Effects of tDCS on the attentional blink revisited: A statistical evaluation of a replication attempt

Leon C. Reteig[1,2], Lionel A. Newman[1,3], K. Richard Ridderinkhof[1,2], Heleen A. Slagter[1,4*]

[1] Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands

[2] Amsterdam Brain and Cognition, University of Amsterdam, Amsterdam, The Netherlands

[3] Department of Artificial Intelligence and Cognitive Engineering, University of Groningen, Groningen, The Netherlands

[4]Department of Applied and Experimental Psychology, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

* Corresponding author

E-mail: h.a.slagter@vu.nl

SHORT TITLE: Replicating tDCS effects on the attentional blink

Abstract

The attentional blink (AB) phenomenon reveals a bottleneck of human information processing: the second of two targets is often missed when they are presented in rapid succession among distractors. In our previous work, we showed that the size of the AB can be changed by applying transcranial direct current stimulation (tDCS) over the left dorsolateral prefrontal cortex (lDLPFC) (London & Slagter, *Journal of Cognitive Neuroscience*, *33*, 756–68, 2021). Although AB size at the group level remained unchanged, the effects of anodal and cathodal tDCS were negatively correlated: if a given individual's AB size decreased from baseline during anodal tDCS, their AB size would increase during cathodal tDCS, and vice versa. Here, we attempted to replicate this finding. We found no group effects of tDCS, as in the original study, but we no longer found a significant negative correlation. We present a series of statistical measures of replication success, all of which confirm that both studies are not in agreement. First, the correlation here is significantly smaller than a conservative estimate of the original correlation. Second, the difference between the correlations is greater than expected due to sampling error, and our data are more consistent with a zero-effect than with the original estimate. Finally, the overall effect when combining both studies is small and not significant. Our findings thus indicate that the effects of lDPLFC-tDCS on the AB are less substantial than observed in our initial study. Although this should be quite a common scenario, null findings can be difficult to interpret and are still under-represented in the brain stimulation and cognitive neuroscience literatures. An important auxiliary goal of this paper is therefore to provide a tutorial for other researchers, to maximize the evidential value from null findings.

Introduction

The *attentional blink* (AB) phenomenon clearly demonstrates that our capacity to process incoming information is easily overwhelmed. The AB occurs when two targets are embedded in a rapidly presented stream of distractors [1] (for reviews, see [2,3]). The first target (T1) is usually reported with little difficulty. When there is a longer lag between the two targets (> 500 ms; [4]), accuracy for the second target (T2) can be on par with the first. However, for shorter lags, T2 is most often missed—as if the attentional system momentarily faltered ("blinked").

While the AB might seem to be a fundamental bottleneck, it can under some circumstances be overcome. For example, the size of the AB can be reduced by distracting activities [5–7], or after following an intensive mental training program [8]. Others have tried to use non-invasive brain stimulation [9] as a means to influence the AB. Several studies have shown that repetitive transcranial magnetic stimulation (TMS) can improve target perception in AB tasks [10–12]. Yet, as TMS did not show a differential effect for targets presented at shorter or longer lags, it did not affect the AB itself.

Transcranial direct current stimulation (tDCS) is another brain stimulation technique that has gained traction in the past two decades. In tDCS, an electrical current flows between an anodal and cathodal electrode, which can affect the excitability of the underlying cortex [13]. Anodal stimulation generally enhances cortical excitability, while cathodal stimulation may have an inhibitory effect [14,15] (though note that this does not hold in all cases [16], and the underlying physiology is complex [17–19]).

A previous study by our group [20] recently examined the effects of tDCS on the AB. In this study, anodal and cathodal tDCS were applied over the left dorsolateral prefontal

cortex (lDPLFC, with the other electrode on the right forehead)—one of the core brain areas implicated in the AB [21,22]. At the group level, tDCS did not appear to have any effects on AB size. However, anodal and cathodal tDCS did appear to systematically affect the AB within individuals, as their effects were negatively correlated. For a given individual, this negative correlation implies that when AB size decreased (compared to a baseline measurement) during anodal tDCS, AB size would increase during cathodal tDCS (or vice versa).

Our previous findings [20] complement earlier literature showing large individual differences in both the AB [23] and effects of tDCS [24]. However, only one other tDCS study of the AB exists to date, which in fact did find a group-level effect of stimulation, in contrast to [20] ([25]). Also, the negative correlation between the effects of anodal and cathodal tDCS has so far only been reported once [20], and was based on an exploratory analysis. We thus decided to conduct another study aiming to replicate this finding. Replication of findings in general is important to establish credibility of scientific claims [111], and in this specific case, reproduction of our original findings would be informative about the ability of tDCS to modulate the apparent bottleneck in information processing captured by the AB, and about the importance of left dorsolateral prefrontal cortex in the AB and attentional filtering more generally. Thus, we aimed to confirm our previous results and conclusions.

We also aimed to examine the neurophysiological mechanisms underlying the individual differences in response to tDCS using EEG [24,45,46]. However, since we did not replicate the behavioral effect of tDCS on the AB as in our original study [20] , we were unable to fulfill this aim.

To foreshadow our results, like in our original study [20], we did not find a group effect of tDCS on the AB. However, the correlation between the effects of anodal and cathodal tDCS was also not significant. Although this indicates that the two studies may differ, a failure to reject the null hypothesis by itself does not tell us much [26]: it is crucial to also take measures of the uncertainty and effect size in both studies into account [27]. That said, there is no consensus on when a replication can be considered successful or not [28], let alone a single statistical test to definitely answer this question conclusively [29].

Therefore, we employ a number of statistical methods to maximize the evidential value in these two studies. We ask three questions (cf. [30]) that all aim to evaluate to what extent the present study is a successful replication [31] (cf. [32,33]) of our previous work [20]. First, while the effect in our study was not significant, there might still be a meaningful effect that is simply smaller than anticipated. Therefore, we used equivalence testing [34] to answer the question "*is the correlation in study 2 significantly small?*" Second, although our current result differs from, it could still be more consistent with our previous findings [20] than with alternative explanations. So in addition, we asked "*is the correlation in study 2 different from study 1?*" and aimed to answer this question using prediction intervals [35,36] and replication Bayes factors [37,38]. Finally, it could be that the effect in our study alone is not sufficiently large, but the overall effect based on both studies is. This raises the question "*is the effect significant when combining study 1 and 2?*" which we addressed through meta-analysis [39,40] and by pooling the data.

These questions address issues of reproducibility that are currently faced by many in the brain stimulation field [41], and in the cognitive neuroscience community at large [42,43]. Therefore, aside from our focus on tDCS and the AB, an important auxiliary goal of this article is to provide a tutorial on the statistical evaluation of replication studies

(also see [44], [30]). We hope this may prove useful to other researchers who find themselves in similar situations.

Methods

Participants

Fourty-eight participants took part in total, eight of whom were excluded after the first session. One participant was excluded as a precaution because they developed an atypical headache after the first session, and we could not rule out this was related to the tDCS. Another stopped responding to our requests to schedule the second session. Another six participants were excluded because their mean T1 accuracy in the first session was too low, which implies they were not engaged in the task in the manner we expected, and would leave fewer trials to analyze, because our T2 accuracy measure included only trials in which T1 was seen. We used a cut-off of 63% T1 accuracy as an exclusion criterium, which was two standard deviations below the mean of a separate pilot study (n = 10).

This left a final sample of 40 participants (29 female, mean age = 20.94, *SD* = 2.45, range = 18–28). This sample size was determined a priori to slightly exceed the original study [20] (n = 34).

The experiment and recruitment took place at the University of Amsterdam. All procedures for this study were approved by the ethics review board of the Faculty for Social and Behavioral Sciences, and complied with relevant laws and institutional guidelines. All participants provided their written informed consent and were compensated with course credit or €10 per hour (typically €65 for completing two full sessions).

Procedure

The study procedures were identical to our previous study [20]: participants received anodal and cathodal tDCS in separate sessions (Fig 1), which typically took place exactly one week apart (cf. minimum of 48 hours previously [20]). The time in between served to keep the sessions as similar as possible, and to minimize the risk of tDCS carry-over effects. 18 participants received anodal tDCS in the first session and cathodal tDCS in the second, and vice versa for the remaining 22 participants.

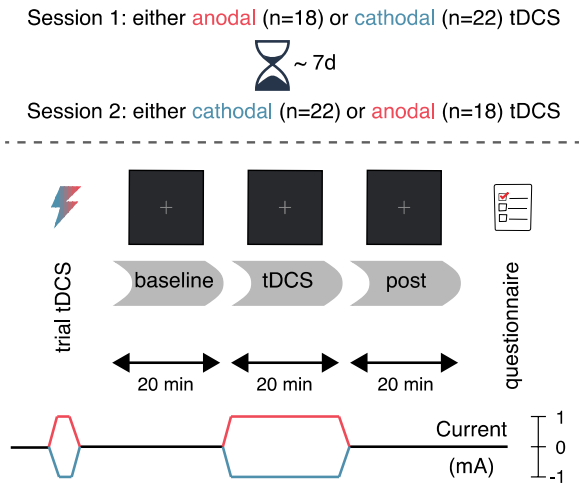


Fig 1. Experimental design. After a baseline block without stimulation, participants performed the attentional blink task during 20 minutes of anodal (red) or cathodal (blue) tDCS, followed by a post-test block (also without stimulation). The second session (typically 7 days later) was identical, except that the tDCS polarity was reversed.

First, participants experienced the sensations induced by tDCS in a brief trial stimulation (see the tDCS section). Next, participants completed 20 practice trials of the task (see the Task section). For the main portion of the experiment, participants performed three blocks of the task (Fig 1): before tDCS (*baseline*), during anodal/cathodal tDCS (*tDCS*),

and after tDCS (*post*). Finally, after completing the three blocks, participants filled in a questionnaire on tDCS-related adverse effects (see S1 Table and S1 Fig).

Within each block of the task, participants took a self-timed break every 50 trials (~5 minutes); between the blocks, the experimenter walked in. Participants performed the task for exactly 20 minutes during the *baseline* and *post* blocks. During the *tDCS* block, the task started after the 1-minute ramp-up of the current was complete, and continued for 21 minutes (constant current, plus 1-minute of ramp-down).

Task

The attentional blink task (Fig 2) was almost identical to the one we used previously [20, 47] and (see S2 Table for a full list of the differences), which in turn was based on an earlier task [48]. A rapid serial visual presentation stream of 15 letters was shown on each trial, using Presentation software (Neurobehavioral Systems, Inc.). Each letter was displayed for 91.7 ms (11 frames at 120 Hz) on a dark gray background. The letters were presented in font size 40 (font: Courier New) at a viewing distance of 90 cm. On each trial, the letters were randomly sampled without replacement from the alphabet, excluding the letters I, L, O, Q, U and V, as they were too similar to each other. All distractor letters were mid-gray, whereas T1 and T2 were colored. T1 was red and always appeared at position 5 in the stream. T2 was green and followed T1 after either 2 distractors (*lag 3*) or 7 distractors (*lag 8*) (cf. lags 2, 4 and 10 in the original study [20]).

