bioRxiv preprint doi: https://doi.org/10.1101/2020.12.06.411850; this version posted December 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Is it all in our head? When subjective beliefs about receiving an intervention are better predictors of experimental results than the intervention itself

Luisa Fassi & Roi Cohen Kadosh*

Wellcome Centre for Integrative Neuroimaging, Department of Experimental Psychology, University of Oxford, UK

*Correspondence:

Luisa Fassi (luisafassi@live.com)

Roi Cohen Kadosh (roi.cohenkadosh@psy.ox.ac.uk)

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.06.411850; this version posted December 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Abstract

In recent years, there has been debate about the effectiveness of interventions from different fields (e.g., non-invasive brain stimulation (NIBS), neurofeedback, cognitive training programs) due to contradictory and nuanced experimental findings. Up to date, studies are focused on comparing the effects of an active form of the intervention to a placebo/control condition. However, a neglected question is how to consider individual differences in response to blinding procedures, and their effect on behavioural outcomes, rather than merely compare the efficacy of blinding using a group-based approach. To address this gap in the literature, we here suggest using subjective intervention—the participants' subjective beliefs about receiving or not receiving an intervention—as a factor. Specifically, we examined whether subjective intervention and subjective dosage (i.e. participants' subjective beliefs about the intensity of the intervention they received) affected performance scores independently, or interacting with, the active experimental condition. We carried out data analysis on an open-access dataset that has shown the efficacy of active NIBS in altering mind wandering. We show that subjective intervention and subjective dosage successfully explained alteration in mind wandering scores, over and beyond the objective intervention. These findings highlight the importance of accounting for the participants' beliefs about receiving interventions at the individual level by demonstrating their effect on human behaviour independently of the actual intervention. Altogether, our approach allows more rigorous and improved experimental design and analysis, which will strengthen the conclusions coming from basic and clinical research, for both NIBS and non-NIBS interventions.

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.06.411850; this version posted December 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Is it all in our head? When subjective beliefs about receiving an intervention are better predictors of experimental results than the intervention itself

A substantial amount of research from the fields of medicine, neuroscience, psychology and education is aimed at establishing the effectiveness of different interventions on both clinical and non-clinical populations. In recent years, a great deal of resources has been invested in the experimental assessment of drugs, vaccines, cognitive training programs, neurofeedback, and brain stimulation. In many cases, research findings have failed to provide definite and convincing results. The nuanced and often contradictory picture emerging from the published literature has led to increasing scepticism concerning the effectiveness of the examined interventions (Lampit et al., 2014; López-Alonso et al., 2014; Sitaram et al., 2017).

Several factors have been pointed at as the plausible reasons underlying the heterogeneity in research findings that arises from different studies. Literature suggests that a big part of the variability in experimental results may be due to the intra-individual and inter-individual variation in the investigated outcome measures, the differences in the spectrum of features that each intervention entails, poor experimental designs, as well as the complex interaction between those factors (Filmer et al., 2020; Gruzelier, 2014; Guerra et al., 2020).

Arguably, experimental design is the factor that falls under researchers’ control to a more considerable extent. Therefore, it is essential to thoroughly consider its effects on the results of a given experiment. Two crucial, yet often neglected, aspects of experimental design are 1) the way effective blinding is assessed and, 2) the lack of modelling individual responses associated with blinding in explaining experimental data. In the present study, we focus on the latter, while also providing an advancement on the former by analysing open-access data from the field of non-invasive brain stimulation (NIBS).

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.06.411850; this version posted December 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

In recent years, NIBS has been highlighted as a promising intervention to treat psychiatric, neurological, and neurodevelopmental disorders (Brunoni et al., 2019; Krause & Cohen Kadosh, 2013; Stefaniak et al., 2020; Vicario et al., 2019).

Among the latest research developments, an increasing number of studies have also investigated the use of NIBS for cognitive enhancement (Cohen Kadosh, 2014; Santarnecchi et al., 2015). In this context, converging evidence shows that NIBS has the potential to improve various cognitive abilities - such as attention, memory, and intelligence - in typical and atypical populations from different age groups (Cohen Kadosh et al., 2012; Hamilton et al., 2011; Santarnecchi et al., 2015).

One of the most popular NIBS techniques, transcranial direct current stimulation (tDCS), involves delivering either a positive (anodal) or negative (cathodal) electric constant current at a fixed amplitude over a selected cortical region (Filmer et al., 2014; Polania et al., 2018; Santarnecchi et al., 2015). Studies that employ this technique have shown promising evidence concerning the benefits brought by a single tDCS session to various cognitive domains, including notable findings on working memory, short-term and long-term memory, as well as speech, language, and mathematical cognition (André et al., 2016; Fregni et al., 2005; Javadi & Cheng, 2013; Monti et al., 2013; Sarkar et al., 2014).

Despite the encouraging results on the cognitive effects of tDCS, scepticism has been raised concerning the effectiveness of this technique (Horvath et al., 2015; Medina & Cason, 2017; Wang et al., 2018; Westwood et al., 2017).

Different factors have been hypothesised to underly the disparity in research findings that emerges across tDCS experiments. Here, we focus on two factors, namely: blinding and *subjective intervention*.

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.06.411850; this version posted December 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Blinding and Subjective Intervention

Blinding refers to the practice of withholding information that may undesirably influence participants' responses. Participants that take part in NIBS studies consistently report various perceptual sensations, such as audible clicks, visual disturbances, and cutaneous feelings (Davis et al., 2013; May et al., 2007). As a result of these sensations, participants can become aware of having received stimulation, and they are, therefore, unblinded to their experimental condition. This, in turn, makes it more likely that expectations about the effect of NIBS and demand characteristics about the aim of the experiment might influence participants' performance (Polania et al., 2018).

