Does Transcranial Direct Current Stimulation Improve Healthy Working Memory?: A Meta-analytic Review

Lauren E. Mancuso[1], **Irena P. Ilieva**[2], **Roy H. Hamilton**[1], and **Martha J. Farah**[1]

Abstract

■ Transcranial direct current stimulation (tDCS) has been reported to improve working memory (WM) performance in healthy individuals, suggesting its value as a means of cognitive enhancement. However, recent meta-analyses concluded that tDCS has little or no effect on WM in healthy participants. In this article, we review reasons why these meta-analyses may have underestimated the effect of tDCS on WM and report a more comprehensive and arguably more sensitive meta-analysis. Consistent with our interest in enhancement, we focused on anodal stimulation. Thirty-one articles matched inclusion criteria and were included in four primary meta-analyses assessing the WM effects of anodal stimulation over the left and right dorsolateral pFC (DLPFC) and right parietal lobe as well as left DLPFC stimulation coupled with WM training. These analyses revealed a small but significant effect of left DLPFC stimulation coupled with WM training. Left DLPFC stimulation alone also enhanced WM performance, but the effect was reduced to nonsignificance after correction for publication bias. No other effects were significant, including a variety of tested moderators. Additional meta-analyses were undertaken with study selection criteria based on previous meta-analyses, to reassess the findings from these studies using the analytic methods of this study. These analyses revealed a mix of significant and nonsignificant small effects. We conclude that the primary WM enhancement potential of tDCS probably lies in its use during training. ■

INTRODUCTION

Working memory (WM) refers to the ability to temporarily maintain and manipulate information in active awareness (Smith, 2001). WM is essential for performing many cognitive tasks (Wiley & Jarosz, 2012), and individual differences in WM capacity are correlated with individual differences in intelligence (Engle, Tuholski, Laughlin, & Conway, 1999). In part for this reason, the feasibility of enhancing WM has been explored with training (Melby-Lervåg & Hulme, 2013), stimulant medications (Ilieva, Hook, & Farah, 2015), and transcranial direct current stimulation (tDCS; Brunoni & Vanderhasselt, 2014). Here, we look more closely into the tDCS and WM enhancement literature.

Early reports of tDCS effects on WM suggest that enhancement is possible through anodal stimulation of the left dorsolateral pFC (DLPFC), a finding that makes sense given the tendency for anodal stimulation to enhance neuronal excitability (Nitsche et al., 2008; Priori, 2003; Nitsche & Paulus, 2000) and the role of the left DLPFC in WM (Carpenter, Just, & Reichle, 2000). In an early study of 15 healthy volunteers, Fregni and colleagues (2005) found a significant effect of left DLPFC anodal stimulation on WM performance in a 3-back task. This was followed by other small studies assessing left DLPFC stimulation on WM in patient groups (see Nitsche et al., 2008) and healthy populations (e.g., Ohn et al., 2008). The literature has continued to grow and has prompted three recent attempts to synthesize the available findings by meta-analysis.

The earliest meta-analysis, combining tDCS and repetitive TMS studies of patients and healthy individuals, focused on the *n*-back task to operationalize WM (Brunoni & Vanderhasselt, 2014). In the *n*-back task, participants monitor a sequence of stimuli and must respond when the current stimulus is identical to the one presented *n* stimuli back. Meta-regression showed significant WM improvements in speed but not accuracy with tDCS. As most of the tDCS studies were done in healthy participants (16/19), these results provide some indication that tDCS probably enhances healthy WM performance as measured by speed of responding, but not accuracy. However, this meta-analytic review included a mix of different stimulation montages, such that the effects of anodal stimulation of the left and right DLPFC were grouped together. Further limiting the informativeness of this analysis with regard to WM, only one type of WM task was included. Performance on other WM tasks, such as digit span or the Sternberg working memory scanning task, was not examined.

A second meta-analysis surveyed the literature on tDCS and WM, along with other cognitive abilities; included multiple WM tasks; and concluded that tDCS has no reliable effect on WM (Horvath, Forte, & Carter, 2015). However, this meta-analysis has been criticized on the

[1]University of Pennsylvania, [2]Weill Cornell Medical College, New York, NY

Journal of Cognitive Neuroscience 28:8, pp. 1063–1089
doi:10.1162/jocn_a_00956

grounds that its design would make it difficult to find positive effects of tDCS whether or not such effects exist. Among the contentious design decisions were: limiting the analysis to tasks that had been studied in relation to tDCS by more than one laboratory, reducing the number of eligible studies, and applying inconsistent criteria for selecting dependent variables to meta-analyze when more than one is available (Chhatbar & Feng, 2015; Nitsche, Bikson, & Bestmann, 2015; Price & Hamilton, 2015; see also Horvath, 2015; Price, McAdams, Grossman, & Hamilton, 2015).

Most recently, Hill, Fitzgerald, and Hoy (2016) reported the results of meta-analyses covering multiple WM tasks in healthy participants and in neuropsychiatric patients. Two separate meta-analyses were carried out on data from healthy participants, focused on reported effects on RTs and accuracies. A drawback of this study is that, of the 34 studies entered into the meta-analyses of tDCS effects on healthy WM, only 12 different samples of participants were tested. Therefore, contrary to the assumptions of meta-analysis, many of the effect sizes entered into the meta-analyses were not independent of one another. The conclusion drawn from this study was that tDCS has small but significant enhancing effects on WM, whether measured by RT or accuracy.

In summary, despite early evidence that tDCS can enhance WM (Andrews, Hoy, Enticott, Daskalakis, & Fitzgerald, 2010; Ohn et al., 2008; Fregni et al., 2005), recent meta-analyses have concluded that the effects are reliable though small (Hill et al., 2016), partial (Brunoni & Vanderhasselt, 2014), or nonexistent (Horvath et al., 2015). The present meta-analysis addresses the effect of tDCS on WM with methods better suited to finding an enhancement effect in healthy people, if it exists, than the previous meta-analyses. Relative to the broad aggregation of Brunoni and Vanderhasselt (2014), who mixed studies of tDCS in healthy and psychiatric populations, we only evaluated WM effects in tDCS studies of healthy participants. Relative to the narrow selectivity of Brunoni and Vanderhasselt (2014), who excluded tasks other than the *n*-back, and Horvath and colleagues (2015), who excluded tasks not reported by multiple laboratories, we included all published data available on WM performance and tDCS in healthy adults.

Five other differences from the earlier meta-analyses would be expected to increase the sensitivity of our analysis in comparison with the three preceding meta-analyses: First, for studies that employed the *n*-back task, we excluded 0-back and 1-back conditions, which place little demand on WM (Braver et al., 1997) and have been used as control conditions in other WM research (Ragland et al., 2002; Carlson et al., 1998). Second, we included results from researchers who had not reported the information necessary to estimate a relevant effect size in their published article, using email requests and measurements of published figures as described under Coding procedures. By obtaining this additional information, we were able to expand the pool of evidence and reduce the influence of publication bias. Third, by excluding studies for which active and sham stimulation sessions were not counterbalanced, we increased the quality of the analyzed research. Fourth, we selected the most appropriate dependent variable from each study and combined them meta-analytically, following Ilieva et al. (2015). As discussed in more detail in the Methods section, under Dependent variables, we used a priori criteria to select dependent variables, rather than separately meta-analyzing accuracy and RT measures and selecting the specific measure of accuracy or RT emphasized by the authors of the original study. Fifth, the fact that the previous three meta-analyses were completed before ours gives our analysis access to later published studies, making it the most comprehensive to date.