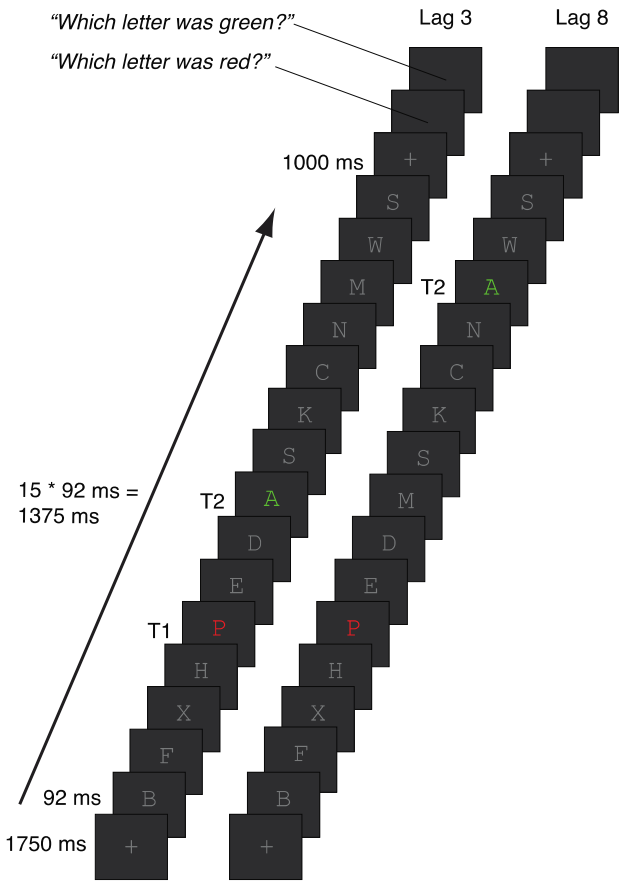


Fig 2. Attentional blink task. Participants viewed rapid serial visual presentation streams of 15 letters, all of which were distractors (gray letters) except for T1 and T2. T1 was presented in red at position 5; T2 was presented in green and followed T1 after 2 distractors (lag 3, inside the AB window) or 7 distractors (lag 8, outside the AB window). At the end of the trial, participants reported the identity of T1 and then T2 (self-paced).

The letter stream was preceded by a fixation cross (same color as the letters) presented for 1750 ms and followed by another fixation cross. Finally, the participant was prompted to type in (using a standard keyboard) the letter they thought was presented as T1 ("Which letter was red?"), followed by T2 ("Which letter was green?").

Trial duration varied slightly because both the T1 and T2 response questions were self-paced, so some participants completed more trials than others depending on their

response times. On average, participants completed 130 short lag trials (*SD* = 17; range = 78–163) and 65 long lag trials (*SD* = 9; range = 39–87) per 20-minute block.

tDCS

Transcranial direct current stimulation was delivered online (i.e. during performance of the attentional blink task) using a DC-STIMULATOR PLUS (NeuroCare Group GmbH). The current was ramped up to 1 mA in 1 minute, stayed at 1 mA for 20 minutes, and was ramped down again in 1 minute.

One electrode was placed at F3 (international 10-20 system) to target the lDLPFC; the other was placed over the right forehead, centered above the eye (approximately corresponding to position Fp2). Both electrodes were 5 x 7 cm in size (35 cm^2), leading to a current density of 0.029 mA/cm^2. The montage and tDCS parameters are identical to our previous study [20], the only exception being the conductive medium. We used Ten20 conductive paste (Weaver and Company), because it was easier to apply concurrently with the EEG equipment (see the EEG section); originally [20,] saline solution was used as a conductive medium, together with rubber straps to keep the electrodes in place.

Participants received either anodal tDCS (anode on F3, cathode on right forehead) or cathodal tDCS (cathode on F3, anode on right forehead) in separate sessions. The procedure was double-blinded: both the participant and the experimenters were unaware which polarity was applied in a given session. The experimenter loaded a stimulation setting on the tDCS device (programmed by someone not involved in data collection), without knowing whether it was mapped to anodal or cathodal tDCS. In the second session, the experimenter loaded a second setting mapped to the opposite

polarity (half the dataset), or simply connected the terminals of the device to the electrodes in the opposite way.

At the start of the experiment, participants received a brief trial stimulation, based on which they decided whether or not they wanted to continue with the rest of the session. The experimenter offered to terminate the experiment in case tDCS was experienced as too uncomfortable, but none of the participants opted to do so. For the trial stimulation, the current was ramped up to 1 mA in 45 seconds, stayed at 1 mA for 15 seconds, and was ramped down again in 45 seconds.

EEG

We also recorded EEG during all three task blocks. As noted in the Introduction, we refrained from further analysis of the EEG data, given the absence of a behavioral effect. Instead, we are making the EEG data publicly available at this time on the OpenNeuro platform: https://doi.org/10.18112/openneuro.ds001810.v1.1.0. The dataset is formatted according to the Brain Imaging Data Structure (BIDS) standard [49] for EEG [50], to facilitate re-use. We include the raw data, as well as the fully preprocessed data and the MATLAB code that generated it.

Data analysis software

Data were analyzed using R [Version 3.6.2; 51] from within RStudio [Version 1.1.463; 63]. In addition to the packages cited in the follow sections, we used the following R-packages: *BayesFactor* [Version 0.9.12.4.2; 52], *broom* [Version 0.7.4.9000; 53], *cowplot* [Version 1.1.1; 54], *emmeans* [Version 1.5.4; 55], *here* [Version 1.0.1; 56], *kableExtra* [Version 1.3.2; 57], *knitr* [Version 1.31; 58], *papaja* [Version 0.1.0.9997; 59], *psychometric* [Version 2.2; 60], *pwr* [Version 1.3.0; 61], and *tidyverse* [Version 1.3.0; 62].

Group-level analysis

Repeated measures ANOVAs were conducted on T1 accuracy (percentage of trials where T1 was reported correctly) and T2|T1 accuracy (percentage of trials where T2 was reported correctly, out of the subset of T1-correct trials) using the *afex* package [Version 0.28-1; 64]. The same factors were included for both repeated measures ANOVAs, (following our previous approach [20]): Lag (3, 8), Block (baseline, tDCS, post), Stimulation (anodal, cathodal), and the between-subject factor Session order (anodal tDCS in the first session vs. cathodal tDCS in the first session). Effect sizes are reported as generalized eta-squared ($\hat{\eta}_G^2$) [65]. Greenhouse-Geisser-adjusted degrees of freedom (df^{GG}) and p-values are reported as a correction for sphericity violations.

Individual differences analysis

We reproduced the analysis behind Figure 4 of London & Slagter [20], which showed a differential effect of anodal vs. cathodal tDCS at the individual participant level. First, we calculated AB magnitude by subtracting T2|T1 accuracy at lag 3 from T2|T1 accuracy at lag 8. Next, change scores were created by subtracting AB magnitude in the *baseline* block from the *tDCS* and the *post* blocks, respectively. The change scores in the anodal and cathodal session were then correlated to each other. Again following our previous approach [20], we computed a partial correlation (using the *ggm* package [Version 2.5; [66]]), attempting to adjust for variance due to Session order.

Replication analyses

In contrast to [20], the analysis described in the previous section did not produce a significant correlation in our dataset. Therefore, we conducted five follow-up analyses

that aim to quantify to what extent our results (do not) replicate [20]. These all provide a complementary perspective on this question. First, we performed an equivalence test (1) to assess whether the effect in the present study was significantly smaller than in our previous study [20]. While this procedure is more focused on hypothesis testing, we also constructed prediction intervals (2) to capture the range of effect sizes we can expect in a replication of our previous study [20]. Both of these procedures are based on frequentist statistics, which cannot directly quantify the amount of evidence for a (null) hypothesis. Therefore, we also computed a replication Bayes factor that expresses whether the data in the present study are more likely under the null hypothesis that the effect is absent, vs. the alternative hypothesis that the effect is comparable to the previous estimate [20]. Finally, we directly combined both studies and estimated the size of the overall effect, through meta-analysis (4) of both correlations, and by computing a new correlation on the pooled dataset (5). More details on each analysis can be found in the following sections, and the provided online resources (see the Data, materials, and code availability section).

Equivalence tests

Equivalence tests can be used to test for the *absence* of an effect of a specific size (see [34] for a tutorial]. Usually, the effect size used for the test is the smallest effect size of interest (the SESOI). Typically, equivalence tests are two one-sided tests: one test of the null hypothesis that the effect exceeds the upper equivalence bound (positive SESOI), and one that the effect exceeds the lower equivalence bound (negative SESOI). However, a one-sided test is more appropriate here: previously we found [20] that the effects of anodal and cathodal tDCS were anticorrelated, so we are only interested in negative effect sizes. This is known as an inferiority test [34].

We follow the "small telescopes" [27] approach to set the SESOI to $r_{33\%}$: the correlation that the original study [20] had 33% power to detect. The reasoning behind this approach is that it is difficult to prove that an effect does not exist at all, but easier to show that it is surprisingly small. An equivalence test can suggest that the effect is unlikely to exceed $r_{33\%}$, such that the odds to detect it were stacked at least 2:1 against the original study [20]. That would not mean the effect does not exist at all, but it would mean the original evidence for the effect is not very convincing, as "too small a telescope" (in this case, an inadequate sample size) was used to reliably detect it.

There are many possible specifications of the SESOI, none of which are necessarily wrong or right [34]. We favored the "small telescopes" approach because it constitutes a relatively strict test—$r_{33\%}$ is much smaller than the original correlation [20]. Because the originally observed correlation [20] could have overestimated the true correlation, it is prudent to set the SESOI to be smaller. Furthermore, the approach was specifically designed to evaluate replication results [27], and has been used previously in large-scale replication studies [e.g. 32]. However, recent simulation studies have also demonstrated that equivalence tests for replications can be either too liberal (especially in the case of publication bias) [67], or too conservative (especially in case of low sample sizes) [68].

We conducted an inferiority test using the *TOSTER* package [Version 0.3.4; 69] against the null hypothesis that the correlation coefficient in the present study is at least as negative as $-r_{33\%}$. At a standard alpha level of 0.05, the test is significant if the 90% confidence interval around the observed correlation does not contain $r_{33\%}$. This would mean that the observed correlation should be considered "statistically inferior," as it is then significantly smaller (i.e. less negative) than $-r_{33\%}$.