In order to tackle the issue of unblinding and the consequent emergence of any bias associated with it, studies have developed experimental designs that, at first sight, seem to control successfully for the sensations produced by active stimulation. In this regard, NIBS experiments most commonly employ sham (placebo) control stimulation (henceforth, sham stimulation). During sham stimulation, a minimal amount or no stimulation is administered to the subject while keeping the experiment otherwise identical. For example, in tDCS, the most prevalent way of administering sham is the *fade in- short stimulation- fade out* approach. Following this procedure, the stimulation intensity is slowly ramped up and, after a few seconds of actual stimulation (e.g., 30 seconds), ramped down promptly. Consequently, no stimulation is given for a total duration that is equal to the duration of the active stimulation (e.g., 20 min). Hence, sham stimulation aims to mimic the perceptual sensations associated with active stimulation without substantially affecting cortical excitability (Fritsch et al., 2010; Nitsche & Paulus, 2000).

Although previous tDCS studies have suggested that sham stimulation delivers sensations that are comparable to the active stimulation, and therefore allows adequate blinding (Ambrus et al., 2010, Ambrus et al., 2012; Gandiga et al., 2006), unblinding has also been reported

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.06.411850; this version posted December 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

(Fonteneau et al., 2019; O'connell et al., 2012; Turi et al., 2019; Wallace et al., 2016). Notably, ineffective blinding is likely to mask or inflate the observed findings. Therefore, it becomes of crucial importance to thoroughly account for its emergence.

Thus far, literature has shown that the experimental manipulation of participants' beliefs about a given intervention can moderate the intervention's efficacy at both the physiological and behavioural level (Benedetti et al., 2003; Boot et al., 2013; Finniss et al., 2010). However, a gap in the current knowledge is not whether blinding was successful or not, but whether the participant's subjective experience of receiving an intervention, whether it is true or not, can impact results over and beyond the actual intervention that participants receive. We term the former *subjective intervention* and the latter *objective intervention*.

The present study

Based on the above literature, we aimed to answer two research questions: (1) Does *subjective intervention* or *subjective dosage* (participants' subjective beliefs about the dose of the intervention, in this case, the strength of NIBS) affect performance scores over and beyond what is explained by *objective* (active/sham) *intervention*?; (2) Is there an interaction between *subjective intervention/dosage* and *objective intervention* in predicting performance scores?. To answer these questions, we conducted secondary data analysis on an open-access dataset of 150 individuals retrieved from a published paper by Filmer et al. (2019).

In the study that we examined, Filmer et al. (2019) tested the effect of tDCS over the prefrontal cortex on mind wandering. Mind wandering has been conceptualised as an executive function that occurs when attention shifts away from the primary task, and it can, therefore, lead to failures in performance and superficial representations of the external environment. Differently from most controlled processes, mind wandering occurs outside of awareness. Therefore, this mental capacity can be defined as a goal-driven process, albeit one that might not be directed

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.06.411850; this version posted December 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

towards the task at hand (Smallwood & Schooler, 2006, but see McVay & Kane, 2010). (). Noteworthy, mind wandering emerges as a critical cognitive process across the literature and is associated with multiple mental disorders, such as anxiety, depression and ADHD (Deng et al., 2014; Figueiredo et al., 2020).

In Filmer et al.'s experiment (2019), the authors showed that active 2mA cathodal tDCS reliably led to an increase in mind wandering scores compared to sham tDCS. In contrast, 1mA anodal, 1mA cathodal tDCS, and 1.5mA cathodal tDCS did not show a robust effect on mind wandering. Filmer et al.'s (2019) results are important, as they provide a potential explanation for the discrepancy in the literature between studies that did and did not find an effect of cathodal tDCS on mind wandering using a lower stimulation threshold (Axelrod et al., 2015; Axelrod et al., 2018; Boayue et al., 2020). Here, we show that two aspects of the study merit further attention. Firstly, the procedure to assess participants' blinding. Secondly, the contribution of adding *subjective intervention* as an explanatory variable to test whether the examined intervention, in this case, tDCS, has a genuine effect on mind wandering.

Methods

Participants and design

One hundred fifty healthy participants (age *M*=23, *SD*=5, 96 females) took part in this study. Participants were tested as part of a between-subject design. Each participant was randomly assigned to either one of the following five conditions: anodal 1 mA, cathodal 1 mA, 1.5 mA, 2 mA, or sham tDCS. The original study's analysis plan, sample size and methodology were all pre-registered on the OSF by the authors (https://osf.io/j6mqa/).

Procedure

The experiment consisted of three main parts. Firstly, participants were familiarised with the experimental paradigm. Secondly, participants were instructed to sit quietly with their eyes

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.06.411850; this version posted December 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

open and stimulation was applied offline to the left prefrontal cortex for 20 min. Lastly, participants performed a sustained attention task for 40 minutes, during which mind wandering, which was the main outcome of this research, was measured. Overall, each participant completed a single session, lasting approximately 1.5 hours.

