EEG, we carried out a methods-of-limits procedure for each of the two conditions to identify the maximum comfortable stimulation level for each individual. We repeatedly applied 5 s of stimulation, after which participants rated the subjective intensity on a scale of 0–10 (0 = no sensation, 3 = light tingling, 6 = strong tingling, 10 = painful). The level of stimulation started at 0.1 mA and increased by 0.2 mA until a rating of 9 or a maximum stimulation level of 5 mA was reached. The intensity corresponding to the subjective rating of 9 (just below painful) was selected.

Next, the participants were instructed to fixate on a white fixation cross on a grey background (RGB: 125, 125, 125). There were eight blocks of 11 trials, each lasting ~5 min (Fig. 1B). Each trial started with 3.4 s of stimulation, followed by an inter-stimulus interval jittered between 25 s and 27 s. After every two blocks participants responded to 11 questions regarding their subjective experience of stimulation and were free to take a break. Then the condition switched between taVNS and sham. The EEG data collected during the breaks (mean length: 40 s) were used to characterize individual alpha activity. At the end of the session we asked participants whether they thought the final block involved taVNS or sham stimulation. Two participants did not respond to this question. Of the other participants, 48% responded correctly, suggesting that participants could not distinguish between taVNS and sham.

2.3. Transcutaneous auricular vagus nerve stimulation

We used the NEMOS® device (tVNS Technologies GmbH) to deliver electrical stimulation to the left cymba conchae (taVNS) or the left ear lobe (sham [34]; Fig. 1A). The titanium electrodes were covered with cotton rings soaked in electrolyte conductive liquid to establish good connection for stimulation and to prevent direct contact with the skin. Pulses (width: 200–300 μs; shape: rectangular, biphasic, asymmetric) were delivered at a rate of 25 Hz and an intensity as described above. To achieve precisely 3.4 s of stimulation, the NEMOS® device was controlled using two linear actuators which pressed the ON/OFF buttons. These 'push' commands were sent at pre-programmed times via an Arduino mini, and were embedded within the task script, which was programmed using the Expyriment Python package [35]. Note that the NEMOS® device used by us and Sharon and colleagues gradually ramps up the intensity during each trial, meaning that we administered the selected stimulation intensity for a period shorter than the 3.4-s stimulation period.

Two checks were implemented to ensure that only trials with good connection between the electrodes and the skin were included. First, the taVNS device automatically stops stimulating once connection has been disrupted. Second, we placed an additional external EEG electrode at the back of the left earlobe, and one on the auricle. Signal from these two electrodes showed a clear artifact during stimulation. Trials in which this artifact was not visible or did not show the pattern of a steady ramp-up of stimulation with 25-Hz pulses were removed from analysis (see EEG section for details).

2.4. Pupillometry

Data acquisition. The eye-tracker was positioned 75 cm from the participant's eyes. Pupil size was recorded from the dominant eye at a sampling rate of 40 Hz using a Tobii Pro eye-tracker. Eye gaze was

---

[2] In case of a discrepancy between the analysis scripts and the information in the article of Sharon et al. [31], we compared both procedures and found essentially the same results. Here we report the results based on the procedures as reported in the scripts.
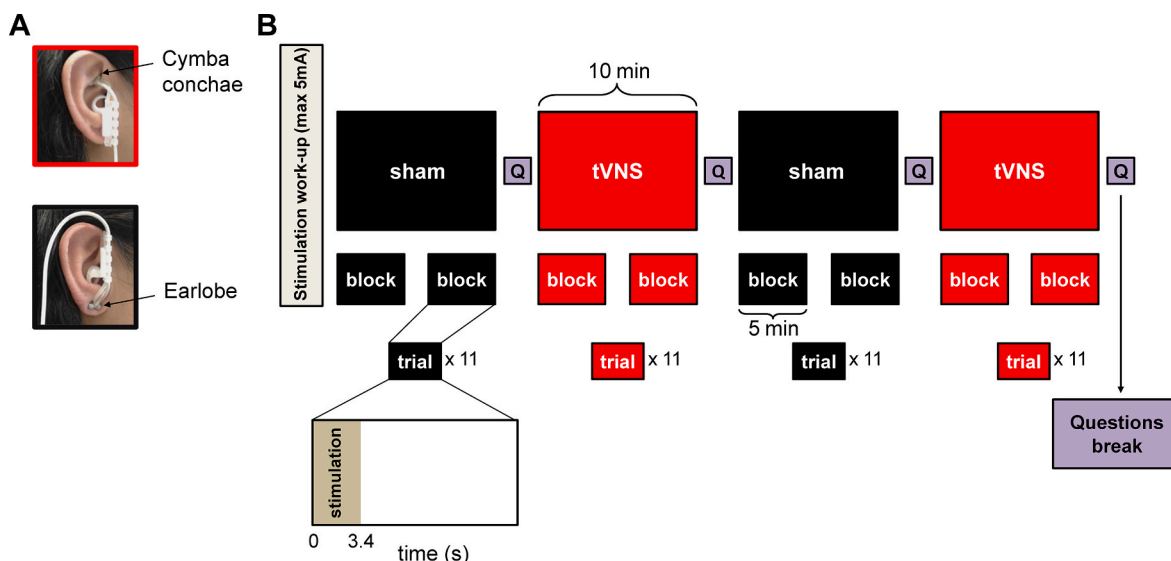
Fig. 1. Study design for testing the effects of taVNS on pupil diameter and alpha activity. A) The placement of the electrodes at the (left) cymba conchae (taVNS condition) and the (left) earlobe (sham condition; picture adapted from Ref. [36]). B) The experimental sequence, as in Ref. [31].

measured at a sampling rate of 120 Hz. Eye positions were transformed to degrees of visual angle based on a five-point calibration procedure. The experiment was carried out under constant ambient light.