Four main issues are addressed by this meta-analysis. First, we ask: For three commonly used anodal stimulation sites, does tDCS have an effect on WM performance in healthy adults, and if so, how large is this effect? Second, we ask: Are certain tDCS setups and contexts more effective for WM enhancement than others? What factors, including reference electrode placement, current density, stimulation before or during task performance, and so forth, moderate WM enhancement by tDCS? Third, we address the issue of tDCS as an adjuvant to WM training: Does tDCS amplify the enhancing effects of WM training, as might be expected given its effects on neuronal excitability and synaptic plasticity (Stagg, 2014; see also Santarnecchi et al., 2015)? Fourth, what role might publication bias play in shaping the literature on tDCS enhancement of WM, and how do the conclusions of that literature differ when the influence of publication bias is estimated and corrected?

METHODS

Literature Search

Online databases PubMed and PsychInfo were searched through December 2014 with the key words *transcranial direct current stimulation* or *tDCS*, combined with each of the following: *working memory, n-back, Sternberg, span,* or *cognition*. The reference sections of relevant reviews and reports were also searched for eligible studies. Articles available on journal sites ahead of print publication were included.

Eligibility Criteria

Publication Type and Language

Empirical investigations in any report format were eligible for inclusion in the meta-analysis. Research on nonhuman participants, qualitative studies, and nonempirical publications (e.g., review articles, meta-analyses, case studies, commentary pieces, articles on modeling methods) were excluded. Empirical studies that only evaluated the effects of other brain stimulation techniques such as TMS or transcranial alternating current stimulation were also excluded at this level. Only reports published in English were included.

Participants

Eligible participants were healthy adults, aged 18 years and older. Research on participants with a history of mental illness, neurological disease, stroke, brain injury, or disorders of consciousness was excluded. Studies that evaluated the effects of tDCS on sleep-deprived participants were also excluded.

Research Design

Studies with double-blind or single-blind, sham-controlled designs were included. For studies using within-participant designs, the order of stimulation sessions was required to be counterbalanced. Because the goal of the analysis was to assess the enhancing potential of tDCS, we focused on anodal stimulation. Although undoubtedly a simplification of the reality linking brain stimulation and cognition, anodal stimulation is thought to increase excitability of cortical neurons (Nitsche et al., 2008; Priori, 2003; Nitsche & Paulus, 2000); cathodal stimulation is generally thought of as being inhibitory and has produced less reliable effects on cortical excitability and behavioral outcomes (Batsikadze, Moliadze, Paulus, Kuo, & Nitsche, 2013; Jacobson, Koslowsky, & Lavidor, 2012).

Cognitive Construct

Eligible studies assessed WM, the ability to temporarily store and manipulate information in the service of other ongoing cognitive functions (Smith, 2001), as operationalized by the *n*-back task, Sternberg task, digit span task, letter number sequencing task, paced auditory serial addition task (PASAT), complex WM span task, operation span (OSPAN) task, symmetry span task, change detection task, delayed WM task, internal shift task, visual STM task, sequential presentation task, and the Corsi block-tapping test.

Dependent Variables

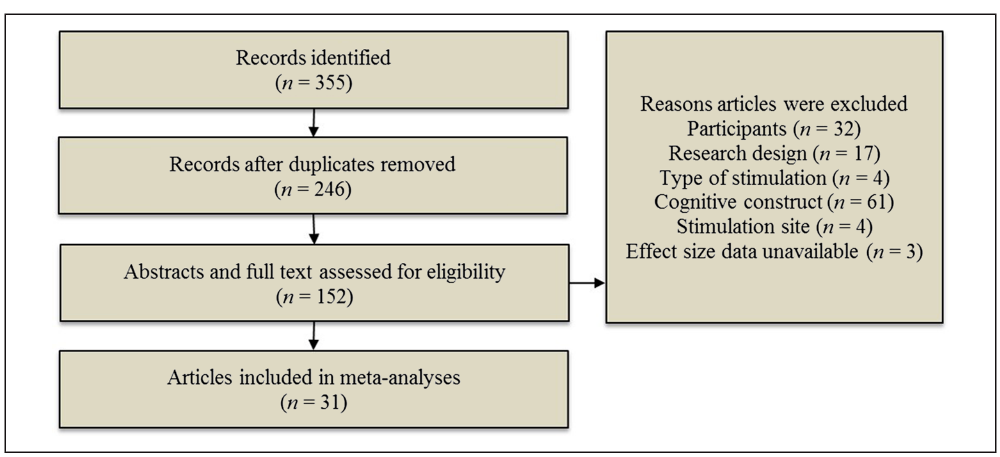
The types and number of dependent variables reported for each task varied across reports. Flexibility in choice of primary dependent variable can predispose to bias (Ioannidis, 2005), so we adopted a priori criteria for selecting the dependent variables to use in the meta-analysis following Ilieva and colleagues (2015). Accuracy measures were favored over RT measures unless overall performance was determined to be at ceiling. When measures of performance are near ceiling, they are less sensitive to manipulations and would therefore underestimate the effects of tDCS. For present purposes, a measure was considered to be at risk of ceiling effects if the smaller of the means was within 1 *SD* of the maximum of the measurement scale. If some, but not all, performance measures were at ceiling (e.g., lower trial loads or easy problem types), data were used for trial types not at ceiling.

Conversely, if accuracy rates are not high, then RT measures are problematic. This is because the RT distribution, even if limited to correct trials, will reflect performance from mixture of different processes: correct use of WM and "lucky" guesses, the latter equal in proportion to the wrong responses when chance performance is 50% (Sternberg, 1998). For studies reporting multiple non-ceiling accuracy-based measures, measures such as d' or Cowan's *K*, which combine hit and false alarm rates, were preferred. If unavailable, overall accuracy followed by hits (accuracy for "yes" trials) was used. For WM span task performance, accuracy was used if available, otherwise longest length correct. This procedure does not introduce bias into the analysis, as would selecting the dependent measure that showed an effect. Furthermore, meta-analysis does not require a common dependent variable (e.g., Rosenthal & Rubin, 1986).

Determining Study Eligibility

The search process, summarized in Figure 1, led to the identification of 355 titles, which were narrowed down

Figure 1. Process for determining study eligibility.



to 246 after 109 duplicate articles were removed. After screening the abstracts of these articles, an additional 94 reports were excluded because they did not report empirical research on tDCS (89 reports) or were not written in English (five reports).

The full text of the remaining 152 articles was reviewed. An additional 32 reports failed to meet the criteria because they looked at the effects of stimulation in depression or affective disorders (n = 11), stroke (n = 5), schizophrenia (n = 5), Alzheimer's disease and mild cognitive impairment (n = 3), traumatic brain injury (n = 2), Parkinson's disease (n = 1), aphasia (n = 1), posttraumatic stress disorder (n = 1), chronic pain (n = 1), amyotrophic lateral sclerosis (n = 1), or following sleep deprivation (n = 1). Seventeen studies did not meet our research design criteria. Of these, 13 studies did not have sham conditions, three studies with within-participant designs did not counterbalance the order of sham and stimulation conditions, and one study evaluated the effects of tDCS delivered intermittently for 15 sec every 15 sec. Four reports were excluded because they measured the effects of cathodal stimulation only. An additional 61 studies were excluded because they measured cognitive constructs outside the scope of the present review (e.g., inhibitory control, episodic memory, creativity, intelligence, motor performance) or evaluated the effects of tDCS on cigarette smoking (n = 2) or alcohol dependence (n = 1). A single article on 1-back, a task that places minimal demand on WM, was also excluded at this level. An additional four investigations were excluded because of too few studies with comparable electrode montages to allow for meta-analysis (active anodal electrode position: cerebellum, n = 2; left parietal, n = 1; occipital, n = 1). Three articles were excluded because the data needed to compute effect sizes were not published, and authors did not respond to our requests.