Prediction interval

Prediction intervals contain a range of values we can expect a new observation to fall within. In our case, the observation of interest is the correlation between the effects of anodal and cathodal tDCS. This correlation is estimated based on a sample, and is thus subject to sampling error: any two estimates of the correlation will almost never be exactly the same. Prediction intervals aim to quantify how dissimilar two estimates can be before we should be surprised.

Here, we construct a prediction interval around the original estimate of the correlation [20]. This prediction interval contains the range of correlation coefficients we can expect to see in the present study, given the results of the original study [20]. The width of the interval depends on the sample sizes of both studies, as larger samples will reduce variation in the estimates, leading to smaller prediction intervals [36].

If the original study were replicated 100 times, 95 of the observed correlation coefficients would fall within the 95% prediction interval [36]. Note that this definition is related to, but different from, a *confidence interval*, which quantifies uncertainty about the (unknown) true correlation in the population (95 out of every hundred constructed 95% confidence intervals contain the true population parameter). Because prediction intervals concern the next single observation, they make a more specific claim, and will be wider than confidence intervals.

We calculated a 95% prediction interval for correlations, following [35], using the *predictionInterval* package [Version 1.0.0; 70].

Replication Bayes factor

Bayes factors can be used to express the relative evidence for the null (H_0) or alternative hypothesis (H_1) [71]. In a default Bayesian hypothesis test, H_0 states the effect size is absent (i.e. exactly zero); H_1 states that the effect is present (specified further by a prior distribution of effect sizes).

In a replication context, we want to decide between two different scenarios [38]. H_0 is the hypothesis of an idealized skeptic, who disregards the information from the original study and believes the effect is absent. The alternative hypothesis H_r belongs to an idealized proponent, who believes that the effect is exactly as in the original study, i.e. their prior distribution is simply the posterior distribution of the original study.

We used a replication Bayes factor test for correlations [37]. The replication Bayes factor BF_{0r} expresses evidence for H_0 : "the correlation is 0" relative to H_r : "the correlation is as in the original study." We use the following interpretation scheme [72], where $1 < BF_{0r} < 3$ consitutes "anecdotal evidence" for H_0, $3 < BF_{0r} < 10 \sim$ "moderate evidence," and $10 < BF_{0r} < 30 \sim$ "strong evidence."

Meta-analysis

The outcomes from multiple studies on the same phenomenon can be combined through meta-analysis. Here we compute a meta-analytic estimate of the correlation based on the correlations observed here and as observed originally [20], using the *metafor* package [Version 2.4-0; [73]]. We weighed the estimate by sample size, as larger studies allow for more precise effect size estimates, and thus should also contribute more to the overall meta-analytic estimate. In our case, this means the present study will have a slightly higher influence on the meta-analytic effect size (because its sample size exceeds our

previous study [20]). We specified the meta-analysis as a fixed-effects model, because both studies are highly similar and from the same population (e.g., the experiments were conducted in the same location, and the sample was from the same university student population). With a fixed-effects analysis, we estimate the size of the effect *in the set of available studies*, meaning our inferences cannot generalize beyond. A random-effects meta-analysis would be appropriate in case the studies were more dissimilar, and if we sought an estimate of *the true effect in the population*, but we would need more than just two studies for this approach. When more studies are available, and it is not immediately clear whether fixed- or random-effects meta-analysis is more appropriate, it is advisable to report both [40]. Note that while meta-analyses are a powerful way to assess the overall effect in a series of studies, they are particularly vulnerable to false positives when there is publication bias [67], or if the selection of studies (or any single study) is biased [74].

Pooling the data

Another approach is to pool the single-subject data from both studies, and to re-calculate the partial correlation on the combined sample (n = 74). The main difference between the two studies is that we originally [20] presented T2 at lags 2, 4 and 10; here we used lags 3 and 8. The long lags (lag 8 vs. lag 10) should be fairly comparable, as they are both well outside the attentional blink window (> 500 ms following T1; [4]). However, there should be a sizeable performance difference at the short lags (lag 2 vs. lag 3), as the attentional blink is larger at lag 2 than lag 3. Therefore, we opted to also create a "lag 3" condition in the previous data [20], by averaging T2|T1 accuracy at lag 2 and lag 4. The difference from lag 2 to 4 (and 4 to 10) in the original study [20] looks fairly linear (see Figure 3 [20]), so this seems a fair approximation of "true" lag 3 performance. Afterwards

we recomputed the partial correlation between AB magnitude change scores as described previously (see the Individual differences analysis section).

Note that this analysis is tailored to this series of studies, and not generally advisable. To get a more accurate estimate of the effect at lag 3, it is necessary to redo the analysis on the larger, combined sample. But repeating a statistical test after collecting more data ("optional stopping") invalidates the interpretation of the p-value and can drastically increase the false positive rate [75]. This would only be acceptable when the analysis plan has been preregistered, and the false positive rate of sequential analyses is controlled (for potential solutions, see [76]). We therefore do not report a p-value for this test, but only the effect size and its confidence interval.

Data, materials, and code availability

All of the data and materials from this study and the data from the original study [20] are available on the Open Science Framework: https://doi.org/10.17605/OSF.IO/Y6HSF. The analysis code is available on GitHub (https://doi.org/10.5281/zenodo.4782446; and also from our OSF page), in the form of an R notebook detailing all the analyses that we ran for this project, along with the results. We also include an Rmarkdown [77] source file for this paper that can be run to reproduce the pdf version of the text, along with all the figures and statistics.

Results

Group-level analysis

Fig 3 shows the attentional blink (T2|T1 accuracy per lag) for each of the three blocks (pre, tDCS, post) and stimulation conditions separately. The summary statistics and ANOVA results for T2|T1 accuracy can be found in Tables 1 and 2. There was a clear attentional blink effect on average (main effect of Lag, $F[1,38] = 432.11$, $p < .001$), as T2|T1 accuracy for lag 8 was higher than lag 3.

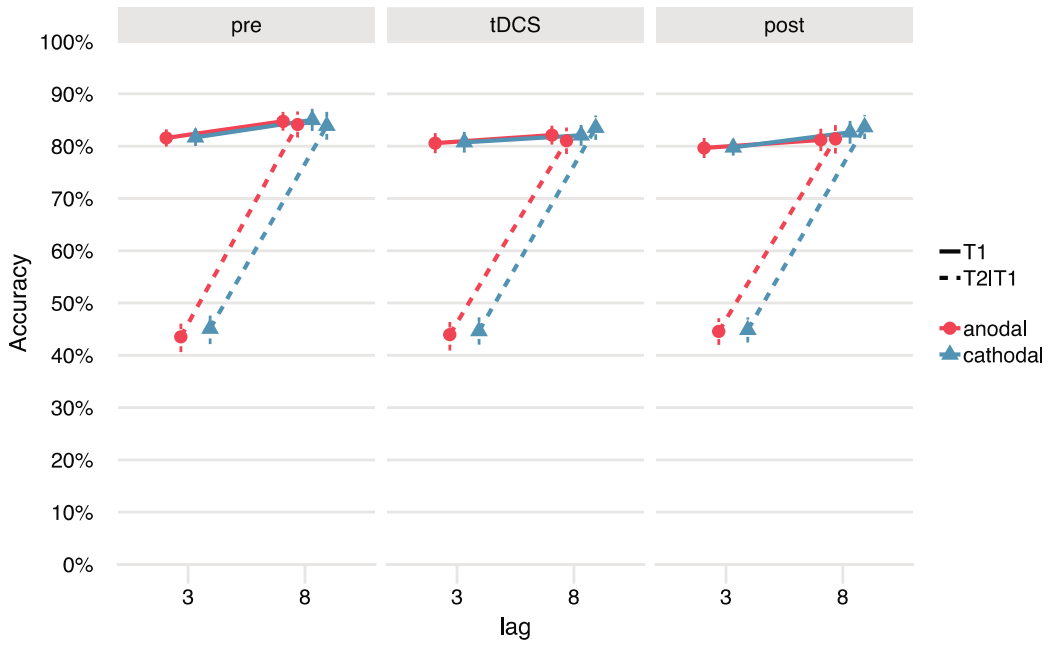


Fig 3. No effects of tDCS on the attentional blink at the group level. There was a clear attentional blink effect: a lower % T2 accuracy (given T1 correct: T2|T1; dashed lines) for lag 3 (T2 presented inside the attentional blink window) than lag 8 (T2 presented outside the attentional blink window, on par with T1 accuracy). However, the attentional blink did not change systematically over stimulation conditions (anodal, cathodal) and blocks (pre, tDCS, post). T1 accuracy (solid lines) was also not affected by tDCS.

	First session: anodal (n=18)		First session: cathodal (n=22)	
	anodal	cathodal	anodal	cathodal
baseline				
lag 3	0.395 (0.149)	0.486 (0.167)	0.469 (0.211)	0.423 (0.168)
lag 8	0.826 (0.119)	0.819 (0.095)	0.854 (0.125)	0.855 (0.087)
tDCS				
lag 3	0.415 (0.168)	0.450 (0.165)	0.460 (0.201)	0.444 (0.188)
lag 8	0.787 (0.131)	0.819 (0.098)	0.830 (0.115)	0.849 (0.096)
post				
lag 3	0.437 (0.171)	0.451 (0.160)	0.453 (0.190)	0.447 (0.164)
lag 8	0.783 (0.118)	0.825 (0.098)	0.840 (0.145)	0.846 (0.120)

Table 1. Descriptive statistics for T2|T1 accuracy.

Values are "Mean (SD)".

Effect	F	df_1^{GG}	df_2^{GG}	p	$\hat{\eta}_G^2$
Session order	0.33	1	38	.568	.006
Block	1.13	1.91	72.71	.328	.001
Stimulation	2.47	1	38	.125	.002
Lag	432.11	1	38	< .001	.634
Session order × Block	0.29	1.91	72.71	.741	.000
Session order × Stimulation	5.55	1	38	.024	.005
Session order × Lag	0.48	1	38	.494	.002
Block × Stimulation	0.28	1.91	72.51	.747	.000

Effect	F	df_1^{GG}	df_2^{GG}	p	$\hat{\eta}_G^2$
Block × Lag	1.67	1.70	64.73	.199	.001
Stimulation × Lag	0.10	1	38	.751	.000
Session order × Block × Stimulation	1.93	1.91	72.51	.154	.001
Session order × Block × Lag	0.24	1.70	64.73	.752	.000
Session order × Stimulation × Lag	5.84	1	38	.021	.002
Block × Stimulation × Lag	2.77	1.77	67.17	.076	.001
Session order × Block × Stimulation × Lag	7.25	1.77	67.17	.002	.003

Table 2. Repeated measures ANOVA on T2|T1 accuracy.