Materials

Experimental task

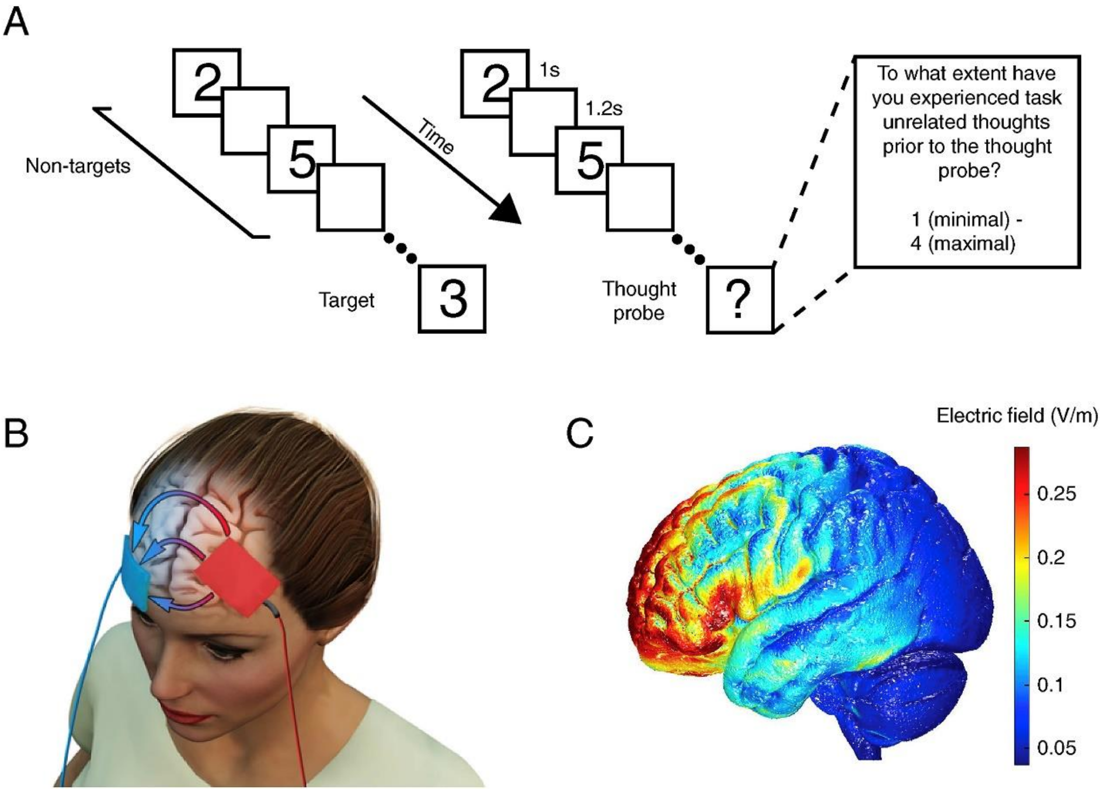
Participants completed a sustained attention task (SART) in which they were asked to respond via a keypress (space bar) to non-target stimuli (single digits excluding the number 3) and withholding responses to target stimuli (the number 3), see Figure 1a. Half of the trials ended in a target stimulus; the other half ended in a task unrelated thought (TUT) probe. The TUT probe asked: “To what extent have you experienced task unrelated thoughts prior to the thought probe? 1 (minimal) – 4 (maximal)”. Participants average response to the probe across trails was taken as a measure of mind wandering performance, with higher scores indicating higher mind wandering.

Objective intervention

Stimulation was delivered with a NeuroConn stimulator (neuroConn GmbH, Ilmenau, Germany). The target was placed over F3 (EEG 10–20 system), and the reference over the right orbitofrontal region (e.g., Figure 2b). For the four groups who received active stimulation (e.g., Figure 2c), tDCS lasted 20 minutes (including 30 s ramping up and down). During stimulation, participants were asked to sit quietly and keep their eyes open. The group who received sham stimulation had the same instructions but only received 15 s of constant current. The current was ramped up for 30 s up to 1.5 mA, and then ramped down for 30 s. Stimulation was single-blinded, meaning that while the participants were blind to the stimulation they received, the experimenters were aware of the participant’s stimulation group.

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.06.411850; this version posted December 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 1. **Experimental design, montage, and modelling of current distribution.** (a) Graphical representation of the sustained attention to response task (SART) for the target and thought probe trials; (b) electrode montage showed here for anodal stimulation targeting the left prefrontal cortex; (c) imaging of the amount of current induced using a standard head model, in the image, modelled current distribution for the 2 mA cathodal stimulation condition. Figure adapted from Filmer et al. (2019) with permission.



Subjective intervention

At the end of the experiment, participants were asked whether they thought they received active or sham stimulation (presented as a binary choice) via a short questionnaire. We here refer to the subject's judgment of whether they received active or sham stimulation as *subjective intervention*, in opposition to *objective intervention*, which indicates the actual type of stimulation that each subject received during the experiment. Moreover, at the end of the study, participants were also asked to guess which stimulation dosage they received, having to choose between the following options: none, weak, moderate, or strong. We here refer to the participants' judgment of stimulation dosage as *subjective dosage*.

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.06.411850; this version posted December 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Statistical Analysis

Statistical analysis was run using both R (version 4.0.1. for Windows) and JASP (version 0.13.1.0 for Windows) on the open-access dataset made available by Filmer et al. (2019). All the analyses were run on average mind wandering scores calculated over the whole experimental session. While Filmer et al. (2019) also examined the effect of tDCS on the first and second half of their experiment, for brevity we do not report such analyses as they are out of the scope of our research. However, when we ran such analyses, the results yielded a similar conclusion to our primary reported findings.

In R, we performed a chi-square test to assess blinding effectiveness by examining the relationship between stimulation guess (*subjective intervention*) and stimulation group (*objective intervention*).

Given that Filmer et al. (2019) used Bayesian ANOVA in their study, for consistency, we decided to employ the same statistical approach chosen by the authors. Therefore, in JASP with default priors, we performed two tests. Firstly, we reproduced Filmer et al.'s (2019) findings by running Bayesian ANOVAs with *objective intervention* as a between-subject factor, followed by the relevant post-hoc pairwise comparisons. Consequently, our primary analysis consisted of two Bayesian ANOVAs that were run to examine model fit. In the first analysis, *objective intervention* and *subjective intervention* were included as between-subject factors; while, in the second one, we examined model fit for *objective intervention* and *subjective dosage*. In both cases, average TUT ratings were taken as the outcome measure. If *subjective dosage* or *subjective intervention* led to better model fit, we separately analysed the effect of each of these factors on mind wandering. Moreover, for the *subjective dosage* factor, we conducted a trend analysis to examine whether the association between an increase in perceived stimulation dosage (from none to strong) and mind wandering scores is linear.

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.06.411850; this version posted December 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Following the original study by Filmer et al. (2019), we interpreted BF_{10} of 1–3 as anecdotal, BF_{10} of 3–10 as moderate, and $BF_{10}>10$ as strong evidence in favour of the alternate hypothesis (namely, that the examined model provided better fit than a null model). BF_{10} values ~1 were interpreted as providing no evidential value. Similarly, for support in favour of the null hypothesis, BF_{01} 1–3 was interpreted as anecdotal, BF_{01} of 3–10 as moderate, and $BF_{01}>10$ as strong evidence. Besides Bayesian statistics, we also reported results based on frequentist statistics for our primary analyses. In such cases, we accepted statistical significance for all tests at $p<.05$. Our choice to report both orthodox statistics, such as *F* and *P* values, alongside *Bayes* factors is based on literature that has highlighted the importance of addressing the relationship between significance testing and Bayesian analysis (Dienes & Mclatchie, 2018). We believe that reporting both statistical approaches will allow wider access to results' interpretability for researchers that may not be familiar with Bayesian methods.