Disagreements at any level of study selection were resolved by consensus after discussion by the authors of this article. Thirty-one articles were included in the meta-analysis.

Coding Procedures

Coded variables included means and standard deviations for performance under anodal and sham stimulations, sample size, dependent variable, and effect direction. For studies with multiple subgroups of qualifying participants we collapsed results across subgroups for a single measure of effect for healthy, nonsleep-deprived adults. Moderating variables were also coded for two of our main meta-analyses, involving left DLPFC stimulation and left DLPFC stimulation coupled with WM training. All studies were coded by the first author. The second author independently coded a random sample of 30% of reports included in the meta-analyses. Analyses of reliability showed excellent agreement (two-way mixed-model intraclass correlation coefficient for absolute agreement > 0.99 in all cases).

Effect sizes were calculated using means and standard deviations. These descriptive statistics provide less biased estimates of effect size, compared with inferential statistics in repeated-measures designs (Dunlap, Cortina, Vaslow, & Burke, 1996). As noted earlier, when unavailable from reports, we sought to obtain these data in other ways. First, we requested them by email from authors. Relevant data were obtained from 9 of the 12 researchers contacted. Second, when numerical information was not available but could be measured from graphs, means and standard deviations or standard errors were estimated from published figures using Image J software (imagej.nih.gov/ij/).

The following moderators were coded:

1. Current density (low vs. high): Current density was calculated based on the current intensity (mA) and size (cm^2) of the active electrode(s). Current density was dichotomized into "low" ($\leq$0.05 mA/cm^2) and "high" (>0.05 mA/cm^2).
2. Reference electrode position (cephalic vs. extracephalic): To determine if the placement of the reference electrode (cathode) influenced the effect of anodal stimulation on WM performance, we grouped the studies into those using cephalic locations on the scalp and those using extracephalic reference locations (e.g., deltoid muscle, contralateral cheek).
3. Stimulation duration (short vs. long): The duration of active stimulation was dichotomized between "short" ($\leq$15 min) and "long" (>15 min).
4. Timing of stimulation relative to task (online vs. offline): We compared designs in which WM tasks took place entirely or mostly "online," that is, during active stimulation, or mostly "offline," that is, mostly after active stimulation.
5. Task type (n-back vs. other): We investigated whether the most widely used, studied WM task, the n-back, was more or less affected by tDCS than the other tasks that have been used.
6. Length of training period: For training studies only, the length of training period was dichotomized into single day and multiday (2–10 days).
7. Transfer: For training studies only, the use of the same task for training and test versus a different (transfer) task was coded.

Statistical Methods

Effect Size Metrics

Effect sizes were calculated for all studies using Hedges' g (Hedges, 1981). The conventions used to interpret Hedges' g are similar to Cohen's d such that effect sizes of 0.2, 0.5, and 0.8 are considered to reflect small, moderate, and large effects, respectively (Cohen, 1988). Hedges' g is calculated by multiplying the effect size Cohen's d by a coefficient J, which corrects for the tendency for studies with small sample sizes to bias the mean effect size positively because of publication bias: $J = 1 - \left(\frac{3}{4 \times df - 1}\right)$. In combining effect sizes, each effect size was weighed by the inverse of the squared standard error.

The most straightforward measure of enhancement is the difference in performance attributable to stimulation, scaled by the standard deviation of the sample's sham performance. This method addresses the question, "How far along the distribution of normal performance does tDCS stimulation push participants?" The difference attributable to stimulation was calculated by taking the difference between stimulation and sham performance or, when baseline performance data were available, the difference between stimulation–baseline and sham–baseline. Most included investigations had within-participant designs, which allow effect sizes to be calculated in a second way as well, scaling performance change by units of variability of change. This addresses the question, "How much of a stimulation-related benefit can one expect, relative to the variability of change scores in the sample?" Both types of effect size analyses are reported here, with primary emphasis placed on the first type.

In our primary analyses, within- and between-participant studies were included. The formulas used, typically employed for between-participant designs, were modified so that the observed standard deviations in the sham condition were entered for both anodal and sham stimulation conditions. In particular:

$$g = J \times \frac{M_{\mathrm{ANODAL}} - M_{\mathrm{SHAM}}}{SD_{\mathrm{POOLED}}}$$

where $SD_{\mathrm{POOLED}} = \sqrt{\frac{\left((N_{\mathrm{ANODAL}} - 1)SD^2_{\mathrm{SHAM}} + (N_{\mathrm{SHAM}} - 1)SD^2_{\mathrm{SHAM}}\right)}{N_{\mathrm{ANODAL}} + N_{\mathrm{SHAM}} - 2}}$ and $SE = SD_{\mathrm{POOLED}} \times \sqrt{\frac{1}{N_{\mathrm{ANODAL}}} + \frac{1}{N_{\mathrm{SHAM}}}}$

Our secondary analyses focused on the change score (anodal minus sham) for within-participant designs, specifically the average benefit associated with anodal stimulation, relative to variability of change within the sample. Hedges' g was calculated differently for these designs using the following formula:

$$g = J \times \left(\frac{(M_{\mathrm{ANODAL}} - M_{\mathrm{SHAM}}) \times \left(\sqrt{2(1 - Corr)}\right)}{SD_{\mathrm{DIFF}}}\right)$$

where $SD_{\mathrm{DIFF}} = \sqrt{SD^2_{\mathrm{ANODAL}} + SD^2_{\mathrm{SHAM}} - 2Corr SD_{\mathrm{ANODAL}} SD_{\mathrm{SHAM}}}$ and $SE = \frac{SD_{\mathrm{DIFF}}}{\sqrt{N}}$

The computation of the effect size of change scores requires the correlation between participants' performance in the stimulation and sham conditions, and these correlations are not included in published reports. We therefore set the value of the correlation to .5 in the analyses to be reported but also performed the analyses with $r = .2$ and $r = .8$ to assess the dependence of results on the assumed correlation.

Handling of Studies with More than One Effect Size

In a meta-analysis, effect sizes can be assumed to be independent if each effect size comes from an independent study sample (Lipsey & Wilson, 2001). If multiple effect sizes are included for the same participant sample, the between-study variance will be underestimated, and as a result, the significance of the overall effect will be overestimated. We therefore limited the number of effect sizes per study sample to 1. This was accomplished by coding the effect sizes from the available data for multiple conditions, such as different current levels, different WM loads, different relevant trial types, different time points, or different tasks, and then averaging the effect sizes. When data were split for analysis within the same sample based on education, WM capacity, or performance, effect sizes were also averaged across groups.

If more than one WM task was evaluated in the same experiment and these tasks were both performed either online or offline, effect sizes were averaged to compute one overall effect. If one task was completed online and another was completed offline, results from the first were included in the online stimulation meta-analysis, and results from the second, performed after a WM task during stimulation, were included in the stimulation with WM training meta-analysis. For example, if a 3-back task was performed during stimulation and a Sternberg task was performed immediately after, separate effect sizes would be included in meta-analyses evaluating the effects of stimulation (on the 3-back task) and effects of stimulation coupled with WM training (on the Sternberg task).