Data analysis. Pupil data were preprocessed using PupCor (http s://github.com/lindvoo/PupCor). Blinks were removed by linearly interpolating over the periods with invalid samples (automatically marked by the device manufacturer) from 100 ms before blink onset to 400 ms after blink offset. Next, the data were manually checked and corrected if any artifacts had not been successfully removed. Pupil-size data were then low-pass filtered using a 10-Hz fourth-order Butterworth filter with zero phase shift, and segmented into trials by extracting the samples from −10 s to +13.4 s around stimulation onset. Trials in which >50% of the samples were marked as invalid were excluded, resulting in an average of 38.9 ± 2 trials remaining for the taVNS condition and 39.1 ± 1.9 trials for the sham condition (88.4% and 88.9%, respectively). As in Sharon et al. [31], pupil size was converted to 'percentage change' values relative to a 10-s baseline before stimulation onset using the formula: $\left(\frac{(x-baseline)}{baseline}\right) * 100$. Although baseline pupil size did not differ between conditions ($M_{taVNS} = 3.69 \pm 0.6$ mm, $M_{sham} = 3.73 \pm 0.72$ mm, $p = .59$), subtractive baseline correction is preferred over the divisive baseline correction used by Sharon and colleagues [37]. Fortunately, we replicated the significant effect of taVNS on pupil size when using subtractive baseline correction (Supplementary Fig. 1).

After this preprocessing procedure, the pupil time series was averaged across trials, separately for the two conditions (taVNS, sham), resulting in two pupil time series per participant that were used for statistical analyses.

2.5. EEG

Data acquisition. The EEG was recorded using a Biosemi Active-Two system (BioSemi) with 64 Ag–AgCl electrodes. To record the vertical and horizontal electrooculogram (EOG), electrodes were placed above and below the left eye and on the outer canthi of both eyes. As described earlier, we also placed electrodes at the back of the earlobe and the auricle to record the stimulation artifact of the taVNS device. EEG, EOG and supporting electrodes were continuously recorded at a sampling rate of 512 Hz and with an electrode impedance <50 kΩ.

EEG analyses were performed using the EEGLAB toolbox [38], the Fieldtrip toolbox [39], and custom-written routines provided by Sharon and colleagues. EEG data were re-referenced to electrode Cz and

high-pass filtered to exclude frequencies below 0.1 Hz. Continuous data were segmented to 15-s epochs (−5 to +10 s) around stimulation onset. Epochs were detrended linearly and notch-filtered at 50 Hz. Then the epoched data were visually inspected to confirm all sham and taVNS trials contained a 25-Hz stimulation artifact. Trials without the artifact, caused by problems with the robot actuator or loss of contact between an electrode and the skin, were excluded from all analyses (overall: 9.4 ± 2.4%; taVNS: 6.9%; sham: 12.6%). To focus on the alpha frequency band, we applied a third-order two-pass Butterworth band-pass filter to exclude frequencies below 5 Hz and above 15 Hz. An additional notch filter at stimulation pulse frequency (25 Hz + harmonics up to 100 Hz) was implemented, to remove possible residual artifact activity. Finally, epochs were excluded whenever activity in a channel exceeded ±100 μV in the automatic iterative process specified in the original study. On average, this procedure led to the interpolation of 3.5 ± 0.90 channels and the removal of 0.33 ± 0.13 trials.

After preprocessing was completed, the data were re-referenced to the average reference. The mean number of valid trials was 40.1 ± 1.1 (out of 44) in the taVNS condition, and 37.2 ± 2.0 in the sham condition. Data were downsampled to 64 Hz (Nyquist frequency of 32 Hz > maximum alpha frequency of 15 Hz) and transformed to the time-frequency domain using the Morlet wavelet approach (number of cycles = 7; standard deviations of Gaussian kernel = 3). Applied frequencies increased from 5 to 15 Hz in 31 equally spaced steps. Taken together, this resulted in a frequency resolution of 0.33 Hz and a temporal resolution of 15.6 ms.

To identify participants' resting-state alpha topography and frequency spectrum, we first extracted data from the breaks between blocks. These were segmented into 5-s epochs, so that there was an overlap of 1 s with both the preceding and subsequent epoch. The obtained epochs underwent the same preprocessing as the stimulation trials and were subsequently reduced to 3-s epochs (discarding the overlap) to avoid filtering artifacts at the edges. This resulted in an average of 34.8 ± 3.6 epochs per participant. Next, we converted the epochs to the time-frequency domain and used them to identify participants' alpha topography and frequency spectrum using the PARAFAC method [40], as implemented in the N-way toolbox [41]. In line with Sharon et al. [31], we set non-negativity as the constraint for each dimension. The number of components was determined using the core consistency diagnostic (CCD), using a minimal CCD value of 55% as the guiding principle [42]. The average CCD was 87.8 ± 3.0%, resulting in an average number of components of 1.88 ± 0.14 per participant. Please

refer to Sharon et al. [31] for a more detailed depiction of the procedure.

The alpha components were visually derived from the PARAFAC models (Supplementary Fig. 2). After selection of the optimal number of components, the topographies of the respective model's components were inspected for occipital clusters. Afterwards, those components were screened for an alpha frequency profile. For 19 subjects this procedure clearly identified a single alpha component. Five participants had to be excluded due to a lack of an identifiable alpha component. Data from five participants contained multiple components that satisfied the above mentioned criteria of occipital/parietal cluster and alpha frequency range. For such ambiguous cases, we always chose the component with a higher order of descriptiveness of the data, as indicated by a lower component number. Supplementary Fig. 2 also includes the following information for each individual: sex, CCD value and size of stimulation effect (taVNS - sham) on alpha power. The identified components' channel and frequency spectrum from the break data was used as weights to be multiplied with the trial data separately for each participant. This procedure resulted in a single channel which represents the weighted activity of both spatial and frequency domain. We then applied baseline correction by subtracting the mean activity in a time window of −4 to 0 s before stimulation onset from the individuals' average power.