Effect	F	df_1^{GG}	df_2^{GG}	p	$\hat{\eta}_G^2$
Session order	0.96	1	38	.332	.018
Block	6.64	1.89	71.87	.003	.010
Stimulation	0.00	1	38	.996	.000
Lag	29.23	1	38	< .001	.013
Session order × Block	0.60	1.89	71.87	.540	.001
Session order × Stimulation	11.24	1	38	.002	.030
Session order × Lag	0.04	1	38	.844	.000
Block × Stimulation	0.24	1.92	73.08	.777	.000
Block × Lag	1.91	1.93	73.20	.158	.001
Stimulation × Lag	0.31	1	38	.584	.000

Effect	F	df_1^{GG}	df_2^{GG}	p	$\hat{\eta}_G^2$
Session order × Block × Stimulation	9.94	1.92	73.08	< .001	.007
Session order × Block × Lag	0.19	1.93	73.20	.821	.000
Session order × Stimulation × Lag	0.96	1	38	.333	.000
Block × Stimulation × Lag	0.31	1.86	70.75	.718	.000
Session order × Block × Stimulation × Lag	0.53	1.86	70.75	.580	.000

Table 3. Repeated measures ANOVA on T1 accuracy.

Effects of tDCS at the group-level should manifest as a three-way interaction between Block, Stimulation, and Lag. As in the original study [20], this interaction was not significant ($F[1.77,67.17] = 2.77, p = .076$). However, the higher-order interaction with Session order did reach significance ($F[1.77,67.17] = 7.25$, $p = .002$). This interaction appears to be mostly driven by a learning effect across sessions and blocks (the Session order by Stimulation by Lag interaction). From the first to the second session, lag 3 performance improved, while lag 8 performance declined somewhat (leading to a smaller attentional blink). This change was stronger in participants that received anodal tDCS in the first session, and less pronounced in the cathodal-first group. We do not consider this a genuine effect of tDCS on the attentional blink, because there is no clear reason why these randomized groups should differ, and because the largest difference between the anodal and cathodal session occurred in the baseline block already (see Table 1).

We ran the same repeated measures ANOVA for T1 accuracy (Table 3). On average, participants correctly identified T1 in 82% of trials (Fig 3), which is comparable to previous studies using this task (86% [20], 82% [47]). T1 accuracy was also slightly lower

for lag 3 than lag 8 (main effect of Lag, $F[1,38] = 29.23$, $p < .001$). There was a main effect of Block, reflecting that T1 accuracy decreased *within* a session ($F[1.89,71.87] = 6.64$, $p = .003$). Finally, T1 performance also decreased *across* the sessions (interaction between Session order and Stimulation: $F[1,38] = 11.24$, $p = .002$).

In sum, we conclude that there is no significant effect of tDCS on the attentional blink or T1 accuracy at the group level, in agreement with our previus study [20].

Individual differences

Our main aim was to replicate our original work [20], that revealed a negative correlation between AB magnitude change scores (comparing the *tDCS* and *baseline* blocks) in the anodal and cathodal sessions ($r(31) = -.45$, 95%CI[-0.68, -0.12], $p = .009$). Their interpretation was that the effects of anodal and cathodal tDCS were anti-correlated: those participants that improve their performance (i.e., show a smaller AB) due to anodal tDCS tend to worsen due to cathodal tDCS, and vice versa.

We ran the same partial correlation (attempting to adjust for Session order) between the anodal and cathodal AB magnitude change scores (*tDCS* - *baseline*) on our data (Fig 4). However, here the resulting correlation did not go in the same direction ($r(37) = .02$, 95%CI[-0.30, 0.33]), and was not significant ($p = .899$). Or, adopting the terminology suggested to describe replications [78]: we obtained *no signal* (the 95% confidence interval includes 0, and therefore the correlation is not significant), and our result is *inconsistent* (the 95% confidence interval excludes the point estimate of the original correlation [20]). In the next sections, we present a number of follow-up analyses that further explore this difference in results between both studies (see the Replication analyses subsection in the Methods section).

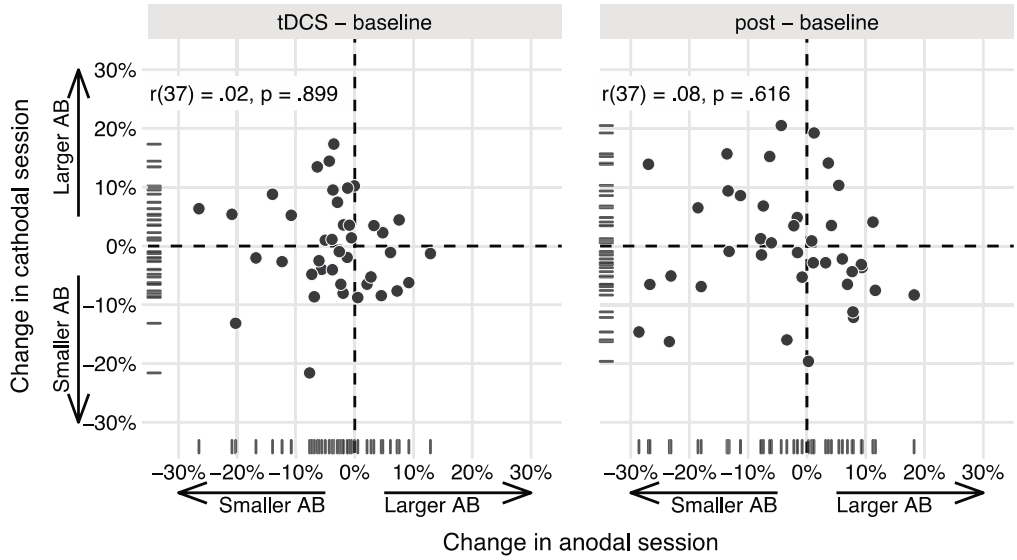


Fig 4. The effects of anodal and cathodal tDCS are not correlated in the present study. Data points show AB magnitude change scores (tDCS - baseline, post - baseline) for each participant in the study, in the anodal session (x-axis) and the cathodal session (y-axis). While we originally [20] found a negative partial correlation (for tDCS - baseline), suggesting opposite effects of anodal and cathodal tDCS, this effect appears to be absent here. The partial correlation coefficients (attempting to adjust for Session order) and p-values are printed in the upper left.

Is the correlation in study 2 significantly small?

We originally [20] found a medium- to large correlation (r = -.45), but the correlation we find here is much smaller (r = .02). We use equivalence testing ([34]; see the Equivalence tests subsection in the Methods section) to assess whether this correlation is significantly smaller than a smallest effect size of interest (SESOI). Following the "small telescopes" approach [27], we set the SESOI to $r_{33\%}$, the correlation the original study [20] had 33% power to detect. Given their sample size of 34 participants, $r_{33\%}$ = .27.

An inferiority test shows that the correlation here is significantly less negative than $-r_{33\%}$ (p = .038) (Fig 6), although only by a small margin. The effect is therefore "statistically inferior": the correlation does not exceed the lower equivalence bound ($-r_{33\%}$ = -.27). By this definition, the correlation is too small (i.e. not negative enough) to be considered meaningful, indicating that we did not successfully replicate the previous study [20].

Is the correlation in this study (study 2) different from the previous study (study 1)?

To evaluate whether the correlation in the present study (study 2) was to be expected based on study 1 [20], we constructed a 95% *prediction interval* (PI) [35], using the correlation in study 1 and the sample size of both studies (see the Prediction interval subsection of the Methods section).

The 95% PI[-0.82, -0.05] around the original correlation (r = -.45) is very wide, so almost any negative correlation would fall within it. However, it does not include the replication result (r = .02) (Fig 6). Assuming the results of both studies differ only due to sampling error, the correlation observed here has a 95% chance to fall within the interval. This means the correlation in our replication study is inconsistent with the correlation in the original study [20], and thus that the replication should not be considered successful.

Another approach to quantify the difference between study 1 and 2 is to construct a replication Bayes Factor [37]. We use this Bayes factor to assess evidence for H_0 : "the effects of anodal and cathodal tDCS are uncorrelated" relative to H_r : "the effects of anodal and cathodal tDCS are anticorrelated [20]" (see the Replication Bayes factor subsection of the Methods section).

This replication Bayes factor expresses that the data are BF_{0r} = 9.66 times more likely under H_0 than under H_r. This constitutes moderate to strong evidence that the effect is absent vs. consistent with the original [20], and thus that the effect did not replicate.

Is the effect significant when combining study 1 and 2?

The sample size in both study 1 (n = 34) and 2 (n = 40) is on the lower end. Therefore, we might get a more accurate estimate of the size of the effect when combining both studies.

The meta-analytic estimate of the correlation is r = -0.20, 95%CI[-0.42, 0.04], p = .094 (Fig 6). So, when combining the original correlation from [20] and the correlation observed here, the overall effect is no longer significant.

In addition, we pooled the data from both studies at the single-subject level and re-calculated the partial correlation on the combined sample (n = 74). To make the datasets as comparable as possible, we first averaged over T2|T1 accuracy at lags 2 and 4 in the previous study [20], to mimic a "lag 3" condition cf. the present study (Fig 5). The resulting T2|T1 accuracies at lag 3 are very comparable across the two studies (study 1: M=0.65, SD=0.26; study 2: M=0.64, SD=0.24)

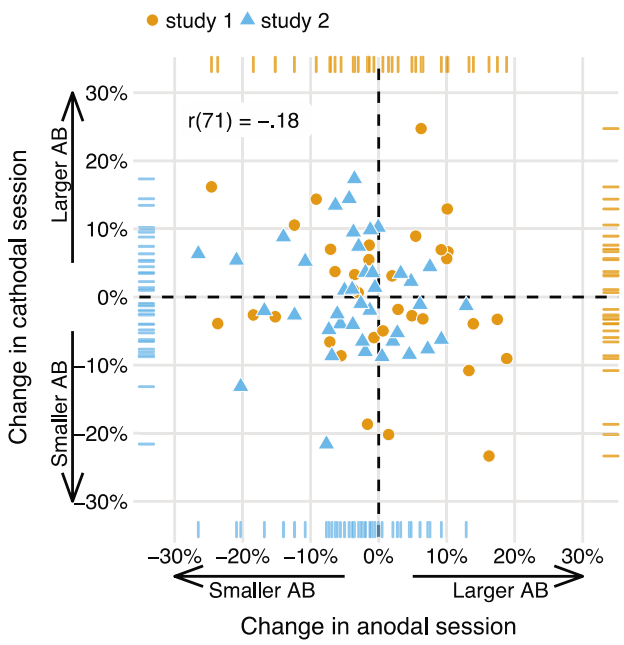


Fig 5. The effects of anodal and cathodal tDCS are also not correlated when pooling data from both studies. As in Fig 4, data points show AB magnitude change scores (tDCS - baseline) in the anodal session (x-axis) and the cathodal session (y-axis), but now for each participant in study 1 [20] and the present study (study 2). While study 1revealed a negative partial correlation, suggesting opposite effects of anodal and cathodal tDCS, this effect appears to be absent when based on the combined data of both studies. The partial correlation coefficient (attempting to adjust for Session order) is printed in the upper left.