Results

Blinding

We would like the reader first to judge the following two scenarios. In the first scenario, the correct guess rate (referring to the percentage of participants that successfully guessed their experimental condition) in the active group is 70%, while it is 40% in the sham group. In the second scenario, the correct guess rate is undisguisable between the active and sham group and is 70%. In which scenario was blinding more effective?

The second scenario is similar to what occurred in Filmer et al.'s experiment (2019). At the end of the experiment, participants were asked whether they thought they received active or sham stimulation. The reported correct guess rate was similar across all stimulation groups (ranging from 53% to 77%, **Table 1**). Given that the correct success rates did not appear to increase with higher stimulation intensity and based on their similarity between the active

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.06.411850; this version posted December 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

stimulation conditions and the sham condition (70%), the authors concluded that blinding was successful.

Table 1. *Correct guess rate for the judgment of active and sham stimulation as in Filmer et al. (2019) versus active stimulation guess rate per stimulation type.*

Stimulation type	***n***	**Correct guess rate**	**Active stimulation guess rate**
Anodal 1 mA	*30*	73%	73%
Cathodal 1 mA	*30*	53%	53%
Cathodal 1.5 mA	*30*	73%	73%
Cathodal 2 mA	*30*	57%	57%
Sham	*30*	70%	30%

Fallacious reasoning may lead to answering that blinding was effective in this study and in the second scenario we presented. However, we here show that Filmer et al.'s approach (2019) allows a justified conclusion regarding blinding effectiveness only when the guess rate is 50%. When the guess rate deviates from 50%, as in the current study, the fact that the correct guess rates are similar across conditions does not necessarily indicate that all groups are successfully blinded. On the contrary, in this case, the high correct guess rate indicates that the number of people in the sham condition who made an incorrect guess and thought they had received active stimulation was substantially lower (30%) than the number of people in the stimulation conditions that thought they had received active stimulation (53%-73%).

We analysed the potential differences in blinding success across all experimental groups using a chi-square test of independence. This test examined the association between the *active stimulation guess rate* and participants' experimental group (**Table 1**). Results showed that the association was significant, $\chi^2(4, N=150)=15.64$, $p=.004$, indicating that the feeling of

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.06.411850; this version posted December 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

receiving active stimulation was dependent upon the stimulation group to which participants were assigned.

To decompose what contributed to the overall association between the investigated chi-square measures, we further examined the standardised residuals. For people in the sham group, significantly more participants than expected by chance guessed to be receiving sham stimulation (z=2.29, p=0.022) and significantly fewer participants than expected guessed to be receiving active stimulation (z=-1.98, p= 0.048). From these findings, we concluded that, in contrast to the claim by Filmer et al. (2019), blinding was ineffective.

The effect of objective and subjective intervention on mind wandering

After examining blinding effectiveness, we tackled the question of whether *subjective intervention* or *subjective dosage* explained participants mind wandering scores over and beyond *objective intervention*. To address such a question, we employed Bayesian ANOVA as in Filmer et al. (2019), and we also report the results based on frequentist statistics.

Firstly, we reproduced the results from the original study by performing Bayesian ANOVA with *objective intervention* only as a between-subject factor. While the frequentist statistics resulted in a positive finding, the Bayesian ANOVA showed no evidential value for overall differences between all tDCS conditions in average TUT ratings over the whole experimental session (BF_{10}=0.984; F(4,145)=2.607, p=0.038). Comparing pairs of groups on average TUT ratings, we found anecdotal evidence for higher mind wandering in the 1.5 mA cathodal tDCS group compared to the sham group (BF_{10}=2.189, t(58)=2.272, p=.027), and moderate evidence for higher mind wandering in the 2 mA cathodal group (BF_{10}=7.436, t(58)=2.875, p=.006) compared to sham. No evidence emerged in support of 1mA anodal or cathodal stimulation, relative to sham, modulating TUTs (anodal stimulation: $BF_{10} < 1$, t(58)=1.619, p=.11; cathodal stimulation: $BF_{10} < 1$, t(58)=1.313, p=.19). These results successfully reproduced Filmer et al.’s (2019) findings.

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.06.411850; this version posted December 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Next, we run Bayesian ANOVA to examine the effect of *subjective intervention* and *subjective dosage* on mind wandering. *Subjective intervention*, alone, led to the best model fit in explaining mind wandering (**Table 2**). Hence, participants' beliefs about having received active or sham stimulation during the experiment were better at explaining mind wandering performance than *objective intervention* only, *objective* and *subjective intervention* together, or the interaction between *objective* and *subjective intervention*.

Table 2. Model comparison run using Bayesian ANOVA to predict average mind wandering scores

Models	P(M)	P (M| data)	BF_{10}	Error %
Subjective intervention	**0.200**	**0.464**	**3.374**	**8.94 e^{-8}**
Objective+Subjective intervention	0.200	0.205	1.492	2.201
Null model	0.200	0.137	1.000	
Objective	0.200	0.135	0.984	1.40 e^{-4}
*Objective*Subjective intervention*	0.200	0.059	0.429	1.000

As shown in **Figure 2**, people that thought they had received active stimulation during the experiment had higher mind wandering scores than people that thought they had received sham stimulation (BF_{10}=3.374, $t(148)=2.55$, $p=0.012$).