Fixed versus Random Effects Model

A fixed effects model assumes that sampling error is the only source of effect size variability, whereas a random effects model assumes that sampling error and between-study variability are potential sources of effect size variability. Effect sizes were estimated using a random effects model because of the variability between individual studies (different stimulation montages, strengths, and durations; measures of WM; and time relative to stimulation that WM was measured) and because we wanted to generalize the findings beyond the examined research.

Estimation of Heterogeneity

Studies are heterogeneous if they differ from one another more than would be expected by the random error of sampling participants, evident in within-study error variance. Heterogeneity of effect sizes was assessed using the Q statistic and the I^2 index. A significant Q statistic indicates that the studies being meta-analyzed are not all of a kind. The I^2 index is an estimate of between-study variance as a percentage of the total variance. I^2 values of 25, 50, and 75 reflect low, medium, and high levels of heterogeneity, respectively (Lipsey & Wilson, 2001).

Moderator Analysis

One of the goals of this meta-analysis is to discover what factors influence the effectiveness of tDCS for the enhancement of WM. We approach this goal with moderator analysis, which tests specific factors for their roles in moderating the effect size. This differs from the more intuitive practice of simply testing different sets of studies separately and reporting the two different significance levels or effect sizes, as Hill et al. (2016) did for online and offline stimulations. In the case of moderator analyses, one can determine not only whether different conditions have different effects but also whether that difference is itself reliable.

Moderator analyses are typically conducted only if significant heterogeneity is found. However, lack of heterogeneity can emerge from either the absence of significant moderation or the presence of two or more moderators whose effects cancel each other out. We therefore planned to conduct moderator analyses regardless of heterogeneity results for the primary analyses. Because lack of a significant moderation effect, like any other effect, would be expected when sample sizes are very small, moderation analyses were only conducted when at least 10 studies were available to analyze.

The effects of the dichotomous moderators described earlier were examined using mixed effects analyses. This type of analysis assumes that effect size variation is because of a combination of systematic associations between moderators and effect sizes, random differences between studies, and participant-level sampling error.

Some moderator analyses were complicated by studies having more than one level of a moderator in a single study. These included analyses of current density, task timing (online vs. offline), and task type (n-back vs. other). This occurred when more than one current density was evaluated, when WM performance was measured at multiple time points within the same study, or when a single study evaluated performance on multiple tasks, one of which was the n-back. These analyses were approached in two ways. First, to satisfy the assumption of independence between effect sizes, we excluded, from moderator analyses, studies that had data for more than one level of any moderator variable. In a separate second analysis, we employed the shifting unit method, in which the same study is allowed to contribute to each level of the moderator (Cooper, 2010). Whereas the first approach leaves meta-analysis assumptions unviolated, the latter approach makes use of all available data. The findings based on both approaches were in agreement, and only data based on the latter approach are reported here.

Publication Bias

Significant findings are more likely to be published than null results, and as a result, the literature may not represent the true set of research findings (Rothstein, Sutton, & Borenstein, 2006). Three methods were used to assess publication bias: funnel plots, trim-and-fill procedure, and fail-safe N (Lipsey & Wilson, 2001). These analyses were conducted without correcting for the factor J, which itself serves to correct for publication bias.

The funnel plot is a qualitative, visual method for assessing publication bias by plotting study effect sizes against standard error (the inverse of study precision). The lower the precision, the greater the dispersion of effect sizes around the true value, making the shape of the scatterplot look like a funnel, if publication bias is absent. If publication bias is present, the funnel plot is negatively skewed, with missing points in the lower left part of the plot.

The trim-and-fill procedure calculates an unbiased estimate of the effect size in case of publication bias. Outliers on the funnel plot, which indicate extreme positive effects, are identified from the analysis, and a mirror image data point is imputed on the left side of the funnel plot. The corrected data are used to obtain an unbiased effect size estimate.

Finally, the fail-safe N indicates the number of studies with a zero effect size that, if added to the analysis, would render the mean effect size nonsignificant. Publication bias is unlikely if the fail-safe N is large relative to the number of studies meta-analyzed. A commonly used threshold is $5k + 10$, where k is the number of meta-analyzed studies (Rothstein et al., 2006).

Test for Outliers

To prevent extreme findings from biasing our results, we tested for outliers ≥ 3 *SD* above or below the mean of all eligible effect sizes within each separate meta-analysis. As noted below, no study met this outlier criterion in any of the analysis.

Software

All analyses were performed using Comprehensive Meta-Analysis 3.0 software (Englewood, NJ).

RESULTS

Overview of Results

Meta-analyses investigating the effects of tDCS on WM were conducted for four anodal stimulation montages: left DLPFC stimulation, right DLPFC stimulation, right parietal lobe stimulation, and left DLPFC stimulation coupled with WM training.

Three additional meta-analyses were carried out, applying our coding and analytic methods to studies selected by the criteria of Brunoni and Vanderhasselt (2014), Horvath and colleagues (2015), and Hill and colleagues (2016), to better understand the relation between the results of the present analysis and theirs.

For each group of studies, two measures of effect size were calculated and reported, as explained in the section on Effect size metrics. The effect size reported as the primary result refers to the amount by which tDCS would be expected to enhance WM measured against variability in normal (nonstimulated) WM performance. We also report the amount by which tDCS enhances WM relative to the variability of change associated with tDCS. The outcomes were similar in all cases. To assess the influence on the latter, change-based, measure of assumptions concerning the correlations between the repeated measures of within-participant designs, we compared the results obtained with different assumptions. Compared with results obtained assuming that $r = .5$, results assuming $r = .2$ and $r = .8$ were very similar. Of the 12 change effect sizes computed with these alternate r values, the largest deviation in value of Hedges' g was 0.05. Therefore, we only report effect sizes calculated based on an imputed correlation of .5 between repeated measures.

Using the main effect size measure, we also report heterogeneity and three assessments of publication bias. In no case did heterogeneity or publication bias results differ qualitatively for the second effect size measure, so we do not report them here. In the absence of significant heterogeneity, we conducted moderator analyses when there were at least 10 studies to analyze, specifically left DLPFC stimulation and left DLPFC stimulation with WM training. There were no outliers identified in any of the meta-analyses.

The Effect of Left DLPFC Stimulation on WM

The 23 studies shown in Table 1 examined the effect of left DLPFC anodal stimulation on WM. Meta-analysis of effect size relative to normal variability indicated a small but significant effect of stimulation on WM: Hedges' $g = 0.17$, 95% confidence interval (CI) [0.03, 0.30] (Figure 2). When effect size was measured relative to the variability of gain scores, the result was similar: Hedges' $g = 0.15$, 95% CI [0.05, 0.26]. There was no evidence for heterogeneity: $Q(22) = 3.80, p > .99, I^2 = 0.00$, or for moderation by any of the factors examined (all ps $> .51$).

The funnel plot revealed a slightly negative skew for this set of studies, represented by the open circles in Figure 3, suggestive of publication bias. Consistent with this, the trim-and-fill procedure trimmed 6 data points. After correction by trim-and-fill, with the imputed effect sizes shown as solid circles, the effect was reduced to a nonsignificant trend with the 95% CI only barely crossing from positive effect sizes to zero, Cohen's $d = 0.12$, 95% CI [−0.001, 0.25]. The fail-safe N procedure indicated that only 14 unpublished studies with an effect size of zero at the time of analysis (i.e., not captured by the present analysis) would nullify the significance level of the uncorrected results. Taken together, the skew of the funnel plot, the reduction of effect size with the trim-and-fill procedure, and the low fail-safe N results suggest the need for caution in interpreting the small enhancement effect found here.