The partial correlation based on the pooled data is $r(71) = -.18$. Thus, the correlation across a combination of both samples is a lot smaller than in the first study [20], and slightly smaller still than the meta-analytic estimate that included the original correlation from the first study (Fig 6).

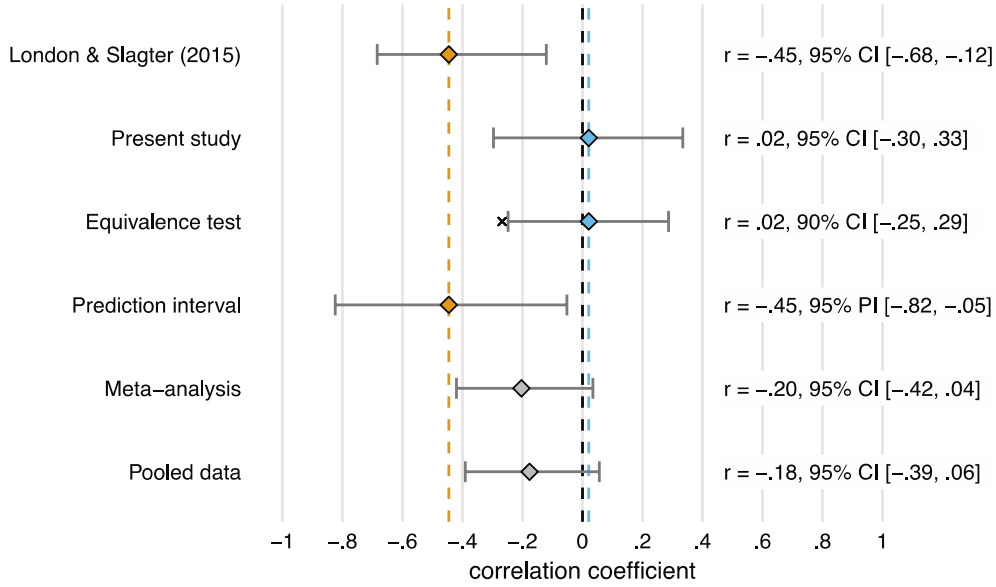


Fig 6. Summary of all the replication analyses (with exception of the replication Bayes Factor). The first two rows show the partial correlation (attempting to adjust for Session order) between the AB magnitude change scores (tDCS - baseline) in the anodal and cathodal sessions, for study 1 [20] (in yellow) and the present study (in blue). The first is significantly negative, the second is slightly positive and not significant, because its 95% confidence interval (CI) overlaps with zero. The third row shows the 90% CI around the correlation in the present study. Because this interval does not overlap with the "small telescopes" effect size, (indicated by the x: $-r_{33\%}$ = -.27), this correlation is significantly smaller. The fourth row shows the 95% prediction interval (PI) around the correlation in [20]. Because this interval does not overlap with the correlation in the present study, both correlations are not consistent. The final two rows show the overall effect when the two correlations are meta-analyzed, and when one correlation is computed over the pooled data from both studies. Neither are significant (95% CI overlaps with zero). Thus, our replication analyses all suggest that we failed to replicate London & Slagter [20], and when the results are examined in combination, no evidence

in support of a negative relationship between the AB magnitude change scores (tDCS - baseline) in the anodal and cathodal sessions is obtained.

In sum, all the analyses we conducted seem to support the same conclusion: that the current study is not a successful replication of the previous [20], and that therefore the purported opposing effect of anodal and cathodal tDCS on the attentional blink within individuals is called into question.

Discussion

Our recent study [20] was the first to uncover potential individual differences in the effect of tDCS on the AB. This original study revealed a negative correlation between the effects of anodal and cathodal tDCS on T2|T1 accuracy across individuals individuals (in the absence of a group-level effect of tDCS). This analysis suggested an interesting pattern of individual differences in response to tDCS: those that tended to benefit from anodal tDCS (i.e., whose AB became smaller compared to baseline) would tend to worsen during cathodal tDCS, and vice versa. This finding seemed plausible given the substantial and well-documented individual differences in both the AB [23] and the effects of tDCS [24].

We conducted a replication study, and again found no effect of tDCS at the group level. However, in contrast to our prior work [20], the correlation between the AB magnitude change scores (tDCS - baseline) in the anodal and cathodal sessions here is not significant, and not in the same direction. We therefore also computed several statistical measures of replication success focused on the negative correlation between anodal and cathodal tDCS. As discussed in more detail below, the field of electrical brain stimulation is

characterized by inconsistent findings across studies. It is hence important that researchers in this field also start to incorporate methods that permit quantification of replication success. It is our hope that the overview of main, complementary analyses that researchers can use to maximize the evidential value from negative findings, provided here, will serve as a guideline for future studies.

The measures of replication success showed that in our study the observed correlation is smaller than we found previously [20], and than the smallest correlation they could have reasonably detected (i.e., an equivalence test for the lower bound of $r_{33\%}$ [27] was significant). The difference between the two studies is greater than expected based on sampling error alone (i.e., the correlation in the present study falls outside of the 95% prediction interval). Further, the data provide moderate to strong evidence for the null hypothesis of zero correlation compared to the alternative that the correlation is as in the original study [20] (i.e., we found a replication Bayes factor in favor of the null hypothesis of ~10). Finally, combining both studies yields a smaller and non-significant correlation (both in a meta-analysis and by pooling the data).

The overall picture is consistent: all measures indicate that the present study is not a successful replication of our original study [20]. But because both studies were similar in design and sample size, it is not warranted to dismiss the previous findings. Our analysis revealed substantial variation between both estimates of the effect, and so the conclusion whether tDCS is effective should not be based on any single study, nor on only a single replication study [79]. In any series of studies, one should expect to sometimes not find a significant result, purely due to random variation—even if there is a true effect [80].

Still, the marked difference in results of both studies is surprising: they are similar enough that the present study [31] could be considered a direct replication of the first

study [20]. We used the exact same electrode montage and tDCS parameters, followed the same experimental design (Fig 1), ran the experiment in the same location with the same participant population, and used a virtually identical task. Nevertheless, there are some discrepancies between the studies that could have contributed to the different outcome (see S2 Table for a full list of the differences).

The most important difference is that in the previous study [20], T2 was presented at lags 2, 4, and 10, whereas we now used lags 3 and 8. This means that AB magnitude (the difference in T2|T1 accuracy between the shortest and longest lag) was on average smaller in the present study. We introduced the change for precisely this reason: to have about as many trials for the EEG analyses in which T2 was seen vs. missed, we needed the AB to be smaller. But this increase in average T2|T1 accuracy may have also reduced the between-subject variability that is essential for analyses that capitalize on individual differences. Indeed, from Fig 5 it seems that the previous sample [20] had a larger spread, at least in the change scores for AB magnitude.

Though the change in lags is probably the most important, the concurrent EEG measurement did introduce other differences as well. Each session took longer to complete, because the EEG setup required extra time, and the pace of the task was slowed down with longer inter-stimulus intervals. Finally, the current flow could have changed due to the presence of the EEG cap and electrodes, as well as the use of conductive paste instead of saline solution.

At the group level, our results are in agreement [20]: we find no differential effects of anodal and cathodal tDCS on the AB. However, another study [25] did report a group-level effect of tDCS on the AB, but this study used different stimulation parameters and was fairly small (see our previous paper [20] for a more extensive discussion of the

differences). At least with our AB task and fairly standard electrode montage and stimulation parameters [81], the AB does not appear to be very amenable to tDCS over the lDLPFC. Our null results are consistent with recent reviews and meta-analyses highlighting there is little evidence that prefrontal tDCS can be used to enhance cognition [82]; or if so, that its effects on attention [83] and cognition more generally [84] are difficult to predict, rather small [85], and restricted to a limited subset of outcome measures and stimulation parameters [86,87].

Interpreting null results is always difficult, especially in brain stimulation studies [88]. Ultimately, a myriad of possible underlying explanations may apply, most of which we have no direct access to. For one, we cannot be sure that the current density in the lDLPFC reached sufficient levels [89,90] based on our montage alone. More precise targeting of the stimulation towards the lDPLFC [91] and modeling of the current flow [92,93] could provide some more confidence. Still, it would be difficult to verify that anodal tDCS and cathodal tDCS indeed had an opposite effect on cortical excitability [94,95]. This assumption holds in many cases, and would provide a plausible explanation [24] for the negative correlation we originally found [20]. But ideally, it should be validated with direct measures of cortical excitability. Such measures are hard to obtain, although some recent studies suggest that combining tDCS with magnetic resonance spectroscopy can be used for this purpose [96–98]. Combining tDCS with MRI may also help predict individual differences in stimulation efficacy. One recent study found that individual variability in the cortical thickness of left prefrontal cortex accounted for almost 35% of the variance in stimulation efficacy among subjects [99].

It is also possible that we do not find an effect because the studies did not have sufficient statistical power [100]. Especially for the individual differences analyses, the sample size

in both studies is on the lower end. For example, to detect a medium-sized correlation (r = 0.3) with 80% power, the required sample size is n = 84. We do approach this sample size when combining both studies (n = 74). Also, the correlation in [20] was larger (r = -0.45), but we cannot have much confidence in the precision of this estimate: : the 95% confidence interval is so wide (-0.68 to -0.12) that it is consistent with a large range of effect sizes, and therefore almost unfalsifiable [101]. To estimate the size of a medium correlation (r = 0.3) within a narrower confidence band (± 0.15 with 90% confidence), a sample size of n = 134 would be required [102]. Although our analyses suggest that the correlation is small if anything, this means we cannot accurately estimate how small—even with the combined sample.

These are all decisions to be made at the design stage, which can increase the value of a null result [88]. However, after the data are in, additional tools are available to increase the value of a null result [26], especially in the case of a replication study [27], which we have demonstrated here.

We hope our paper provides inspiration for others in the fields of brain stimulation and cognitive neuroscience. Many speak of a crisis of confidence [41] and fear that the current literature is lacking in evidential value [82]. This is certainly not a unique development, as these sentiments [103] and low rates of replication abound in many research areas [32,33,104]. But it is perhaps aggravated by the fact that the brain stimulation field has not matured yet [105]. At the same time, this also leaves room for the field to improve, and more work is necessary to determine the potential of electrical brain stimulation [106].