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.06.411850; this version posted December 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

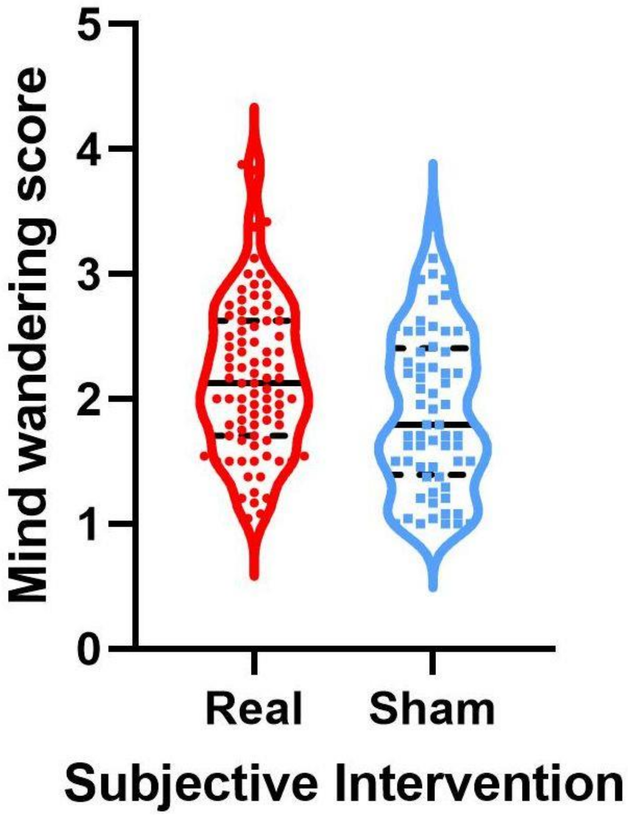


Figure 2. A violin plot showing the full distribution of the mind wandering score highlighting the differences as a function of subjective intervention. Solid and dotted black lines indicate the medians and quartiles, respectively.

Given that, at the end of Filmer et al.'s experiment (2019), participants were also asked to guess stimulation dosage (none, weak, moderate or strong), we re-run Bayesian ANOVA replacing *subjective intervention* with *stimulation dosage*. Our results showed that *subjective dosage*, alone, also led to the best model fit in explaining mind wandering (**Table 3**, BF_{10}=3.708). Therefore, we concluded that participants' beliefs about stimulation dosage during the experiment were better at explaining mind wandering performance than *objective intervention* only, *objective* and *subjective intervention* together, or the interaction between *objective* and *subjective intervention*.

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.06.411850; this version posted December 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Table 3. Model comparison run using Bayesian ANOVA to predict average mind wandering scores

Models	P(M)	P (M| data)	BF_{10}	Error %
Subjective dosage	**0.200**	**0.489**	**3.708**	**$1.715\ e^{-4}$**
Objective + Subjective dosage	0.200	0.219	1.658	0.587
Null model	0.200	0.132	1.000	
Objective	0.200	0.130	0.984	$1.402\ e^{-4}$
*Objective *Subjective dosage*	0.200	0.031	0.232	0.582

When running Bayesian ANOVA on the *subjective dosage* factor only, results showed that people that thought they had received no stimulation, weak, moderate or strong stimulation during the experiment presented significant differences in average mind wandering (BF_{10}=5.911, $F(3,146)=4.198$, $p=0.007$). This association between *subjective dosage* and mind wandering was best explained by a significant linear trend ($SE=0.13$, $t(146)=2.63$, $p=0.009$, **Figure 3**). Further analysis using a linear regression analysis showed that subjective dosage was a significant predictor of mind wandering (BF_{10}=24.26, β=.263, $t(148)=3.13$, p=.001). These results indicate that as the *subjective dosage* increased (from none to strong), mind wandering increased proportionally.

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.06.411850; this version posted December 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

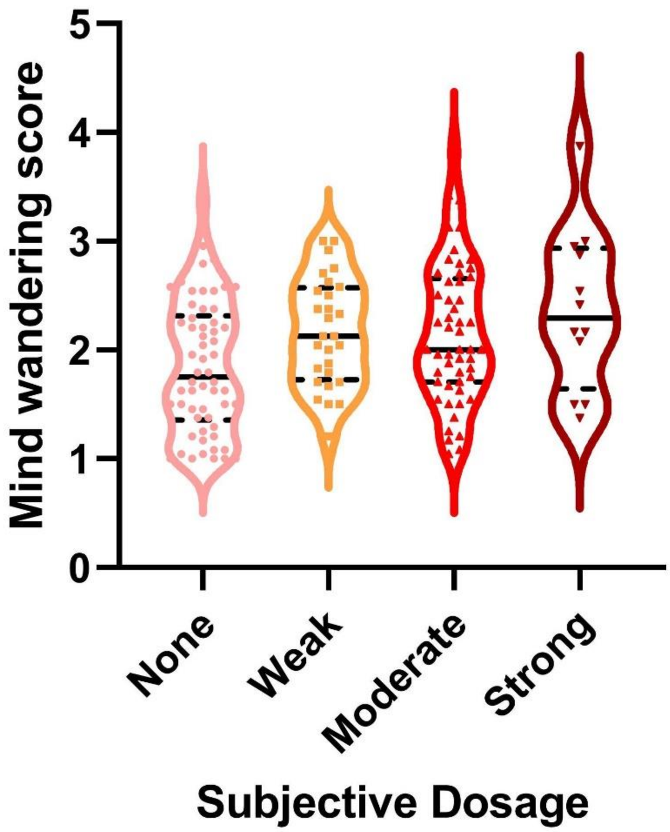


Figure 3. A violin plot showing the full distribution of the mind wandering score as a function of subjective dosage, showing an increase in mind wandering score as a function of subjective dosage. Solid and dotted black lines indicate the medians and quartiles, respectively.

Discussion

In the present study, we highlighted the importance of an effective procedure to assess blinding, as well as accounting for the individual subjective belief about receiving an intervention, in this case, NIBS.

Our results showed that in the experiment by Filmer et al. (2019), blinding of the sham condition was compromised. Notably, in that study participants received NIBS offline, before the task was delivered. The fact that participants could entirely focus on the sensations associated with the delivery of stimulation might have contributed to unblinding. Importantly, ineffective blinding can affect experimental outcomes in two ways. Firstly, it could mask positive or negative effects that the intervention may have on performance. Alternatively, it

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.06.411850; this version posted December 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

could inflate such effects. Therefore, it is of crucial importance to thoroughly account for its emergence.