The Effect of Right DLPFC Stimulation on WM

The eight studies of Table 2 examined the effect of right DLPFC anodal stimulation on WM. No effect was found in the primary analysis: Hedges' $g = 0.04$, 95% CI [−0.19, 0.27] (Figure 4). Similar results were obtained for effect size measured relative to the variability of gain scores: Hedges' $g = 0.07$, 95% CI [−0.11, 0.26]. There was no evidence of heterogeneity: $Q(7) = 2.17, p = .95, I^2 = 0.00$. The funnel plot, shown in Figure 5, appears slightly skewed, and the trim-and-fill procedure trimmed 3 data points, causing the average effect size to become nonsignificantly negative, Cohen's $d = -0.05$, 95% CI [−0.25, 0.14]. Although the relatively small number of studies calls for caution, the present analysis suggests that anodal tDCS of the right DLPFC does not enhance WM.

The Effect of Right Parietal Stimulation on WM

The seven studies shown in Table 3 examined the effect of right parietal stimulation on WM. No significant effect was found in the primary analysis, Hedges' $g = 0.17$, 95% CI [−0.09, 0.44] (Figure 6), or when effect sizes were measured relative to the variability of gain scores, Hedges' $g = 0.16$, 95% CI [−0.06, 0.38]. There was no evidence of heterogeneity: $Q(6) = 5.22, p = .52, I^2 = 0.00$. The funnel plot, shown in Figure 7, does not appear negatively skewed, and no points were trimmed for the trim-and-fill procedure. In summary, we do not find evidence that anodal right parietal stimulation enhances WM. However, the small number of studies analyzed prevents strong conclusions, and the effect size was similar in magnitude to the effect shown in the left DLPFC stimulation analysis.

The Effect of Left DLPFC Stimulation and Training on WM

Ten studies, shown in Table 4, examined the effect of WM training accompanied by tDCS over left DLPFC on subsequent WM performance, compared with WM training with sham stimulation. Included under the rubric of training studies are any that assess WM performance after performing at least one training session, with the training carried out with left DLPFC stimulation or sham.

The primary analysis indicated a small but significant effect of stimulation on training benefit: Hedges' $g = 0.29$, 95% CI [0.06, 0.52] (Figure 8). This was also found when effect size was measured relative to change scores: Hedges' $g = 0.30$, 95% CI [0.08, 0.52]. There was no evidence of heterogeneity, $Q(9) = 1.46, p > .99, I^2 = 0.00$, or of moderation by any of the factors examined all (ps $> .34$).

Table 1. Left DLPFC Stimulation on WM: Study Characteristics and Effect Sizes

Study	*N (Anode/Sham)*	*Mean Age (y)*	*Age Range (y)*	*% Male*	*Stimulation Site (Anode)*	*Reference Site (Cathode)*	*Reference Site (Cephalic)*	*Current (mA)*	*Current Density (mA/cm²)*	*Duration (min)*	*Test Timing*	*Test*	*Measure Coded*	*Design*	*Ceiling Effects?*	*Hedges' g*
Berryhill and Jones (2012)	25	63.7	56–80	Not reported	F3	Right cheek	No	1.5	0.04	10	Offline	2-back	Accuracy	Within-participant	No	0.031
Fregni et al. (2005)	15	20.2	19–22	26.7	F3	CSA	Yes	1	0.03	10	Online	3-back	Accuracy	Within-participant	No	0.335
Gill, Shah-Basak, and Hamilton (2015), Experiment 1	11	21.8	18–25	72.7	F3	CSA	Yes	2	0.08	20	Online	3-back	Accuracy	Within-participant	No	0.139
Gill et al. (2015), Experiment 2	11	19.8 (*N* = 12)	18–25 (*N* = 12)	58.3 (*N* = 12)	F3	CSA	Yes	2	0.08	20	Offline	PASAT	Accuracy	Within-participant	Yes	0.175
Gladwin, den Uyl, Fregni, and Wiers (2012)	14	22	Not reported	42.9	F3	CSA	Yes	1	0.03	10	Offline	Sternberg	Accuracy	Within-participant	Yes	0.305
Hoy et al. (2013)	18	24.7	Not reported	38.9	F3	CSA	Yes	1	0.03 0.06	20	Offline	2-back 3-back	Accuracy	Within-participant	No	0.123
Jeon and Han (2012)	16 (8/8)	39.5/36.5	20–59 (*N* = 32)	37.5/37.5	F3	CSA	Yes	1	0.03	20	Offline	Digit span forward Digit span backward	Accuracy	Between-participant	No	0.185
Jones, Gözenman, and Berryhill (2015), Experiment 1	24	23.8	Not reported	50	F3/F7	CSA	Yes	1.5	0.04	10	Offline	Change detection	Accuracy	Within-participant	No	0.224
Jones et al. (2015), Experiment 2	20	22.0	Not reported	40	F3/F7	CSA	Yes	1.5	0.04	10	Offline	Change detection	Accuracy	Within-participant	No	0.249
Keeser et al. (2011)	10	28.9	Not reported	50	F3	CSA	Yes	2	0.06	20	Offline	2-back	Accuracy	Within-participant	No	0.442
Keshvari, Pouretemad, and Ekhtiari (2013)	30	22.3	Not reported	50	F3	F4	Yes	2	0.08	20	Offline	2-back	Accuracy	Within-participant	No	0.148

Lally, Nord, Walsh, and Roiser (2013)	21 (10/11)	23.1	Not reported	33.3	F3	Right cheek	No	1	0.03	10	Online Offline	3-back	Accuracy	Between-participant	No	0.292
Martin et al. (2013)	42 (21/21)	23.1/23.2	Not reported	57.1/61.9	F3	Right deltoid	No	2	0.06	30	Online	Adaptive dual *n*-back	Accuracy	Between-participant	No	0.159
Meiron and Lavidor (2013)	25 (14/11)	24.9/23.1	18–36 ($N = 41$)	50/36.4	F3	Cz	Yes	2	0.13	15	Online	Modified *n*-back	Accuracy	Between-participant	Yes	0.315
Mulquiney, Hoy, Daskalakis, and Fitzgerald (2011)	10	29.4	Not reported	40	F3	CSA	Yes	1	0.03	10	Online	Sternberg	Accuracy	Within-participant	No	0.048
Mylius et al. (2012)	12	25.1	20–25 ($N = 24$)	50	F3	CSA	Yes	2	0.06	20	Offline	2-back	RT	Within-participant	Yes	0.110
Nozari and Thompson-Schill (2013)	24	Not reported	19–30	45.8	F3	F4	Yes	1.5	0.06	20	Online	2-back 3-back	Accuracy	Within-participant	No	0.248
Ohn et al. (2008)	15	26.5	Not reported	33.3	F3	CSA	Yes	1	0.04	30	Online Offline	3-back	Accuracy	Within-participant	No	0.302
Richmond, Wolk, Chein, and Olson (2014)	40 (20/20)	20.7/20.7	18–30	35/35	F3	F4	Yes	1.5	0.04	15	Online	Complex WM span	Accuracy	Between-participant	No	0.312
Tadini et al. (2011)	22	Not reported	18–64	63.6	F3	CSA	Yes	1.3	0.04	30	Offline	Digit span forward Digit span backward	Accuracy	Within-participant	No	−0.196
Teo, Hoy, Daskalakis, and Fitzgerald (2011)	12	Not reported	Not reported	Not reported	F3	CSA	Yes	1	0.03 0.06	20	Online	3-back	Accuracy	Within-participant	No	−0.003
Vanderhasselt, Brunoni, Loeys, Boggio, and De Raedt (2013)	32	22.3	18–36	37.5	F3	CSA	Yes	2	0.06	20	Online	Internal shift task	RT	Within-participant	Yes	0.041
Wolkenstein and Plewnia, 2012	22	31.9	Not reported	22.7	F3	Right deltoid	No	1	0.03	20	Online	Delayed WM	Accuracy	Within-participant	No	0.191