To make sure the literature accurately reflects the efficacy of the technique, it is crucial to combat publication bias. Positive results are heavily over-represented in most of the

scientific literature, [107–109], which has recently prompted a call to brain stimulation researchers to also publish their null results (https://www.frontiersin.org/research-topics/5535/non-invasive-brain-stimulation-effects-on-cognition-and-brain-activity-positive-lessons-from-negativ). Preregistration and registered reports are other important methods to combat publication bias that the field of electrical brain stimulation should embrace (see e.g., [110]). Publishing null results and replication studies—and making the most of their interpretation—is crucial to better this situation.

Acknowledgements

We thank Raquel London for sharing all her experience and materials, as well as Daphne Box and Esther van der Giessen for their assistance in data collection.

References

1. Raymond JE, Shapiro KL, Arnell KM. Temporary suppression of visual processing in an RSVP task: An attentional blink? Journal of Experimental Psychology: Human Perception and Performance. 1992;18: 849–860.

2. Dux PE, Marois R. The attentional blink: A review of data and theory. Attention, Perception, and Psychophysics. 2009;71: 1683–1700. doi:10.3758/APP.71.8.1683

3. Martens S, Wyble B. The attentional blink: Past, present, and future of a blind spot in perceptual awareness. Neuroscience and Biobehavioral Reviews. Elsevier Ltd; 2010;34: 947–957. doi:10.1016/j.neubiorev.2009.12.005

4. MacLean MH, Arnell KM. A conceptual and methodological framework for measuring and modulating the attentional blink. Attention, Perception, and Psychophysics. 2012;74: 1080–1097. doi:10.3758/s13414-012-0338-4

5. Olivers CNL, Nieuwenhuis S, Olivers CNL, Nieuwenhuis S. The beneficial effect of concurrent task-irrelevant mental activity on temporal attention. Psychological Science. 2005;16: 265–269. doi:10.1111/j.0956-7976.2005.01526.x

6. Olivers CNL, Nieuwenhuis S. The beneficial effects of additional task load, positive affect, and instruction on the attentional blink. Journal of Experimental Psychology: Human Perception and Performance. 2006;32: 364–379. doi:10.1037/0096-1523.32.2.364

7. Thomson DR, Ralph BCW, Besner D, Smilek D. The more your mind wanders, the smaller your attentional blink: An individual differences study. Quarterly Journal of Experimental Psychology. 2015;68: 181–191. doi:10.1080/17470218.2014.940985

8. Slagter HA, Lutz A, Greischar LL, Francis AD, Nieuwenhuis S, Davis JM, et al. Mental training affects distribution of limited brain resources. PLoS biology. 2007;5: e138. doi:10.1371/journal.pbio.0050138

9. Dayan E, Censor N, Buch ER, Sandrini M, Cohen LG. Noninvasive brain stimulation: from physiology to network dynamics and back. Nature neuroscience. Nature Publishing Group; 2013;16: 838–44. doi:10.1038/nn.3422

10. Cooper ACG, Humphreys GW, Hulleman J, Praamstra P, Georgeson M. Transcranial magnetic stimulation to right parietal cortex modifies the attentional blink. Experimental Brain Research. 2004;155: 24–29. doi:10.1007/s00221-003-1697-9

11. Arasanz CP, Staines WR, Schweizer TA. Isolating a cerebellar contribution to rapid visual attention using transcranial magnetic stimulation. Frontiers in Behavioral Neuroscience. 2012;6: 55. doi:10.3389/fnbeh.2012.00055

12. Esterman M, Thai M, Okabe H, DeGutis J, Saad E, Laganiere SE, et al. Network-targeted cerebellar transcranial magnetic stimulation improves attentional control. NeuroImage. Elsevier; 2017;156: 190–198. doi:10.1016/j.neuroimage.2017.05.011

13. Gebodh N, Esmaeilpour Z, Bikson M, Adair D, Fregni F, Schestattsky P. Transcranial direct current stimulation among technologies for low-intensity transcranial electrical stimulation: classification, history, and terminology. In: Knotkova H, Nitsche MA, Bikson M, Woods AJ, editors. Practical guide to transcranial direct current stimulation. Cham (Switzerland): Springer; 2019. pp. 3–43. doi:10.1007/978-3-319-95948-1_1

14. Nitsche MA, Paulus W. Excitability changes induced in the human motor cortex by weak transcranial direct current stimulation. The Journal of physiology. 2000;527 Pt 3: 633–9. doi:10.1111/j.1469-7793.2000.t01-1-00633.x

15. Nitsche MA, Paulus W. Sustained excitability elevations induced by transcranial DC motor cortex stimulation in humans. Neurology. 2001;57: 1899–1901. doi:10.1212/WNL.57.10.1899

16. Parkin BL, Bhandari M, Glen JC, Walsh V. The physiological effects of transcranial electrical stimulation do not apply to parameters commonly used in studies of cognitive neuromodulation. Neuropsychologia. Elsevier Ltd; 2018; 1–8. doi:10.1016/j.neuropsychologia.2018.03.030

17. Bikson M, Paulus W, Esmaeilpour Z, Kronberg G, Nitsche MA. Mechanisms of acute and after effects of transcranial direct current stimulation. In: Knotkova H, Nitsche MA, Bikson M, Woods AJ, editors. Practical guide to transcranial direct current stimulation. Cham (Switzerland): Springer; 2019. pp. 81–113. doi:10.1007/978-3-319-95948-1

18. Liu A, Vöröslakos M, Kronberg G, Henin S, Krause MR, Huang Y, et al. Immediate neurophysiological effects of transcranial electrical stimulation. Nature Communications. Springer US; 2018;9: 5092. doi:10.1038/s41467-018-07233-7

19. Stagg CJ, Antal A, Nitsche MA. Physiology of transcranial direct current stimulation. Journal of ECT. 2018;34: 144–152. doi:10.1097/YCT.0000000000000510

20. London RE, Slagter HA. No Effect of Transcranial Direct Current Stimulation over Left Dorsolateral Prefrontal Cortex on Temporal Attention. Journal of Cognitive Neuroscience. 2021;33: 756–768. doi:10.1162/jocn_a_01679

21. Slagter HA, Johnstone T, Beets IAM, Davidson RJ. Neural competition for conscious representation across time: An fMRI study. PLoS ONE. 2010;5: e10556. doi:10.1371/journal.pone.0010556

22. Hommel B, Kessler K, Schmitz F, Gross J, Akyürek E, Shapiro K, et al. How the brain blinks: towards a neurocognitive model of the attentional blink. Psychological Research. 2006;70: 425–435. doi:10.1007/s00426-005-0009-3

23. Willems C, Martens S. Time to see the bigger picture: Individual differences in the attentional blink. Psychonomic Bulletin and Review. Psychonomic Bulletin & Review; 2016;23: 1289–1299. doi:10.3758/s13423-015-0977-2

24. Krause B, Cohen Kadosh R. Not all brains are created equal: the relevance of individual differences in responsiveness to transcranial electrical stimulation. Frontiers in systems neuroscience. 2014;8: 25. doi:10.3389/fnsys.2014.00025

25. Sdoia S, Conversi D, Pecchinenda A, Ferlazzo F. Access to consciousness of briefly presented visual events is modulated by transcranial direct current stimulation of left dorsolateral prefrontal cortex. Scientific Reports. 2019;9: 10950. doi:10.1038/s41598-019-47527-4

111. Schmidt S. Shall we really do it again? The powerful concept of replication is neglected in the social sciences. Review of General Psychology. 2009;13: 90-100. doi: 10.1037/a0015108

26. Harms C, Lakens D. Making 'null effects' informative: statistical techniques and inferential frameworks. Journal of Clinical and Translational Research. 2018;3: 382–393. doi:10.18053/jctres.03.2017s2.007

27. Simonsohn U. Small telescopes: Detectability and the evaluation of replication results. Social Science Research Network. 2015; 1–11. doi:http://dx.doi.org/10.2139/ssrn.2259879

28. Brandt MJ, IJzerman H, Dijksterhuis A, Farach FJ, Geller J, Giner-Sorolla R, et al. The Replication Recipe: What makes for a convincing replication? Journal of Experimental Social Psychology. 2014;50: 217–224. doi:10.1016/j.jesp.2013.10.005

29. Anderson SF, Maxwell SE. There's more than one way to conduct a replication study: Beyond statistical significance. Psychological Methods. 2016;21: 1–12. doi:10.1037/met0000051

30. Zondervan-Zwijnenburg M, Rijshouwer D. Testing Replication with Small Samples. In: Schoot R van de, Miočević M, editors. Small sample size solutions: A guide for applied researchers and practitioners. Abingdon, Oxon ; New York, NY: Routledge; 2020.

31. Zwaan RA, Etz A, Lucas RE, Donnellan MB. Making replication mainstream. Behavioral and Brain Sciences. 2018;e120. doi:10.1017/S0140525X17001972

32. Camerer CF, Dreber A, Holzmeister F, Ho TH, Huber J, Johannesson M, et al. Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. Nature Human Behaviour. 2018;2: 637–644. doi:10.1038/s41562-018-0399-z

33. Open Science Collaboration. Estimating the reproducibility of psychological science. Science. 2015;349: aac4716-1-aac4716-8. doi:10.1126/science.aac4716

34. Lakens D, Scheel AM, Isager PM. Equivalence Testing for Psychological Research: A Tutorial. Advances in Methods and Practices in Psychological Science. 2018;1: 259–269. doi:10.1177/2515245918770963

35. Spence JR, Stanley DJ. Prediction interval: What to expect when you're expecting … a replication. Dubé M-P, editor. PLOS ONE. 2016;11: e0162874. doi:10.1371/journal.pone.0162874

36. Patil P, Peng RD, Leek JT. What should researchers expect when they replicate studies? A statistical view of replicability in psychological science. Perspectives on Psychological Science. 2016;11: 539–544. doi:10.1177/1745691616646366

37. Wagenmakers E-J, Verhagen J, Ly A. How to quantify the evidence for the absence of a correlation. Behavior research methods. Behavior Research Methods; 2016;48: 413–426. doi:10.3758/s13428-015-0593-0

38. Verhagen J, Wagenmakers EJ. Bayesian tests to quantify the result of a replication attempt. Journal of Experimental Psychology: General. 2014;143: 1457–1475. doi:10.1037/a0036731

39. Quintana DS. From pre-registration to publication: A non-technical primer for conducting a meta-analysis to synthesize correlational data. Frontiers in Psychology. 2015;6: 1549. doi:10.3389/fpsyg.2015.01549

40. Goh JX, Hall JA, Rosenthal R. Mini Meta-Analysis of Your Own Studies: Some Arguments on Why and a Primer on How. Social and Personality Psychology Compass. 2016;10: 535–549. doi:10.1111/spc3.12267