We would like to highlight that the way in which Filmer et al. (2019) assessed the efficacy of their blinding is not uncommon in the literature (for recent clinical trials, e.g., Schecklmann et al., 2020; Sebastian et al., 2020). For instance, in a recent clinical trial that assessed the effect of NIBS on depression, and concluded its inefficacy (Schecklmann et al., 2020), the authors used the same procedure to assess blinding efficacy and wrote: “*In the group of patients, 39% (2 sham, 8 real treatment) subjects rated the treatment correctly” (p. 119)*. Again, the correct analysis of this blinding shows that there is a significant difference between the groups, and that blinding was ineffective ($\chi^2(1, N{=}36){=}4.12$, $p{=}.042$).

While highlighting the fallacy of using such procedure for assessing the success of blinding in interventions, the more notable contribution of our study is demonstrating how subjective beliefs about the type of intervention received explained the participants’ performance better than *objective intervention*, *subjective* and *objective intervention* together, or the interaction between *subjective* and *objective intervention*. Hence, in Filmer et al.’s study (2019), the fact that participants’ believed to have or have not received the real intervention affected their performance to a more considerable extent than the actual intervention to which they were assigned.

The same pattern of results emerged when we replaced *subjective intervention* with *subjective dosage*. In this case, a higher *subjective dosage* was linearly associated with higher mind wandering.

One question that emerges from this study is whether our results that were observed with self-report measures would apply to more objective behavioural outcomes and neural functions. In our view, the answer for this is likely to be positive, given that placebo effects have been shown

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.06.411850; this version posted December 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

to impact objective behavioural outcomes and neural activity (Hashmi, 2018; Schmidt et al., 2014; Oken et al., 2008). Independently of this possibility, we argue that the effect of *subjective intervention* on self-reported outcomes shall not be underestimated. Noteworthy, in randomised controlled trials that investigate the effect of NIBS on clinical and subclinical populations (e.g., depression, chronic pain, ADHD), some of which have also been approved by US Food and Drug Administration (FDA), the evidence is based on self-reported questionnaires. This consideration makes the case of *subjective intervention* even stronger, hinting to the potential role played by this factor in explaining experimental results across a variety of experiments.

In our view, it is a crucial responsibility of the relevant researchers to ensure that results from previous studies, especially those approved by the FDA, are not solely due *to subjective intervention*. Therefore, we strongly encourage the scientists in these studies to re-examine their data or make it publicly available in a format that will allow testing the effect of subjective intervention as in the present study. Of note, when we approached researchers in the field with a request to access their data, some of them could not share it due to restrictions by their ethical committee, while others reported that the assessment of blinding, aside from side effects, was not made.

As a point of improvement for the quality of research in the field of interventions, future experiments that investigate the effect of *subjective intervention* on performance should strive to record participants' expectations about the effect of stimulation thoroughly (Boot et al., 2013). Some participants may expect an intervention to improve their capabilities, while others could expect even the opposite, and the level of these expectations may vary as well. Accounting for participant's expectations would allow to more accurately model the directionality and strength of the effect of *subjective intervention* on the outcome of interest.

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.06.411850; this version posted December 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

In a nutshell, we showed that participants' subjective beliefs about receiving an intervention affect the primary outcome independently of the actual intervention condition to which participants are assigned. While we focused in this study on NIBS, our approach applies to other types of interventions (e.g., neurofeedback, pharmacological studies) that collect data on participant's subjective experience concerning whether they received a placebo or active intervention to assess blinding effectiveness. Our finding holds twofold importance. Firstly, it introduces two concepts in the intervention literature, namely: *subjective intervention* and *subjective dosage*. Secondly, it highlights the importance of assessing participants' blinding, independent of side effects. Unfortunately, different studies, including pre-registered reports, that claim for the efficacy or inefficacy of a given intervention do not measure or report data on the success of blinding (for examples in the field of NIBS see: Axelrod et al., 2015; Axelrod et al., 2018; Boayue et al., 2020; Horne et al., 2020). It is therefore impossible to draw strong conclusions from those studies, without knowing whether blinding was effective, and how beliefs and/or expectations played a role in inflating or masking the observed results. In this respect, it should be highlighted that reports on side effects are an unsuitable measure to assess blinding efficacy, as participants might subjectively feel being in the active intervention group, while not reporting side effects.

We call for future studies to systematically collect data on participants' subjective beliefs, as well as previous studies, to examine the potential effect of such beliefs on their results. Aside from estimating the contribution of subjective beliefs about belonging to the active or control condition, future research shall also strive to collect and analyse data on participants' expectations about the directionality and strength of the effect of *subjective intervention* on expected outcomes. Such data will allow to thoroughly examine the effect of subjective beliefs, yielding more valid and replicable results in order to progress scientific and clinical studies for

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.06.411850; this version posted December 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

the benefit of human wellbeing and cognition. Otherwise, the effect of those interventions, or the lack of it, might all only be in our head.

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.06.411850; this version posted December 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Acknowledgements

We appreciate the transparency that Hannah L. Filmer, Ashleigh Griffin, and Paul E. Dux provided by making their data accessible, without that the production of this work and progress of knowledge would not be possible.

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.06.411850; this version posted December 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

References

Ambrus, G. G., Al-Moyed, H., Chaieb, L., Sarp, L., Antal, A., & Paulus, W. (2012). The fade-in–short stimulation–fade out approach to sham tDCS–reliable at 1 mA for naive and experienced subjects, but not investigators. *Brain stimulation*, *5*(4), 499-504.

Ambrus, G. G., Paulus, W., & Antal, A. (2010). Cutaneous perception thresholds of electrical stimulation methods: comparison of tDCS and tRNS. *Clinical Neurophysiology*, *121*(11), 1908-1914.

André, S., Heinrich, S., Kayser, F., Menzler, K., Kesselring, J., Khader, P. H., ... & Mylius, V. (2016). At-home tDCS of the left dorsolateral prefrontal cortex improves visual short-term memory in mild vascular dementia. *Journal of the Neurological Sciences*, *369*, 185-190.

Axelrod, V., Rees, G., Lavidor, M., & Bar, M. (2015). Increasing propensity to mind-wander with transcranial direct current stimulation. *Proceedings of the National Academy of Sciences*, *112*(11), 3314-3319.

Axelrod, V., Zhu, X., & Qiu, J. (2018). Transcranial stimulation of the frontal lobes increases propensity of mind-wandering without changing meta-awareness. *Scientific reports*, *8*(1), 1-14.