Study	Hedges' g	Lower Limit	Upper Limit	SE
Berryhill and Jones (2012)	0.031	−0.516	0.578	0.279
Fregni et al. (2005)	0.335	−0.367	1.036	0.358
Gill et al. (2014), Experiment 1	0.139	−0.666	0.944	0.411
Gill et al. (2014), Experiment 2	0.175	−0.631	0.981	0.411
Gladwin et al. (2012)	0.305	−0.420	1.029	0.369
Hoy et al. (2013)	0.123	−0.518	0.764	0.327
Jeon and Han (2012)	0.185	−0.744	1.114	0.474
Jones et al. (2015), Experiment 1	0.224	−0.339	0.786	0.287
Jones et al. (2015), Experiment 2	0.249	−0.366	0.863	0.313
Keeser et al. (2011)	0.442	−0.409	1.293	0.434
Keshvari et al. (2013)	0.148	−0.353	0.648	0.255
Lally et al. (2013)	0.292	−0.541	1.125	0.425
Martin et al. (2013)	0.159	−0.435	0.754	0.303
Meiron and Lavidor (2012)	0.315	−0.454	1.083	0.392
Mulquiney et al.(2011)	0.048	−0.793	0.889	0.429
Mylius et al. (2011)	0.110	−0.663	0.883	0.394
Nozari and Thompson-Schill (2013)	0.248	−0.312	0.807	0.286
Ohn et al. (2008)	0.302	−0.401	1.005	0.359
Richmond et al. (2014)	0.312	−0.300	0.924	0.312
Tadini et al. (2011)	−0.196	−0.786	0.395	0.301
Teo et al. (2011)	−0.003	−0.775	0.770	0.394
Vanderhasselt et al. (2013)	0.041	−0.447	0.529	0.249
Wolkenstein and Plewnia (2013)	0.191	−0.391	0.773	0.297

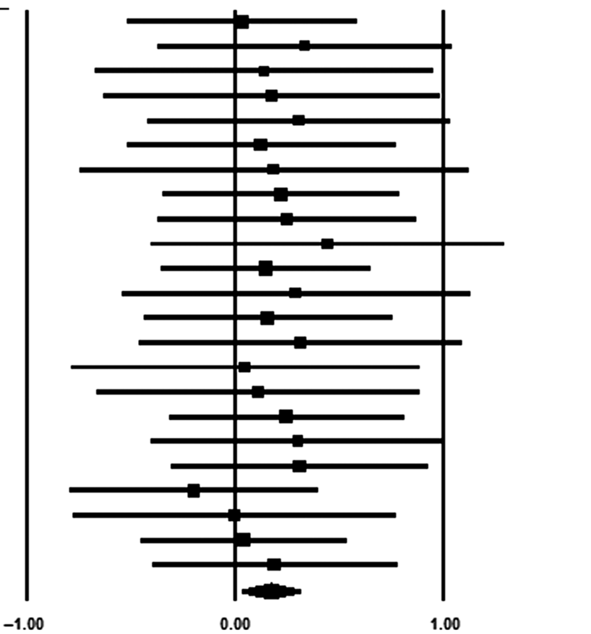


Figure 2. Forest plot: left DLPFC anodal stimulation.

The funnel plot, shown in Figure 9, appears slightly skewed, and the trim-and-fill procedure trimmed 2 data points. After adding back these points and imputing the missing points, the effect size was reduced but remained significant: Cohen's $d = 0.25$, 95% CI [0.04, 0.46]. However, at the time of analysis, the fail-safe *N* procedure indicated that merely seven unpublished studies with an effect size of zero added to the 10 studies analyzed here would nullify the effect. In summary, there is evidence that tDCS enhances the effects of training on WM. However, it would not take a large number of null results to eliminate the effect.

Reanalysis of Anodal Left DLPFC Stimulation Modeled on Brunoni and Vanderhasselt (2014)

Brunoni and Vanderhasselt (2014) meta-analyzed the literature on noninvasive brain stimulation and *n*-back performance, including both TMS and tDCS, and multiple sites of stimulation. Here, we attempt to relate the present results to theirs by meta-analyzing the effect of left anodal DLPFC tDCS on *n*-back performance, thus focusing on the stimulation site that appears most promising. The 14 studies included in this analysis are shown in Table 5.

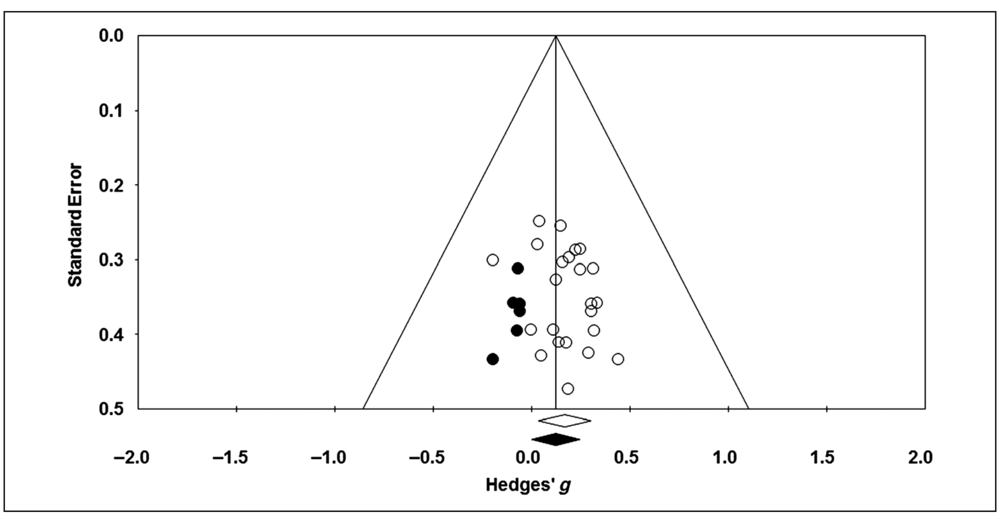


Figure 3. Funnel plot of publication bias: left DLPFC anodal stimulation. Dark data points represent studies imputed by the trim-and-fill procedure.