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

42. Munafò MR, Nosek BA, Bishop DVM, Button KS, Chambers CD, Percie du Sert N, et al. A manifesto for reproducible science. Nature Human Behaviour. Macmillan Publishers Limited; 2017;1: 0021. doi:10.1038/s41562-016-0021

43. Huber DE, Potter KW, Huszar LD. Less “story” and more “reliability” in cognitive neuroscience. Cortex. Elsevier Ltd; 2019; doi:10.1016/j.cortex.2018.10.030

44. Lakens D, McLatchie N, Isager PM, Scheel AM, Dienes Z. Improving Inferences About Null Effects With Bayes Factors and Equivalence Tests. The Journals of Gerontology: Series B. 2020;75: 45–57. doi:10.1093/geronb/gby065

45. Li LM, Uehara K, Hanakawa T. The contribution of interindividual factors to variability of response in transcranial direct current stimulation studies. Frontiers in Cellular Neuroscience. 2015;9: 181. doi:10.3389/fncel.2015.00181

46. Harty S, Sella F, Cohen Kadosh R. Transcranial Electrical Stimulation and Behavioral Change: The Intermediary Influence of the Brain. Frontiers in Human Neuroscience. 2017;11: 1–5. doi:10.3389/fnhum.2017.00112

47. Slagter HA, Georgopoulou K. Distractor inhibition predicts individual differences in recovery from the attentional blink. PloS one. 2013;8: e64681. doi:10.1371/journal.pone.0064681

48. Dux PE, Marois R. Distractor inhibition predicts individual differences in the attentional blink. Holcombe AO, editor. PloS one. Public Library of Science; 2008;3: e3330. doi:10.1371/journal.pone.0003330

49. Gorgolewski KJ, Auer T, Calhoun VD, Craddock RC, Das S, Duff EP, et al. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. Scientific Data. 2016;3: 160044. doi:10.1038/sdata.2016.44

50. Pernet CR, Appelhoff S, Gorgolewski KJ, Flandin G, Phillips C, Delorme A, et al. EEG-BIDS, an extension to the brain imaging data structure for electroencephalography. Scientific Data. 2019;6: 103. doi:10.1038/s41597-019-0104-8

51. R Core Team. R: A language and environment for statistical computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2019. Available: https://www.R-project.org/

52. Morey RD, Rouder JN. BayesFactor: Computation of bayes factors for common designs [Internet]. 2018. Available: https://CRAN.R-project.org/package=BayesFactor

53. Robinson D, Hayes A, Couch S. Broom: Convert statistical objects into tidy tibbles. 2021.

54. Wilke CO. Cowplot: Streamlined plot theme and plot annotations for 'ggplot2' [Internet]. 2020. Available: https://CRAN.R-project.org/package=cowplot

55. Lenth RV. Emmeans: Estimated marginal means, aka least-squares means [Internet]. 2021. Available: https://CRAN.R-project.org/package=emmeans

56. Müller K. Here: A simpler way to find your files [Internet]. 2020. Available: https://CRAN.R-project.org/package=here

57. Zhu H. kableExtra: Construct complex table with 'kable' and pipe syntax [Internet]. 2021. Available: https://CRAN.R-project.org/package=kableExtra

58. Xie Y. Dynamic documents with R and knitr [Internet]. 2nd ed. Boca Raton, Florida: Chapman; Hall/CRC; 2015. Available: https://yihui.org/knitr/

59. Aust F, Barth M. papaja: Create APA manuscripts with R Markdown [Internet]. 2020. Available: https://github.com/crsh/papaja

60. Fletcher TD. Psychometric: Applied psychometric theory [Internet]. 2010. Available: https://CRAN.R-project.org/package=psychometric

61. Champely S. Pwr: Basic functions for power analysis [Internet]. 2020. Available: https://CRAN.R-project.org/package=pwr

62. Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, et al. Welcome to the tidyverse. Journal of Open Source Software. 2019;4: 1686. doi:10.21105/joss.01686

63. RStudio Team. RStudio: Integrated Development Environment for R [Internet]. Boston, MA: RStudio, Inc.; 2016. Available: http://www.rstudio.com/

64. Singmann H, Bolker B, Westfall J, Aust F, Ben-Shachar MS. Afex: Analysis of factorial experiments [Internet]. 2021. Available: https://CRAN.R-project.org/package=afex

65. Bakeman R. Recommended effect size statistics for repeated measures designs. Behavior research methods. 2005;37: 379–384. doi:10.3758/BF03192707

66. Marchetti GM, Drton M, Sadeghi K. Ggm: Graphical markov models with mixed graphs [Internet]. 2020. Available: https://CRAN.R-project.org/package=ggm

67. Muradchanian J, Hoekstra R, Kiers H, Ravenzwaaij D van. How Best to Quantify Replication Success? A Simulation Study on the Comparison of Replication Success Metrics [Internet]. MetaArXiv; 2020 Aug. doi:10.31222/osf.io/wvdjf

68. Linde M, Tendeiro J, Selker R, Wagenmakers E-J, Ravenzwaaij D van. Decisions About Equivalence: A Comparison of TOST, HDI-ROPE, and the Bayes Factor. PsyArXiv. 2020; doi:10.31234/osf.io/bh8vu

69. Lakens D. Equivalence tests: A practical primer for t-tests, correlations, and meta-analyses. Social Psychological and Personality Science. 2017;1: 1–8. doi:10.1177/1948550617697177

70. Stanley D. predictionInterval: Prediction interval functions for assessing replication study results [Internet]. 2016. Available: https://CRAN.R-project.org/package=predictionInterval

71. Wagenmakers EJ, Marsman M, Jamil T, Ly A, Verhagen J, Love J, et al. Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. Psychonomic Bulletin and Review. 2018;25: 35–57. doi:10.3758/s13423-017-1343-3

72. Wagenmakers EJ, Love J, Marsman M, Jamil T, Ly A, Verhagen J, et al. Bayesian inference for psychology. Part II: Example applications with JASP. Psychonomic Bulletin and Review. 2018;25: 58–76. doi:10.3758/s13423-017-1323-7

73. Viechtbauer W. Conducting meta-analyses in R with the metafor package. Journal of Statistical Software. 2010;36: 1–48. Available: https://www.jstatsoft.org/v36/i03/

74. Ueno T, Fastrich GM, Murayama K. Meta-analysis to integrate effect sizes within an article: Possible misuse and Type I error inflation. Journal of Experimental Psychology: General. 2016;145: 643–654. doi:10.1037/xge0000159

75. Simmons JP, Nelson LD, Simonsohn U. False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. Psychological Science. 2011;22: 1359–1366. doi:10.1177/0956797611417632

76. Lakens D. Performing high-powered studies efficiently with sequential analyses. European Journal of Social Psychology. 2014;44: 701–710. doi:10.1002/ejsp.2023

77. Xie Y, Allaire JJ, Grolemund G. R Markdown: The Definitive Guide. Boca Raton, FL: CRC Press; 2018. doi:10.1016/B978-0-12-814447-3.00041-0

78. LeBel EP, Vanpaemel W, Cheung I, Campbell L. A Brief Guide to Evaluate Replications. Meta-Psychology. 2019;3. doi:10.15626/MP.2018.843

79. Hedges LV, Schauer JM. More Than One Replication Study Is Needed for Unambiguous Tests of Replication. Journal of Educational and Behavioral Statistics. 2019;44: 543–570. doi:10.3102/1076998619852953

80. Amrhein V, Trafimow D, Greenland S. Inferential Statistics as Descriptive Statistics: There Is No Replication Crisis if We Don't Expect Replication. The American Statistician. 2019;73: 262–270. doi:10.1080/00031305.2018.1543137

81. Santarnecchi E, Brem A-K, Levenbaum E, Thompson T, Kadosh RC, Pascual-Leone A. Enhancing cognition using transcranial electrical stimulation. Current Opinion in Behavioral Sciences. Elsevier Ltd; 2015;4: 171–178. doi:10.1016/j.cobeha.2015.06.003

82. Medina J, Cason S. No evidential value in samples of transcranial direct current stimulation (tDCS) studies of cognition and working memory in healthy populations. Cortex. Elsevier Ltd; 2017;94: 131–141. doi:10.1016/j.cortex.2017.06.021

83. Reteig LC, Talsma LJ, Schouwenburg MR van, Slagter HA. Transcranial Electrical Stimulation as a Tool to Enhance Attention. Journal of Cognitive Enhancement. 2017;1: 10–25. doi:10.1007/s41465-017-0010-y

84. Tremblay S, Lepage J-F, Latulipe-Loiselle A, Fregni F, Pascual-Leone A, Théoret H. The uncertain outcome of prefrontal tDCS. Brain Stimulation. 2014;7: 773–783. doi:10.1016/j.brs.2014.10.003

85. Dedoncker J, Brunoni AR, Baeken C, Vanderhasselt MA. A systematic review and meta-analysis of the effects of transcranial direct current stimulation (tDCS) over the dorsolateral prefrontal cortex in healthy and neuropsychiatric samples: influence of stimulation parameters. Brain Stimulation. Elsevier Inc.; 2016;9: 501–517. doi:10.1016/j.brs.2016.04.006

86. Boer NS de, Schluter RS, Daams JG, Werf YD van der, Goudriaan AE, Holst RJ van. The effect of non-invasive brain stimulation on executive functioning in healthy controls: A systematic review and meta-analysis. Neuroscience & Biobehavioral Reviews. 2021;125: 122–147. doi:10.1016/j.neubiorev.2021.01.013

87. Imburgio MJ, Orr JM. Effects of prefrontal tDCS on executive function: Methodological considerations revealed by meta-analysis. Neuropsychologia. 2018;117: 156–166. doi:10.1016/j.neuropsychologia.2018.04.022

88. Graaf TA de, Sack AT. When and how to interpret null Results in NIBS: A taxonomy based on prior expectations and experimental design. Frontiers in Neuroscience. 2018;12: 915. doi:10.3389/fnins.2018.00915

89. Kim J-H, Kim D-W, Chang WH, Kim Y-H, Kim K, Im C-H. Inconsistent outcomes of transcranial direct current stimulation may originate from anatomical differences among individuals: electric field simulation using individual MRI data. Neuroscience letters. 2014;564: 6–10. doi:10.1016/j.neulet.2014.01.054

90. Opitz A, Paulus W, Will S, Antunes A, Thielscher A. Determinants of the electric field during transcranial direct current stimulation. NeuroImage. Elsevier Inc.; 2015;109: 140–150. doi:10.1016/j.neuroimage.2015.01.033