In line with Sharon et al. [31], we conducted two more analyses on the stimulation data. First, we investigated the time-frequency changes irrespective of the spatial dimension. Therefore, we multiplied the data only with the alpha topographies derived from the PARAFAC decomposition, but ignored the frequency distribution. This allowed us to plot the whole spectrogram at 5–15 Hz as percentage change relative to the same baseline (−4 to 0 s). Second, we investigated the spatial patterns irrespective of frequency distributions. Therefore, we multiplied the data only with the alpha frequency distribution derived from the PARAFAC decomposition, but ignored the topographies. This enabled us to plot the topographical changes relative to the same baseline (−4 to 0 s). Data from both analyses was submitted to cluster-based permutation testing.

2.6. Statistical analysis

Pupil data and data from subjective questionnaires were analysed using Python 3. To estimate the potential effects of taVNS on pupil dilation and subjective feelings of stimulation, we carried out Wilcoxon signed-rank tests using the function *scipy.stats.wilcoxon*. For analyzing the pupil time series, non-parametric *t*-tests were carried out using *neurotools.stats.permtest_rel*. To correct for multiple comparisons, alpha levels (set at 0.05) were adjusted by controlling the false discovery rate (FDR; [60]). Cluster-based permutation analyses for the EEG data were performed using Fieldtrip's ft_freqstatistics function. We used the Monte-Carlo method and a dependent-samples *t* statistic with 10,000 permutations. Alpha level was set to 0.05 after FDR cluster correction. The time window for the cluster-based permutation analysis was −1 to 6 s. To quantify the effect of taVNS on pupil dilation and alpha power (extracted from the period in which Sharon and colleagues found significant effects of taVNS), we carried out a Bayesian paired samples *t*-tests (using R version 4.0.3, library: *BayesFactor*). We used the default Cauchy prior r = 0.707 in JASP v0.9 [43]. All correlations reported are Spearman's rank correlation coefficients. Throughout the paper, data are expressed as the mean ± standard error of the mean (SEM).

3. Results

Stimulation intensity and participants' subjective ratings of stimulation are presented in Fig. 2. We found no significant differences between the taVNS and sham conditions in levels of subjective averseness (i.e., pain, irritation, mood, alertness; $ps > .05$). Individual stimulation intensities were higher for the sham condition ($M_{sham} = 3.0 \pm 1.1$ mA) than for the taVNS condition ($M_{taVNS} = 2.3 \pm 1.3$ mA, $p = .003$), in line with Sharon et al. [31].

3.1. Pupillometry

taVNS led to a significant increase in pupil dilation compared to sham stimulation (Fig. 3A). Pupil size reached a maximum of $5.6 \pm 1.6\%$ above baseline 4.2 s after taVNS onset, in comparison to sham stimulation which peaked at $3.8 \pm 1.6\%$ after 4.1 s. Pupil dilation was significantly stronger following taVNS between 2.88 s and 4.90 s ($ps < .05$, two-sided Wilcoxon signed-rank test, FDR-corrected across all timepoints). To determine the replicability of the effect found by Sharon et al. [31], we extracted the average pupil dilation between the two time points in which taVNS led to significantly larger pupil dilation compared to sham in their report (i.e., 2.88–5.96 s). Average pupil size was significantly larger for taVNS compared to sham (Wilcoxon signed-rank test; $p = .009$; Fig. 3B; Supplementary Fig. 1). Furthermore, a Bayesian two-sided *t*-test revealed moderate evidence in favor of this effect of taVNS on pupil dilation ($BF_{10} = 7.45$).

Eye gaze variability and blink rates did not differ between conditions ($ps > 0.51$). Indeed, our key findings remained robust and significant in
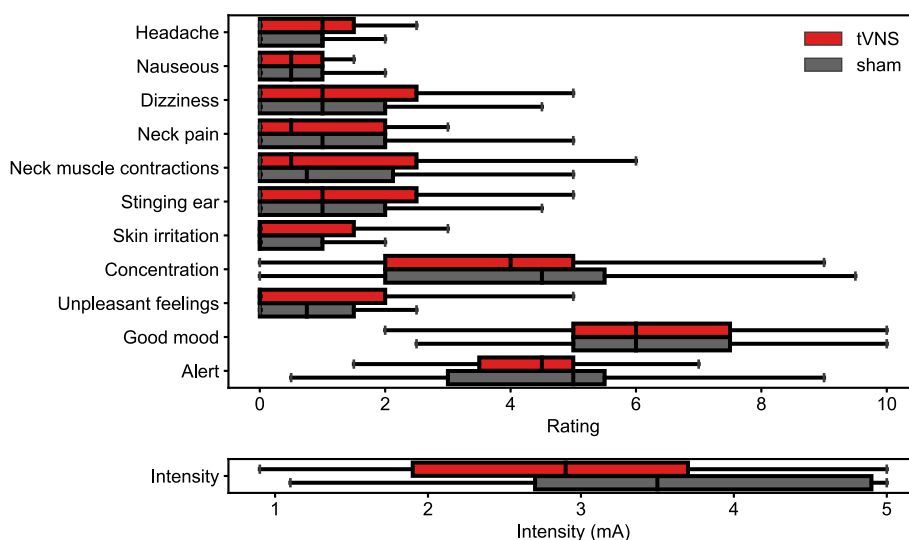

Fig. 2. Box plots indicating participants' subjective ratings of the stimulation, and the objective stimulation intensity levels. Only objective intensity level differed between taVNS and sham.
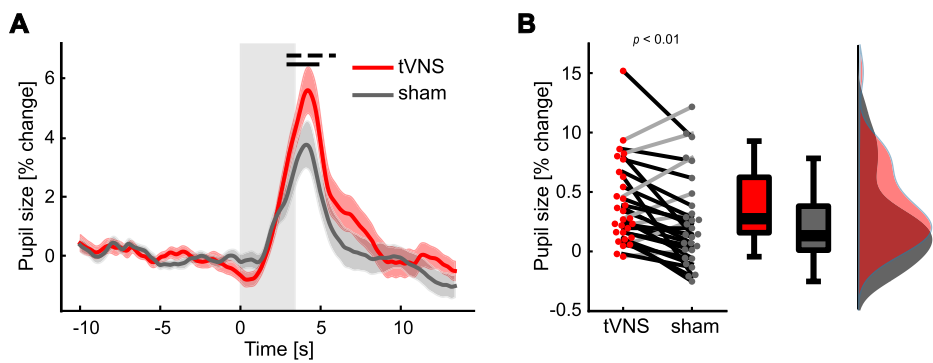
Fig. 3. Stronger pupil dilation following taVNS stimulation compared to sham. A) Grand-average pupil dilation waveforms (% change) relative to the 10 s before stimulation onset. Shaded red and grey areas indicate ± SEM. Grey-shaded rectangle indicates the period of stimulation (3.4 s). The horizontal black lines indicate the time period for which pupil dilation differed between taVNS and sham stimulation in the current study (solid line) and for [31] (dashed line). B) Individual data points, box plots and density plots indicating average pupil dilation during the period with a significant taVNS effect in Sharon et al. [31]. Black lines depict participants showing the expected effect (taVNS > sham, $n = 22$), light grey lines depict those who show the opposite effect (sham > taVNS, $n = 7$). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