Table 2. Right DLPFC Stimulation on WM: Study Characteristics and Effect Sizes

Study	*N (Anode/Sham)*	*Mean Age (y)*	*Age Range (y)*	*% Male*	*Stimulation Site (Anode)*	*Reference Site (Cathode)*	*Reference Site (Cephalic)*	*Current (mA)*	*Current Density* (mA/cm^2)	*Duration (min)*	*Test Timing*	*Test*	*Measure Coded*	*Design*	*Ceiling Effects?*	*Hedges' g*
Berryhill and Jones (2012)	25	63.7	56–80	Not reported	F4	Left cheek	No	1.5	0.04	10	Offline	2-back	Accuracy	Within-participant	No	0.122
Berryhill, Peterson, Jones, and Stephens (2014)	18 (low ASRS)	Not reported	18–37 (*N* = 36)	38.9 (*N* = 36)	F4/Fz	Left cheek	No	1.5	0.04	10	Offline	2-back OSPAN	Accuracy	Within-participant	No	0.006
Bona and Silvanto (2014)	15	25.1	Not reported	46.7	F4	CSA	Yes	2	0.06	20	Offline	VSTM	Accuracy	Within-participant	No	0.018
Jeon and Han (2012)	16 (8/8)	35.1/37.9	20–59 (*N* = 32)	50/62.5	F4	CSA	Yes	1	0.03	20	Offline	Digit span forward Digit span backward	Accuracy	Between-participant	No	−0.029
Keshvari et al. (2013)	30	21.2	Not reported	50	F4	F3	Yes	2	0.08	20	Offline	2-back	Accuracy	Within-participant	No	−0.220
Meiron and Lavidor (2013)	27 (16/11)	25.4/23.1	18–36 (*N* = 41)	50/36.4	F4	Cz	Yes	2	0.13	15	Online	Modified *n*-back	Accuracy	Between-participant	Yes	0.372
Mylius et al. (2012)	12	23.5	20–25 (*N* = 24)	16.7	F4	CSA	Yes	2	0.06	20	Offline	2-back	RT	Within-participant	Yes	0.236
Wu et al. (2014)	20	26	24–31	60	F4	Left cheek	No	1.5	0.06	15	Offline	CBT	Accuracy	Within-participant	No	0.076

CBT = Corsi block-tapping test; VSTM = visual STM.

Study	Hedges' g	Lower Limit	Upper Limit	SE
Berryhill and Jones (2012)	0.122	−0.425	0.668	0.279
Berryhill et al. (2014)	0.006	−0.635	0.647	0.327
Bona and Silvanto (2014)	0.018	−0.682	0.719	0.358
Jeon and Han (2012)	−0.029	−0.960	0.902	0.475
Keshvari et al. (2013)	−0.220	−0.721	0.281	0.256
Meiron and Lavidor (2012)	0.372	−0.379	1.123	0.383
Mylius et al. (2011)	0.236	−0.540	1.011	0.396
Wu et al. (2014)	0.076	−0.536	0.688	0.312

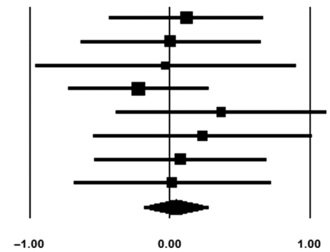


Figure 4. Forest plot: right DLPFC anodal stimulation.

Unlike their analysis, we focus on tDCS in healthy participants, include only 2-back and greater WM loads (including modified and adaptive *n*-back tasks), and select the performance measure (speed or accuracy) according to whether or not accuracy is susceptible to a ceiling effect, as in the first four meta-analyses reported here. Because their study was published in 2014, we were also able to include more recently published studies. In this way, we present an analysis of the portion of the literature Brunoni and Vanderhasselt (2014) were interested in, after applying our more sensitive analytic approach.

Our main analyses indicated a small but significant effect of stimulation on WM: Hedges' $g = 0.20$, 95% CI [0.02, 0.38] (Figure 10). When measured relative to the variability of gain scores, the effect size was estimated to be similarly small and significant: Hedges' $g = 0.19$, 95% CI [0.04, 0.33]. There was no significant evidence of heterogeneity: $Q(13) = 1.73$, $p > .99$, $I^2 = 0.00$.

The funnel plot, shown in Figure 11, appears slightly skewed, and the trim-and-fill procedure trimmed 2 data points. After restoring these points and adding the imputed missing points, the effect size was reduced but remained only barely significant: Cohen's $d = 0.18$, 95% CI [0.002, 0.35]. The fail-safe *N* procedure indicated that only four studies with an effect size of zero at the time of analysis would nullify the effect. In summary, whereas Brunoni and Vanderhasselt's analysis suggested that tDCS enhanced the speed but not the accuracy of performing the *n*-back task, the present results allow us to draw a potentially more general conclusion supporting a small but significant effect of anodal left pFC stimulation on *n*-back performance, albeit an effect that could easily be reduced to nonsignificance by a relatively small number of null results.

Reanalysis of Anodal Left DLPFC Stimulation Modeled on Horvath et al. (2015)

Horvath and colleagues (2015) meta-analyzed the literature on tDCS and a wide variety of cognitive tests. Here, we apply our analytic methods to the WM tasks that fit the task selection criteria of these authors, that is, that tasks must have been used by more than one laboratory. We also included studies published too recently to have been included in their meta-analysis and other studies that appear to meet their criteria. The 16 studies included in this analysis are shown in Table 6.

Our main analyses indicated a small effect that only missed significance: Hedges' $g = 0.16$, 95% CI [−0.01, 0.34] (Figure 12). Relative to the variability of gain scores,

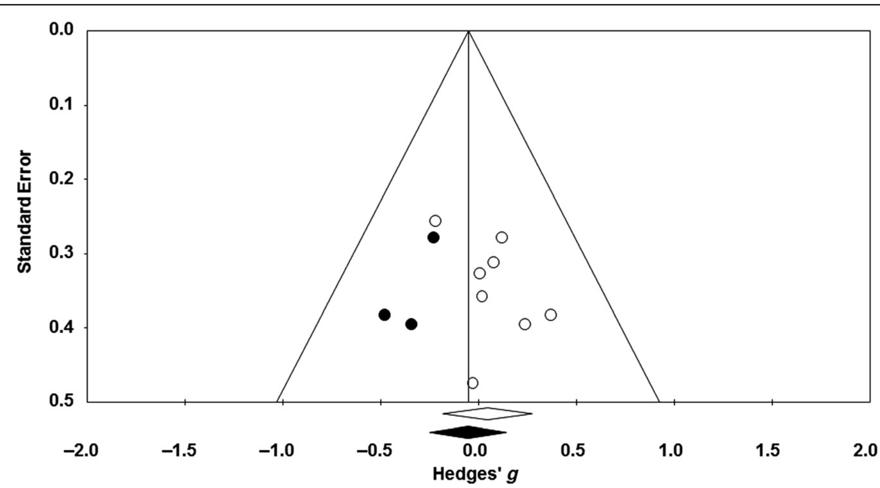


Figure 5. Funnel plot of publication bias: right DLPFC anodal stimulation. Dark data points represent studies imputed by the trim-and-fill procedure.

Table 3. Right Parietal Stimulation on WM: Study Characteristics and Effect Sizes

Study	*N (Anode/ Sham)*	*Mean Age (y)*	*Age Range (y)*	*% Male*	*Stimulation Site (Anode)*	*Reference Site (Cathode)*	*Reference Site (Cephalic)*	*Current (mA)*	*Current Density (mA/cm²)*	*Duration (min)*	*Test Timing*	*Test*	*Measure Coded*	*Design*	*Ceiling Effects?*	*Hedges' g*
Berryhill, Wencil, Coslett, and Olson (2010)	11	25	Not reported	54.5	P4	Left cheek	No	1.5	0.04	10	Offline	Sequential presentation	Accuracy	Within-participant	No	−0.335
Hsu, Tseng, Liang, Cheng, and Juan (2014)	20	22	Not reported	35	P4	Left cheek	No	1.5	0.09	15	Offline	Change detection	Accuracy	Within-participant	No	0.262
Jones and Berryhill (2012), Experiment 1	20	23.25	Not reported	40	P4	Left cheek	No	1.5	0.04	10	Offline	Sequential presentation Change detection	Accuracy	Within-participant	No	−0.028
Jones and Berryhill (2012), Experiment 2	28	22.29	Not reported	14.3	P4	Left cheek	No	1.5	0.04	10	Offline	Sequential presentation Change detection	Accuracy	Within-participant	No	0.427
Sandrini, Fertonani, Cohen, and Miniussi (2011)	18 (9/9)	25	20–30	55.6/55.6	P4	P3	Yes	1.5	0.04	13	Offline	2-back	Accuracy	Between-participant	No	−0.346
Tseng et al. (2012), Experiment 1	10	21	Not reported	50	P4	Left cheek	No	1.5	0.09	15	Offline	Change detection	Accuracy	Within-participant	No	0.572
Tseng et al. (2012), Experiment 2	20	22	Not reported	35	P4	Left cheek	No	1.5	0.09	15	Offline	Change detection	Accuracy	Within-participant	No	0.363