Benedetti, F., Pollo, A., Lopiano, L., Lanotte, M., Vighetti, S., & Rainero, I. (2003). Conscious expectation and unconscious conditioning in analgesic, motor, and hormonal placebo/nocebo responses. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 23(10), 4315–4323.

Boayue, N. M., Csifcsák, G., Aslaksen, P., Turi, Z., Antal, A., Groot, J., ... & Mittner, M. (2020). Increasing propensity to mind-wander by transcranial direct current stimulation? A registered report. *European Journal of Neuroscience*, *51*(3), 755-780.

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.06.411850; this version posted December 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Boot, W. R., Simons, D. J., Stothart, C., & Stutts, C. (2013). The pervasive problem with placebos in psychology: Why active control groups are not sufficient to rule out placebo effects. *Perspectives on Psychological Science*, *8*(4), 445-454.

Brunoni, A. R., Sampaio-Junior, B., Moffa, A. H., Aparício, L. V., Gordon, P., Klein, I., ... & Valiengo, L. (2019). Non-invasive brain stimulation in psychiatric disorders: a primer. *Brazilian Journal of Psychiatry*, *41*(1), 70-81.

Davis, N. J., Gold, E., Pascual-Leone, A., & Bracewell, R. M. (2013). Challenges of proper placebo control for non-invasive brain stimulation in clinical and experimental applications. *European Journal of Neuroscience*, *38*(7), 2973-2977.

Deng, Y. Q., Li, S., & Tang, Y. Y. (2014). The relationship between wandering mind, depression and mindfulness. *Mindfulness*, *5*(2), 124-128.

Dienes, Z., & Mclatchie, N. (2018). Four reasons to prefer Bayesian analyses over significance testing. *Psychonomic Bulletin & Review*, *25*(1), 207-218.

Figueiredo, T., Lima, G., Erthal, P., Martins, R., Corção, P., Leonel, M., ... & Mattos, P. (2020). Mind-wandering, depression, anxiety and ADHD: Disentangling the relationship. *Psychiatry Research*, *285*, 112798.

Filmer, H. L., Dux, P. E., & Mattingley, J. B. (2014). Applications of transcranial direct current stimulation for understanding brain function. *Trends in neurosciences*, *37*(12), 742-753.

Filmer, H. L., Griffin, A., & Dux, P. E. (2019). For a minute there, I lost myself… dosage dependent increases in mind wandering via prefrontal tDCS. *Neuropsychologia*, *129*, 379-384.

Filmer, H. L., Mattingley, J. B., & Dux, P. E. (2020). Modulating brain activity and behaviour with tDCS: Rumours of its death have been greatly exaggerated. *Cortex*, *123*, 141-151.

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.06.411850; this version posted December 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Filmer, H.L., Griffin, A., Dux, P.E., (2019). Dosage Dependent Increases in Mind Wandering via Prefrontal tDCS. UQ eSpace. https://doi.org/10.14264/uql.2019.295.

Finniss, D. G., Kaptchuk, T. J., Miller, F., & Benedetti, F. (2010). Biological, clinical, and ethical advances of placebo effects. *The Lancet*, *375*(9715), 686-695.

Fonteneau, C., Mondino, M., Arns, M., Baeken, C., Bikson, M., Brunoni, A. R., ... & Poulet, E. (2019). Sham tDCS: A hidden source of variability? Reflections for further blinded, controlled trials. *Brain Stimulation*, *12*(3), 668-673.

Fregni, F., Boggio, P. S., Nitsche, M., Bermpohl, F., Antal, A., Feredoes, E., ... & Pascual-Leone, A. (2005). Anodal transcranial direct current stimulation of prefrontal cortex enhances working memory. *Experimental Brain Research*, *166*(1), 23-30.

Fritsch, B., Reis, J., Martinowich, K., Schambra, H. M., Ji, Y., Cohen, L. G., & Lu, B. (2010). Direct current stimulation promotes BDNF-dependent synaptic plasticity: potential implications for motor learning. *Neuron*, *66*(2), 198-204.

Gandiga, P. C., Hummel, F. C., & Cohen, L. G. (2006). Transcranial DC stimulation (tDCS): a tool for double-blind sham-controlled clinical studies in brain stimulation. *Clinical Neurophysiology*, *117*(4), 845-850.

Gruzelier, J. H. (2014). EEG-neurofeedback for optimising performance. III: a review of methodological and theoretical considerations. *Neuroscience & Biobehavioral Reviews*, *44*, 159-182.

Guerra, A., López-Alonso, V., Cheeran, B., & Suppa, A. (2020). Variability in non-invasive brain stimulation studies: reasons and results. *Neuroscience Letters*, *719*, 133330.

Hamilton, R., Messing, S., & Chatterjee, A. (2011). Rethinking the thinking cap: ethics of neural enhancement using non-invasive brain stimulation. *Neurology*, *76*(2), 187-193.

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.06.411850; this version posted December 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Hashmi, J. A. (2018). Placebo effect: theory, mechanisms and teleological roots. *International Review of Neurobiology*, *139*,233-253.

Horne, K. S., Filmer, H. L., Nott, Z. E., Hawi, Z., Pugsley, K., Mattingley, J. B., & Dux, P. E. (2020). Evidence against benefits from cognitive training and transcranial direct current stimulation in healthy older adults. *Nature Human Behaviour*.

Horvath, J. C., Forte, J. D., & Carter, O. (2015). Quantitative review finds no evidence of cognitive effects in healthy populations from single-session transcranial direct current stimulation (tDCS). *Brain Stimulation*, *8*(3), 535-550.

Javadi, A. H., & Cheng, P. (2013). Transcranial direct current stimulation (tDCS) enhances reconsolidation of long-term memory. *Brain Stimulation*, *6*(4), 668-674.

Kadosh, R. C. (Ed.). (2014). *The stimulated brain: cognitive enhancement using non-invasive brain stimulation*. Elsevier.

Kadosh, R. C., Levy, N., O'Shea, J., Shea, N., & Savulescu, J. (2012). The neuroethics of non-invasive brain stimulation. *Current Biology*, *22*(4), R108-R111.