control analyses in which we either (1) removed pupil data samples where eye gaze in the x or y direction was more than 3 SD away from the mean gaze co-ordinates (i.e., the position of the fixation cross); or (2) did not interpolate over blinks but removed invalid samples from the time series. We also correlated the difference between conditions (taVNS – sham) in objective stimulation intensity and the corresponding difference in evoked pupil response and found no correlation ($R = 0.18$, $p = .35$). However, correlating pupil response and stimulation intensity
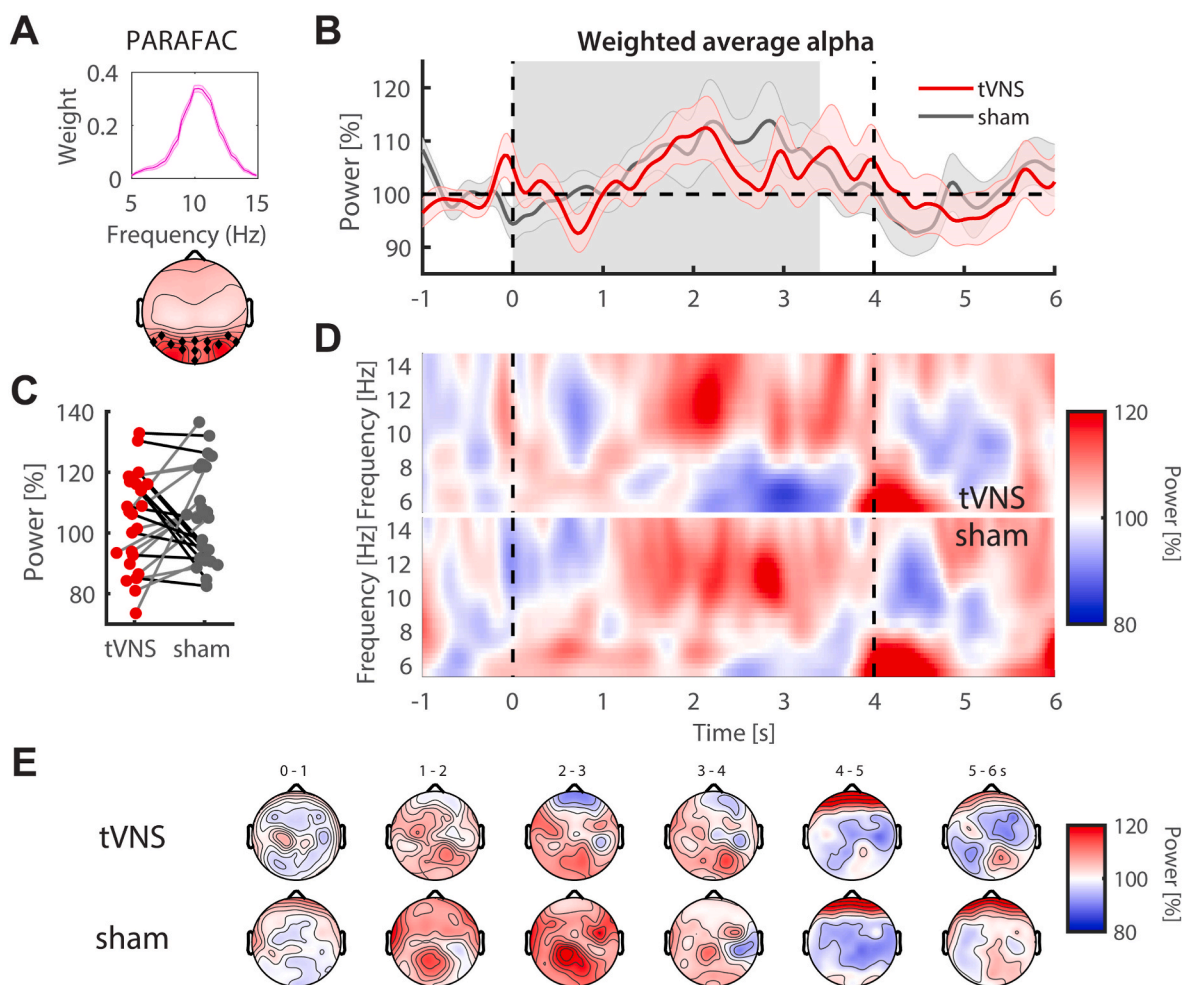

Fig. 4. No difference between taVNS and sham in occipital alpha power. A) Median alpha component spectral and spatial profile, as derived from the unbiased break data. B) Grand-average alpha power waveforms, using a weighted average of spectral and spatial profiles in panel A. Dashed vertical lines indicate the time window for the test of significance between taVNS and sham. The grey area denotes the time window of active stimulation (3.4 s). C) Individual data points for the weighted average alpha power in the time window between 0 and 4 s. Grey lines indicate participants with lower alpha for taVNS compared to sham (12 out of 24). Black lines indicate the reversed pattern. D) Mean induced spectrograms, using the weighting of the spatial profile in panel A. We did not identify any frequencies that showed a significant difference in power between stimulation conditions. E) Mean topographies, using the weighting of the spectral profile in panel A. We did not identify any channels that showed a significant difference in alpha power between stimulation conditions.

separately for each condition, we found a positive correlation for taVNS ($R = 0.39$, $p = .035$) but not for sham ($R = -0.15$, $p = .44$, statistical comparison between these correlations based on permutation testing: $p = .24$). These exploratory, uncorrected, tests reveal tentative evidence that objective stimulation intensity mediates the effect of stimulation on pupil size for taVNS, but not for sham.