91. Datta A, Bansal V, Diaz J, Patel J, Reato D, Bikson M. Gyri-precise head model of transcranial direct current stimulation: Improved spatial focality using a ring electrode versus conventional rectangular pad. Brain Stimulation. Elsevier Inc; 2009;2: 201–207.e1. doi:10.1016/j.brs.2009.03.005

92. Berker AO de, Bikson M, Bestmann S. Predicting the behavioral impact of transcranial direct current stimulation: issues and limitations. Frontiers in human neuroscience. 2013;7: 613. doi:10.3389/fnhum.2013.00613

93. Karabanov AN, Saturnino GB, Thielscher A, Siebner HR. Can transcranial electrical stimulation localize brain function? Frontiers in Psychology. 2019;10: 213. doi:10.3389/fpsyg.2019.00213

94. Bestmann S, de Berker AO, Bonaiuto J. Understanding the behavioural consequences of noninvasive brain stimulation. Trends in Cognitive Sciences. Elsevier Ltd; 2015;19: 13–20. doi:10.1016/j.tics.2014.10.003

95. Bestmann S, Walsh V. Transcranial electrical stimulation. Current Biology. Elsevier; 2017;27: R1258–R1262. doi:10.1016/j.cub.2017.11.001

96. Filmer HL, Ehrhardt S, Bollmann S, Mattingley JB, Dux PE. Accounting for individual differences in the response to tDCS with baseline levels of neurochemical excitability. Cortex. Elsevier Ltd; 2019;115: 324–334. doi:10.1016/j.cortex.2019.02.012

97. Antonenko D, Thielscher A, Saturnino G, Aydin S, Ittermann B, Grittner U, et al. Towards precise brain stimulation: Is electric field simulation related to neuromodulation? Brain Stimulation. 2019; doi:10.1016/j.brs.2019.03.072

98. Talsma LJ, Broekhuizen JA, Huisman J, Slagter HA. No evidence that baseline prefrontal cortical excitability (3T-MRS) predicts the effects of prefrontal tDCS on WM performance. Frontiers in Neuroscience. 2018;12: 481. doi:10.3389/fnins.2018.00481

99. Filmer HL, Ehrhardt SE, Shaw TB, Mattingley JB, Dux PE. The efficacy of transcranial direct current stimulation to prefrontal areas is related to underlying cortical morphology. NeuroImage. 2019;196: 41–48. doi:10.1016/j.neuroimage.2019.04.026

100. Minarik T, Berger B, Althaus L, Bader V, Biebl B, Brotzeller F, et al. The importance of sample size for reproducibility of tDCS effects. Frontiers in Human Neuroscience. 2016;10: 453. doi:10.3389/FNHUM.2016.00453

101. Morey R, Lakens D. Why most of psychology is statistically unfalsifiable. 2016; doi:10.5281/zenodo.838685

102. Schönbrodt FD, Perugini M. At what sample size do correlations stabilize ? Journal of Research in Personality. 2013;47: 609–612. doi:10.1016/j.jrp.2013.05.009

103. Baker M. 1,500 scientists lift the lid on reproducibility. Nature. 2015;533: 52–454.

104. Klein RA, Vianello M, Hasselman F, Adams BG, Adams RB, Alper S, et al. Many Labs 2: Investigating variation in replicability across samples and settings. Advances in Methods and Practices in Psychological Science. 2018;1: 443–490. doi:10.1177/2515245918810225

105. Parkin BL, Ekhtiari H, Walsh VF. Non-invasive human brain stimulation in cognitive neuroscience: a primer. Neuron. Elsevier Inc.; 2015;87: 932–945. doi:10.1016/j.neuron.2015.07.032

106. Filmer HL, Mattingley JB, Dux PE. Modulating brain activity and behaviour with tDCS: Rumours of its death have been greatly exaggerated. Cortex. 2020;123: 141–151. doi:10.1016/j.cortex.2019.10.006

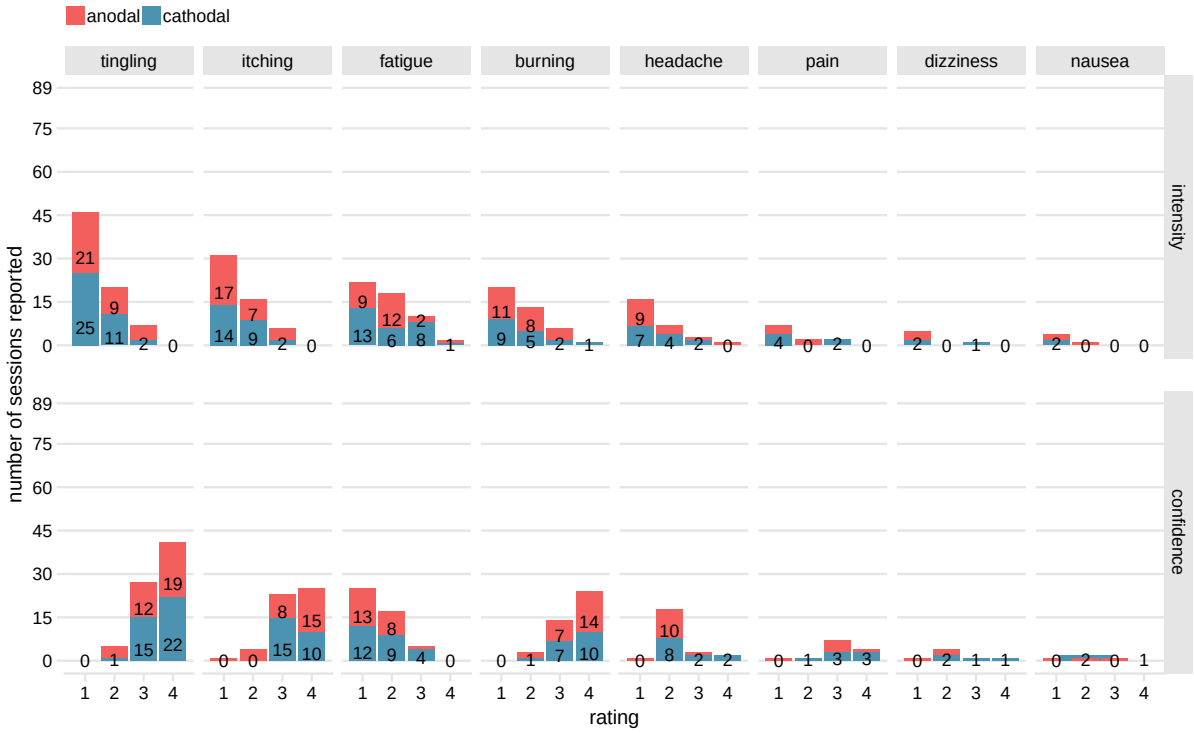
107. Ferguson CJ, Heene M. A vast graveyard of undead theories: Publication bias and psychological science's aversion to the null. Perspectives on Psychological Science. 2012;7: 555–561. doi:10.1177/1745691612459059

108. Franco A, Malhotra N, Simonovits G. Publication bias in the social sciences: Unlocking the file drawer. Science. 2014;345: 1502–1505. doi:10.1126/science.1255484

109. Fanelli D. Negative results are disappearing from most disciplines and countries. Scientometrics. 2012;90: 891–904. doi:10.1007/s11192-011-0494-7

110. Horne KS, Filmer HL, Nott ZE, Hawi Z, Pugsley K, Mattingley JB, et al. Evidence against benefits from cognitive training and transcranial direct current stimulation in healthy older adults. Nature Human Behaviour. 2021;5: 146–158. doi:10.1038/s41562-020-00979-5

Supporting information



S1 Fig. tDCS adverse events. Number of reports out of 89 sessions (either anodal or cathodal tDCS). Top row shows intensity ratings [little, moderate, strong, very strong]; bottom row shows participant's confidence that event was related to tDCS [unlikely, possibly, likely, very likely]. Adverse events are sorted in descending order of number of reports (for very rare events (five reports or fewer for a given polarity), some text counts have been removed to prevent overlap).

Table S1

Number of reports of tDCS side effects

	Intensity rating					Confidence rating				
	none	a little	moderate	strong	very strong	n/a	unlikely	possibly	likely	very likely
anodal session										
burning	23	11	8	4	0	23	0	2	7	14
dizziness	43	3	0	0	0	43	1	2	0	0
fatigue	22	9	12	2	1	24	13	8	1	0
headache	32	9	3	1	1	34	1	10	1	0
itching	18	17	7	4	0	18	1	4	8	15
nausea	43	2	1	0	0	43	1	1	1	0
pain	41	3	2	0	0	40	1	0	4	1
tingling	11	21	9	5	0	11	0	4	12	19
cathodal session										
burning	26	9	5	2	1	25	0	1	7	10
dizziness	40	2	0	1	0	39	0	2	1	1
fatigue	15	13	6	8	1	18	12	9	4	0
headache	30	7	4	2	0	31	0	8	2	2
itching	18	14	9	2	0	18	0	0	15	10
nausea	41	2	0	0	0	40	0	2	0	1
pain	37	4	0	2	0	36	0	1	3	3
tingling	5	25	11	2	0	5	0	1	15	22

[1] Intensity rating: To which degree were the following sensations present during stimulation?

[2] Confidence rating: To which degree do you believe this was caused by the stimulation?

S1 Table. Number of reports of tDCS side effects. The intensity rating question read: "To which degree were the following sensations present during stimulation?". The confidence rating question read: "To which degree do you believe this was caused by the stimulation?".

Difference	Study 2	Study 1	Motivation
Sample: size	40	34	
Sample: age	Mean: 20.94, SD: 4.25	Mean: 22.4, SD: 2.8	
Sample: gender	29 female (73%)	20 female (59%)	
Design: concurrent measurements	EEG	None	
Design: inter-session interval	1 week	min. 48 hours	Increase similarity between sessions; longer washout time
Design: trials per 20-minute block	Short-lag trials: M=130, SD=17; Long-lag trials: M=65; SD=9	Short-lag trials: 40; Long-lag trials: 40	Allow as many trials as possible in 20 minutes by making the task self-paced
Task: RSVP stream	15 letters	17 letters	Shorten the total length of a trial, to allow completion of more trials
Task: inter-stimulus interval	1750 ms, + 1000 ms	480 ms	Allow EEG responses to return to baseline
Task: priming condition	Absent	Present	Remove priming condition as it is orthogonal to main contrast, and showed no effects in @London2021
Task: attentional blink conditions	lag 3, 8	lag 2, 4, 10	Increase proportion of trials where T2 was seen, for EEG analysis
tDCS: conductive medium	Conductive paste	Saline solution	Prevent bridging EEG electrodes with (spreading) saline solution

S2 Table. Differences in the methodology and participant samples of London & Slagter [20] (Study 1) and the present study (Study 2).