Krause, B., & Kadosh, R. C. (2013). Can transcranial electrical stimulation improve learning difficulties in atypical brain development? A future possibility for cognitive training. *Developmental Cognitive Neuroscience*, *6*, 176-194.

Lampit, A., Hallock, H., & Valenzuela, M. (2014). Computerised cognitive training in cognitively healthy older adults: a systematic review and meta-analysis of effect modifiers. *PLoS medicine*, *11*(11), e1001756.

López-Alonso, V., Cheeran, B., Río-Rodríguez, D., & Fernández-del-Olmo, M. (2014). Inter-individual variability in response to non-invasive brain stimulation paradigms. *Brain Stimulation*, *7*(3), 372-380.

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.06.411850; this version posted December 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

May, A., Hajak, G., G€anßbauer, S., Steffens, T., Langguth, B., Kleinjung, T., & Eichhammer, P. (2007). Structural brain alterations following 5 days of intervention: dynamic aspects of neuroplasticity. *Cerebral Cortex*, 17, 205–210.

McVay, J. C., & Kane, M. J. (2010). Does mind wandering reflect executive function or executive failure? Comment on Smallwood and Schooler (2006) and Watkins (2008). *Psychological Bulletin, 136*(2), 188–197.

Medina, J., & Cason, S. (2017). No evidential value in samples of transcranial direct current stimulation (tDCS) studies of cognition and working memory in healthy populations. *Cortex*, *94*, 131-141.

Monti, A., Ferrucci, R., Fumagalli, M., Mameli, F., Cogiamanian, F., Ardolino, G., & Priori, A. (2013). Transcranial direct current stimulation (tDCS) and language. *Journal of Neurology, Neurosurgery & Psychiatry*, *84*(8), 832-842.

Nitsche, M. A., & Paulus, W. (2000). Excitability changes induced in the human motor cortex by weak transcranial direct current stimulation. *The Journal of physiology*, *527*(Pt 3), 633.

O’connell, N. E., Cossar, J., Marston, L., Wand, B. M., Bunce, D., Moseley, G. L., & De Souza, L. H. (2012). Rethinking clinical trials of transcranial direct current stimulation: participant and assessor blinding is inadequate at intensities of 2mA. *PloS One*, *7*(10), e47514.

Oken, B. S., Flegal, K., Zajdel, D., Kishiyama, S., Haas, M., & Peters, D. (2008). Expectancy effect: impact of pill administration on cognitive performance in healthy seniors. *Journal of Clinical and Experimental Neuropsychology*, *30*(1), 7-17.

Polania, R., Nitsche, M. A., & Ruff, C. C. (2018). Studying and modifying brain function with non-invasive brain stimulation. *Nature Neuroscience*, *21*(2), 174-187.

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.06.411850; this version posted December 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Santarnecchi, E., Brem, A.-K., Levenbaum, E., Thompson, T., Kadosh, R. C., & Pascual-Leone, A. (2015). Enhancing cognition using transcranial electrical stimulation. *Current Opinion in Behavioral Sciences*, 4, 171–178.

Sarkar, A., Dowker, A., & Kadosh, R. C. (2014). Cognitive enhancement or cognitive cost: trait-specific outcomes of brain stimulation in the case of mathematics anxiety. *Journal of Neuroscience*, *34*(50), 16605-16610.

Schecklmann, M., Nejati, V., Poeppl, T. B., Peytard, J., Rupprecht, R., Wetter, T. C., ... & Kreuzer, P. M. (2020). Bifrontal high-frequency transcranial random noise stimulation is not effective as an add-on treatment in depression. *Journal of Psychiatric Research*, *132*, 116-122.

Schmidt, L., Braun, E. K., Wager, T. D. & Shohamy, D. (2014). Mind matters: placebo enhances reward learning in parkinson's disease. *Nature Neuroscience* 17, 1793.

Sitaram, R., Ros, T., Stoeckel, L., Haller, S., Scharnowski, F., Lewis-Peacock, J., ... & Birbaumer, N. (2017). Closed-loop brain training: the science of neurofeedback. *Nature Reviews Neuroscience*, *18*(2), 86-100.

Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin*, *132*(6), 946.

Stefaniak, J. D., Halai, A. D., & Ralph, M. A. L. (2020). The neural and neurocomputational bases of recovery from post-stroke aphasia. *Nature Reviews Neurology*, *16*(1), 43-55.

Turi, Z., Bjørkedal, E., Gunkel, L., Antal, A., Paulus, W., & Mittner, M. (2018). Evidence for cognitive placebo and nocebo effects in healthy individuals. *Scientific Reports*, *8*(1), 1-14.

Turi, Z., Csifcsák, G., Boayue, N. M., Aslaksen, P., Antal, A., Paulus, W., ... & Thielscher, A. (2019). Blinding is compromised for transcranial direct current stimulation at 1 mA for 20 min in young healthy adults. *European Journal of Neuroscience*, *50*(8), 3261-3268.

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.06.411850; this version posted December 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Vicario, C. M., Salehinejad, M. A., Felmingham, K., Martino, G., & Nitsche, M. A. (2019). A systematic review on the therapeutic effectiveness of non-invasive brain stimulation for the treatment of anxiety disorders. *Neuroscience & Biobehavioral Reviews*, *96*, 219-231.

Wallace, D., Cooper, N. R., Paulmann, S., Fitzgerald, P. B., & Russo, R. (2016). Perceived comfort and blinding efficacy in randomised sham-controlled transcranial direct current stimulation (tDCS) trials at 2 mA in young and older healthy adults. *PloS One*, *11*(2), e0149703.

Wang, J., Wen, J. B., & Li, X. L. (2018). No effect of transcranial direct current stimulation of the dorsolateral prefrontal cortex on short-term memory. *CNS Neuroscience & Therapeutics*, *24*(1), 58-63.

Westwood, S. J., Olson, A., Miall, R. C., Nappo, R., & Romani, C. (2017). Limits to tDCS effects in language: failures to modulate word production in healthy participants with frontal or temporal tDCS. *Cortex*, *86*, 64-82.