3.2. EEG

To identify individual frequency and topography profiles of alpha oscillations for each participant, we employed PARAFAC decomposition to the unbiased break data. In line with Sharon et al. [31], we found individual alpha frequencies between 7 and 13 Hz and an expected occipital topography (Fig. 4A). Next, we used these individually derived profiles to obtain weighted average data for the frequency and spatial domain during stimulation trials. Contrary to the original study, we did not find an attenuation of alpha power during taVNS compared to sham in a 4-s time window after stimulation onset ($p = .89$, $BF_{10} = 0.23$; Fig. 4B and C). Numerically, electrical stimulation even *increased* weighted average alpha power, although this effect was not significant for taVNS ($105 \pm 3\%$, $p = .12$, $BF_{10} = 0.42$) or sham ($104 \pm 3\%$, $p = .23$, $BF_{10} = 0.65$). Similar results were obtained with selection of alternative components, dichotomized participant-specific channel and frequency weights, fixed weights for each participant, and for a neighboring frequency band (Supplementary Fig. 3).

Contrary to the original work of Sharon et al. [31], we did not find a significant correlation between the difference in weighted average alpha power (taVNS vs. sham) and the difference in stimulation intensity ($R = -0.04$, $p = .86$). Correlating alpha power and stimulation intensity separately for each condition, we also did not find a link for taVNS ($R = -0.21$, $p = .33$) and sham ($R = -0.30$, $p = .15$).

Following the original study, we also carried out follow-up cluster-based permutation analyses, separately for the time-frequency data (Fig. 4D; weighted average over spatial domain) and temporospatial data (Fig. 4E; weighted average over frequency domain). In line with our previous results, we did not identify significant clusters of differences between stimulation conditions for the time-frequency data (all $ps > 0.44$) and the temporospatial data (all $ps > 0.44$).

4. Discussion

Establishing effective biomarkers of taVNS that are consistent with central noradrenergic activity is relevant for our understanding of how taVNS can be used to study basic cognitive functions and improve clinical applications. We found, in line with Sharon et al. [31], that short taVNS pulses induced a transient pupil response that was larger than that induced by sham stimulation. However, we did not replicate the effects of taVNS on EEG alpha activity. Together, our results contribute to the ongoing research on taVNS by showing that taVNS can alter pupil size, a physiological marker associated with central noradrenergic [44] (and cholinergic [11]) activity. However, it raises questions about the replicability of a taVNS effect on EEG alpha activity.

In accordance with Sharon et al. [31], we found stronger pupil dilation following short bursts of taVNS compared to sham stimulation ($BF_{10} = 7.45$). This finding is inconsistent with other taVNS studies which used stimulation epochs on the order of minutes [13,21,27–30, 45], in some cases with a 30-sec on/30-sec off rhythm. Sharon et al. [31] were the first to assess the effects of short-term taVNS. Since then, others have reported increased pupil dilation following brief (600-ms) taVNS pulse trains, especially when stimulation was applied to the external ear canal [46], although this is not the optimal location for auricular tVNS [47]. The use of short bursts of stimulation could be crucial to trigger phasic, stimulus-evoked noradrenergic responses, as indexed by pupil dilation. For now, there is no evidence that taVNS can modulate *baseline* pupil size.

In addition, Sharon et al. [31] and we used a stimulation intensity

level that was individually tailored to be just below painful, while several previous taVNS studies applied substantially lower intensity levels (e.g. 0.5 mA) [13,21,28] without catering to individual sensitivity thresholds [30]. Low stimulation intensities may not be capable of activating the vagus nerve afferent fibers near the ear [48]. Indeed, two taVNS studies found parametric effects of stimulation intensity on pupil dilation, which nicely scaled with pulse amplitude [46,49]. These findings, together with evidence from rodents [11], suggest that stronger intensities are more likely to engage the vagus nerve and induce pupil dilation. Future taVNS studies must therefore ensure sufficient intensity—well above the perceptual threshold and below the pain threshold.

Although we found that at the group level taVNS was associated with greater pupil size than sham stimulation, the sham stimulation nevertheless elicited a large pupil response. This finding and a similar but less prominent pattern in Sharon et al. [31] suggest that short bursts of sham stimulation elicit a tactile pupillary orienting response [50,51]; and that a large part of the pupil response to short bursts of taVNS also reflects this orienting response, rather than the effect of vagus nerve stimulation. Furthermore, the finding that 7 out of 29 participants showed a larger pupil response to sham stimulation than to taVNS (Fig. 3B) questions the validity of pupil dilation as a biomarker for taVNS. That is, a biomarker that is of clinical value needs to differentiate between the presence and absence of vagus nerve activation at the level of individuals. The current findings resonate with prior unsuccessful attempts to identify valid biomarkers for taVNS [52].

We did not replicate the finding of Sharon et al. [31] that short-lasting taVNS transiently attenuated EEG alpha activity; our data were four times more likely under the null hypothesis. This discrepancy is reflective of the state of the literature on VNS and alpha power as measured over the scalp. Little is known about the effect of iVNS on alpha power, with the exception of a single-case study that found a significant reduction in alpha power, not during but immediately after 30 s of stimulation [53]. tVNS studies have yielded mixed results. Two of these studies, one of which used cervical instead of auricular tVNS, had a limited sample size (n = 8) [19,54]. In two other studies the results depended on the measurement location over the scalp [55], or on the exact alpha frequency selected and whether the participants' eyes were open or closed [56].

We believe that the most parsimonious explanation for these previous findings and the current results is that taVNS does not have a robust effect on alpha power. Although our EEG analyses were based on a similar sample size as those of Sharon and colleagues, the present results should arguably receive a higher weight because of the confirmatory nature of the study [57]. A possible explanation for the lack of a robust effect is that the connection between central noradrenergic activity and EEG alpha power is less strong than is often assumed. While in recent years the connection with locus coeruleus activity and pupil size has gained strong support (e.g., based on optogenetics or simultaneous measurements; [26]), the connection with alpha power is mainly based on a host of indirect evidence, including concurrent alpha and pupil size measurements [32]. Note that, in the current study we did not observe a significant correlation between the effects of taVNS (vs. sham) on pupil size and EEG alpha power ($R = -0.26$, $p = .21$; Supplementary Fig. 4).

It is important to mention a limitation of the current study. Whereas Sharon et al. [31] examined only male participants, we tested mainly female participants. Although this difference could potentially account for the discrepancy between our EEG results, the limited available evidence suggests that, if anything, female animals and human participants show a larger response to iVNS [12,58] and taVNS [59] than males, which is inconsistent with our null finding in a group of mainly female participants (Supplementary Fig. 5).

In general, our findings highlight the importance of direct replication studies, and the need for parametric variation of taVNS parameters [49] to determine the most effective set of parameters and help reconcile mixed findings in the literature.

Data and code availability statement

Preprocessing and analyses code can be found here: https://github.com/bethlloyd/sharon_replication. Raw pupil and EEG data is available here: https://figshare.com/projects/Short-term_transcutaneous_vagus_nerve_stimulation_increases_pupil_size_but_does_not_affect_EEG_alpha_power_a_replication/169890

Funding

This work was supported by the Netherlands Organization for Scientific Research (grant No. VI·C.181.032).

CRediT authorship contribution statement

Beth Lloyd: Conceptualization, Data curation, Visualization, Formal analysis, Writing – original draft. **Franz Wurm:** Conceptualization, Data curation, Visualization, Formal analysis, Writing – review & editing. **Roy de Kleijn:** Software, Writing – review & editing. **Sander Nieuwenhuis:** Conceptualization, Supervision, Funding acquisition, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

We thank Omer Sharon for his support and transparency while we carried out this replication study. Omer generously provided scripts to validate our analyses and contributed insightful comments on methodological issues. Moreover, we thank Elio Sjak-Shie for technical support with the robot. We also thank our thesis students and interns who helped with piloting and data collection.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.brs.2023.06.010.

References

[1] Howland RH. Vagus nerve stimulation. Curr Behav Neurosci Rep 2014;1:64–73. https://doi.org/10.1007/s40473-014-0010-5.

[2] Cristancho P, Cristancho MA, Baltuch GH, Thase ME, O'Reardon JP. Effectiveness and safety of vagus nerve stimulation for severe treatment-resistant major depression in clinical practice after FDA approval: outcomes at 1 year. J Clin Psychiatry 2011;72:1376–82. https://doi.org/10.4088/JCP.09m05888blu.

[3] Morris GL, Gloss D, Buchhalter J, Mack KJ, Nickels K, Harden C. Evidence-based guideline update: vagus nerve stimulation for the treatment of epilepsy: report of the guideline development subcommittee of the american academy of neurology. Neurology 2013;81:1453–9. https://doi.org/10.1212/WNL.0b013e3182a393d1.

[4] Yuan H, Silberstein SD. Vagus nerve and vagus nerve stimulation, a comprehensive review: Part i. Headache 2016;56:71–8. https://doi.org/10.1111/head.12647.

[5] Aston-Jones GS. Chapter 4 - afferent regulation of locus coeruleus neurons: anatomy, physiology and pharmacology. Prog Brain Res 1991;88:47–75.

[6] Collins L, Boddington L, Steffan PJ, McCormick D. Vagus nerve stimulation induces widespread cortical and behavioral activation. Curr Biol 2021;31:2088–2098.e3. https://doi.org/10.1016/j.cub.2021.02.049.

[7] Dorr AE, Debonnel G. Effect of vagus nerve stimulation on serotonergic and noradrenergic transmission. J Pharmacol Exp Therapeut 2006;318:890–8. https://doi.org/10.1124/jpet.106.104166.

[8] Hassert DL, Miyashita T, Williams CL. The effects of peripheral vagal nerve stimulation at a memory-modulating intensity on norepinephrine output in the basolateral Amygdala. Behav Neurosci 2004;118:79–88. https://doi.org/10.1037/0735-7044.118.1.79.

[9] Hulsey DR, Riley JR, Loerwald KW, Rennaker RL, Kilgard MP, Hays SA. Parametric characterization of neural activity in the locus coeruleus in response to vagus nerve stimulation. Exp Neurol 2017;289:21–30. https://doi.org/10.1016/j.expneurol.2016.12.005.

[10] Raedt R, Clinckers R, Mollet L, Vonck K, El Tahry R, Wyckhuys T, et al. Increased hippocampal noradrenaline is a biomarker for efficacy of vagus nerve stimulation in a limbic seizure model. J Neurochem 2011;117:461–9. https://doi.org/10.1111/j.1471-4159.2011.07214.x.

[11] Mridha Z, de Gee JW, Shi Y, Alkashgari R, Williams J, Suminski A, et al. Graded recruitment of pupil-linked neuromodulation by parametric stimulation of the vagus nerve. Nat Commun 2021;12:1–30. https://doi.org/10.1038/s41467-021-21730-2.

[12] Farmer AD, Strzelczyk A, Finisguerra A, Gourine AV, Gharabaghi A, Hasan A, et al. International consensus based review and recommendations for minimum reporting standards in research on transcutaneous vagus nerve stimulation (version 2020). Front Hum Neurosci 2021;14. https://doi.org/10.3389/fnhum.2020.568051.

[13] Burger AM, Van der Does W, Brosschot JF, Verkuil B. From ear to eye? No effect of transcutaneous vagus nerve stimulation on human pupil dilation: a report of three studies. Biol Psychol 2020;152:1–39. https://doi.org/10.1016/j.biopsycho.2020.107863.

[14] De Taeye L, Vonck K, van Bochove M, Boon P, Van Roost D, Mollet L, et al. The P3 event-related potential is a biomarker for the efficacy of vagus nerve stimulation in patients with epilepsy. Neurotherapeutics 2014;11:612–22. https://doi.org/10.1007/s13311-014-0272-3.

[15] Neuhaus AH, Luborzewski A, Rentzsch J, Brakemeier EL, Opgen-Rhein C, Gallinat J, et al. P300 is enhanced in responders to vagus nerve stimulation for treatment of major depressive disorder. J Affect Disord 2007;100:123–8. https://doi.org/10.1016/j.jad.2006.10.005.

[16] Nieuwenhuis S, Aston-Jones G, Cohen JD. Decision making, the P3, and the locus coeruleus-norepinephrine system. Psychol Bull 2005;131:510–32. https://doi.org/10.1037/0033-2909.131.4.510.

[17] Vazey EM, Moorman DE, Aston-Jones G. Phasic locus coeruleus activity regulates cortical encoding of salience information. Proc Natl Acad Sci U S A 2018;115:E9439–48. https://doi.org/10.1073/pnas.1803716115.

[18] Gadeyne S, Mertens A, Carrette E, Van den Bossche F, Boon P, Raedt R, et al. Transcutaneous auricular vagus nerve stimulation cannot modulate the P3b event-related potential in healthy volunteers. Clin Neurophysiol 2022;135:22–9. https://doi.org/10.1016/j.clinph.2021.11.079.

[19] Lewine JD, Paulson K, Bangera N, Simon BJ. Exploration of the impact of brief noninvasive vagal nerve stimulation on EEG and event-related potentials. Neuromodulation 2019;22:564–72. https://doi.org/10.1111/ner.12864.

[20] Ventura-Bort C, Wirkner J, Genheimer H, Wendt J, Hamm AO, Weymar M. Effects of transcutaneous vagus nerve stimulation (tVNS) on the P300 and alpha-amylase level: a pilot study. Front Hum Neurosci 2018;12:1–12. https://doi.org/10.3389/fnhum.2018.00202.

[21] Warren CM, Tona KD, Ouwerkerk L, van Paridon J, Poletiek F, van Steenbergen H, et al. The neuromodulatory and hormonal effects of transcutaneous vagus nerve stimulation as evidenced by salivary alpha amylase, salivary cortisol, pupil diameter, and the P3 event-related potential. Brain Stimul 2019;12:635–42. https://doi.org/10.1016/j.brs.2018.12.224.

[22] Gurtubay IG, Perez-Rodriguez DR, Fernandez E, Librero-Lopez J, Calvo D, Bermejo P, et al. Immediate effects and duration of a short and single application of transcutaneous auricular vagus nerve stimulation on P300 event related potential. Front Neurosci 2023:17.

[23] Nater UM, Rohleder N. Salivary alpha-amylase as a non-invasive biomarker for the sympathetic nervous system: current state of research. Psychoneuroendocrinology 2009;34:486–96. https://doi.org/10.1016/j.psyneuen.2009.01.014.

[24] Warren CM, van den Brink RL, Nieuwenhuis S, Bosch JA. Norepinephrine transporter blocker atomoxetine increases salivary alpha amylase. Psychoneuroendocrinology 2017;78:233–6. https://doi.org/10.1016/j.psyneuen.2017.01.029.

[25] Giraudier M, Ventura-Bort C, Burger AM, Claes N, D'Agostini M, Fischer R, et al. Evidence for a modulating effect of transcutaneous auricular vagus nerve stimulation (taVNS) on salivary alpha-amylase as indirect noradrenergic marker: a pooled mega-analysis. Brain Stimul 2022;15:1378–88. https://doi.org/10.1016/j.brs.2022.09.009.

[26] Joshi S, Gold JI. Pupil size as a window on neural substrates of cognition. Trends Cognit Sci 2020. https://doi.org/10.1016/j.tics.2020.03.005.

[27] Borges U, Pfannenstiel M, Tsukahara J, Laborde S, Klatt S, Raab M. Transcutaneous vagus nerve stimulation via tragus or cymba conchae: are its psychophysiological effects dependent on the stimulation area? Int J Psychophysiol 2021;161:64–75. https://doi.org/10.1016/j.ijpsycho.2021.01.003.

[28] D'Agostini M, Burger AM, Franssen M, Claes N, Weymar M, von Leupoldt A, et al. Effects of transcutaneous auricular vagus nerve stimulation on reversal learning, tonic pupil size, salivary alpha-amylase, and cortisol. Psychophysiology 2021;58:1–20. https://doi.org/10.1111/psyp.13885.

[29] D'Agostini M, Burger AM, Villca Ponce G, Claes S, von Leupoldt A, Van Diest I. No evidence for a modulating effect of continuous transcutaneous auricular vagus nerve stimulation on markers of noradrenergic activity. Psychophysiology 2022;59:1–20. https://doi.org/10.1111/psyp.13984.

[30] Keute M, Demirezen M, Graf A, Mueller NG, Zaehle T. No modulation of pupil size and event-related pupil response by transcutaneous auricular vagus nerve stimulation (taVNS). Sci Rep 2019;9:1–10. https://doi.org/10.1038/s41598-019-47961-4.

[31] Sharon O, Fahoum F, Nir Y. Transcutaneous vagus nerve stimulation in humans induces pupil dilation and attenuates alpha oscillations. J Neurosci 2021;41:320–30. https://doi.org/10.1523/JNEUROSCI.1361-20.2020.

[32] Dahl MJ, Mather M, Werkle-Bergner M. Noradrenergic modulation of rhythmic neural activity shapes selective attention. Trends Cognit Sci 2022;26:38–52. https://doi.org/10.1016/j.tics.2021.10.009.

[33] Thompson SL, O'Leary GH, Austelle CW, Gruber E, Kahn AT, Manett AJ, et al. A review of parameter settings for invasive and non-invasive vagus nerve stimulation (VNS) applied in neurological and psychiatric disorders. Front Neurosci 2021;15.

[34] Peuker ET, Filler TJ. The nerve supply of the human auricle. Clin Anat 2002;15: 35–7. https://doi.org/10.1002/ca.1089.

[35] Krause F, Lindemann O. Expyriment: a Python library for cognitive and neuroscientific experiments. Behav Res Methods 2014;46:416–28. https://doi.org/10.3758/s13428-013-0390-6.

[36] Yap JYY, Keatch C, Lambert E, Woods W, Stoddart PR, Kameneva T. Critical review of transcutaneous vagus nerve stimulation : challenges for translation to clinical practice, vol. 14; 2020. https://doi.org/10.3389/fnins.2020.00284.

[37] Mathôt S, Fabius J, Van Heusden E, Van der Stigchel S. Safe and sensible preprocessing and baseline correction of pupil-size data. Behav Res Methods 2018; 50:94–106. https://doi.org/10.3758/s13428-017-1007-2.

[38] Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. 2004.

[39] Oostenveld R, Fries P, Maris E, Schoffelen JM. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput Intell Neurosci 2011;2011. https://doi.org/10.1155/2011/156869.

[40] Harshman Ra. Foundations of the PARAFAC procedure: models and conditions for an "explanatory" multimodal factor analysis. UCLA Work Pap Phonetics 1970;16: 1–84.

[41] Andersson CA, Bro R. The N-way toolbox for MATLAB. Chemometr Intell Lab Syst 2000;52:1–4. https://doi.org/10.1016/S0169-7439(00)00071-X.

[42] Bro R, Kiers HAL. A new efficient method for determining the number of components in PARAFAC models. J Chemom 2003;17:274–86. https://doi.org/10.1002/cem.801.

[43] JASP Team. JASP (Version 0.17.2). 2023 [Computer software], https://jasp-stats.org/faq/how-do-i-cite-jasp/ [Computer software].

[44] Joshi S, Li Y, Kalwani RM, Gold JI. Relationships between pupil diameter and neuronal activity in the locus coeruleus, colliculi, and cingulate cortex. Neuron 2016. https://doi.org/10.1016/j.neuron.2015.11.028.

[45] Zhu S, Qing Y, Zhang Y, Zhang X, Ding F, Zhang R, et al. Transcutaneous auricular vagus nerve stimulation increases eye-gaze on salient facial features and oxytocin release. Psychophysiology 2022;59:e14107. https://doi.org/10.1111/psyp.14107.

[46] Urbin MA, Lafe CW, Simpson TW, Wittenberg GF, Chandrasekaran B, Weber DJ. Electrical stimulation of the external ear acutely activates noradrenergic mechanisms in humans. Brain Stimul 2021;14:990–1001. https://doi.org/10.1016/j.brs.2021.06.002.

[47] Yakunina N, Kim SS, Nam EC. Optimization of transcutaneous vagus nerve stimulation using functional MRI. Neuromodulation 2017;20:290–300. https://doi.org/10.1111/ner.12541.

[48] Yoo PB, Lubock NB, Hincapie JG, Ruble SB, Hamann JJ, Grill WM. High-resolution measurement of electrically-evoked vagus nerve activity in the anesthetized dog. J Neural Eng 2013;10. https://doi.org/10.1088/1741-2560/10/2/026003.

[49] D'Agostini M, Burger AM, Franssen M, Perkovic A, Claes S, von Leupoldt A, et al. Short bursts of transcutaneous auricular vagus nerve stimulation enhance evoked pupil dilation as a function of stimulation parameters. Cortex J Devoted Study Nerv Syst Behav 2023;159:233–53. https://doi.org/10.1016/j.cortex.2022.11.012.

[50] Gusso M de M, Serur G, Nohama P. Pupil reactions to tactile stimulation: a systematic review. Front Neurosci 2021;15.

[51] Nieuwenhuis S, De Geus EJ, Aston-Jones G. The anatomical and functional relationship between the P3 and autonomic components of the orienting response. Psychophysiology 2011;48:162–75. https://doi.org/10.1111/j.1469-8986.2010.01057.x.

[52] Burger AM, D'Agostini M, Verkuil B, Van Diest I. Moving beyond belief: a narrative review of potential biomarkers for transcutaneous vagus nerve stimulation. Psychophysiology 2020;57:1–24. https://doi.org/10.1111/psyp.13571.

[53] Kavakbasi E, Gross J, Wollbrink A, Fellmeth R, Baune BT. Acute effect of vagus nerve stimulation (VNS) on brain function. J Psychiatr Res 2021;141:136–9. https://doi.org/10.1016/j.jpsychires.2021.06.042.

[54] Ricci L, Croce P, Lanzone J, Boscarino M. Brain sciences transcutaneous vagus nerve stimulation modulates EEG microstates and delta activity in. Brain Sci 2020; 10:668.

[55] Konjusha A, Colzato L, Mückschel M, Beste C. Auricular transcutaneous vagus nerve stimulation diminishes alpha-band – related inhibitory gating processes during conflict monitoring. Frontal Cortices 2022;25:457–67.

[56] Chen Y, Lu X, Hu L. Transcutaneous auricular vagus nerve stimulation facilitates cortical Arousal and alertness. Int J Environ Res Publ Health 2023;20:1402. https://doi.org/10.3390/ijerph20021402.

[57] Forstmeier W, Wagenmakers E-J, Parker TH. Detecting and avoiding likely false-positive findings - a practical guide. Biol Rev Camb Phil Soc 2017;92:1941–68.

[58] Yokota H, Edama M, Hirabayashi R, Sekine C, Otsuru N, Saito K, et al. Effects of stimulus frequency, intensity, and sex on the autonomic response to transcutaneous vagus nerve stimulation. Brain Sci 2022;12:1038. https://doi.org/10.3390/brainsci12081038.

[59] De Couck M, Cserjesi R, Caers R, Zijlstra WP, Widjaja D, Wolf N, et al. Effects of short and prolonged transcutaneous vagus nerve stimulation on heart rate variability in healthy subjects. Auton Neurosci 2017;203:88–96. https://doi.org/10.1016/j.autneu.2016.11.003.

[60] Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Royal Stat Soc: Ser B (Methodological) 1995;57(1):289–300.