Study	Hedges' g	Lower Limit	Upper Limit	SE
Berryhill et al. (2010)	−0.335	−1.149	0.479	0.415
Hsu et al. (2014)	0.262	−0.360	0.884	0.317
Jones and Berryhill (2012), Experiment 1	−0.028	−0.637	0.581	0.311
Jones and Berryhill (2012), Experiment 2	0.427	−0.098	0.952	0.268
Sandrini et al. (2012)	−0.346	−1.234	0.541	0.453
Tseng et al. (2012), Experiment 1	0.572	−0.286	1.430	0.438
Tseng et al. (2012), Experiment 2	0.363	−0.517	1.242	0.449

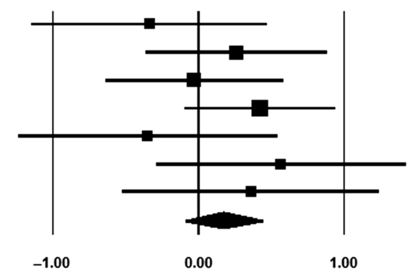


Figure 6. Forest plot: right parietal anodal stimulation.

the effect size was similarly small but did reach significance: Hedges' $g = 0.16$, 95% CI [0.03, 0.29]. There was no evidence of heterogeneity: $Q(15) = 3.06$, $p > .99$, $I^2 = 0.00$.

The funnel plot, shown in Figure 13, appears slightly skewed. The trim-and-fill procedure trimmed five studies, reducing the estimated effect size to Cohen's $d = 0.10$, 95% [−0.05, 0.26]. Fail-safe N is not reported because the overall effect size was not significant according to our main analysis.

Reanalysis of Anodal Left DLPFC Stimulation Modeled on Hill et al. (2016)

Hill and colleagues (2016) meta-analyzed the literature on anodal tDCS on WM tasks. Here, we apply our analytic methods to the WM tasks that fit the task selection criteria of these authors, specifically the n-back (excluding 1-back), Sternberg, and digit span tasks. Similar to the previous analyses modeled on different meta-analyses, we included more recently published studies and other studies that appear to meet their criteria. The 16 studies included in this analysis are shown in Table 7. Coincidentally, the criteria of Hill and colleagues netted the same set of studies as those selected by Horvath and colleagues' criteria. We derived a single effect size from studies in which multiple tasks were performed by the same set of participants. In addition, because Hill et al. (2016) reported finding effects only for offline stimulation, we also separately analyzed the effects of online and offline stimulations using our main effect size measure.

Our main analysis found a small effect that only missed significance: Hedges' $g = 0.16$, 95% CI [−0.01, 0.34] (Figure 14). Relative to the variability of gain scores, the effect size was similarly small but did reach significance: Hedges' $g = 0.16$, 95% CI [0.03, 0.29].

Contrary to expectation, given the earlier group's finding of significant effects for offline but not online stimulation, there was no evidence of heterogeneity: $Q(15) = 3.06$, $p > .99$, $I^2 = 0.00$. Separately analyzing the effects of offline ($n = 13$) and online ($n = 7$) tDCS yielded nonsignificant effects in both cases: Hedges' $g = 0.15$, 95% CI [−0.04, 0.34], and Hedges' $g = 0.18$, 95% CI [−0.09, 0.46], respectively.

Figure 7. Funnel plot of publication bias: right parietal anodal stimulation.

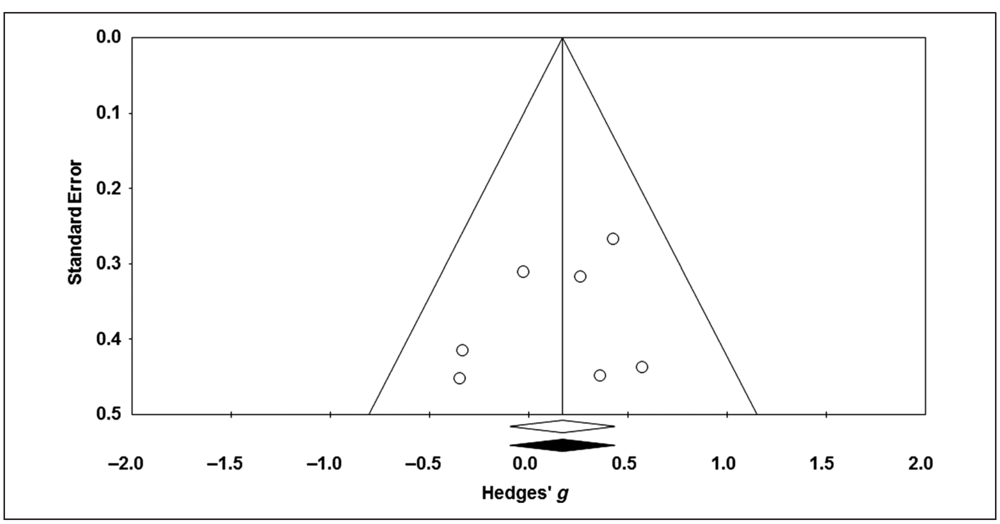


Table 4. Left DLPFC Stimulation and Cognitive Training on WM: Study Characteristics and Effect Sizes

Study	*N (Anode/ Sham)*	*Mean Age (y)*	*Age Range (y)*	*% Male*	*Stimulation Site (Anode)*	*Reference Site (Cathode)*	*Reference Site (Cephalic)*	*Current (mA)*	*Current Density (mA/cm²)*	*Duration (min)*	*Training Test*	*Test*	*Training (Days)*	*Measure Coded*	*Design*	*Ceiling Effects?*	*Hedges' g*
Andrews et al. (2010)	11 (11/10)	28.1	20–51	40	F3	CSA	Yes	1	0.03	10	2-back	Digit span forward	1	Accuracy	Within-participant	No	0.327
											3-back	Digit span backward					
Gill et al. (2015), Experiment 1	11	21.8	18–25	72.7	F3	CSA	Yes	2	0.08	20	3-back	PASAT	1	Accuracy	Within-participant	Yes	0.565
Lally et al. (2013)	21 (10/11)	23.09	Not reported	33.3	F3	Right cheek	No	1	0.03	10	3-back	3-back	2	Accuracy	Between-participant	No	0.227
Martin et al. (2013)	42 (21/21)	23.1/23.2	Not reported	57.1/61.9	F3	Right deltoid	No	2	0.06	30	Adaptive dual *n*-back	Digit span forward	10	Accuracy	Between-participant	No	0.251
												Digit span backward					
												Letter number sequencing					
Meiron and Lavidor (2013)	25 (14/11)	24.9/23.1	18–36 ($N = 41$)	50/36.4	F3	Cz	Yes	2	0.13	15	Modified *n*-back	Poststimulation modified *n*-back	1	Accuracy	Between-participant	No	0.322
Mulquiney et al. (2011)	10	29.4	Not reported	40	F3	CSA	Yes	1	0.03	10	Sternberg	2-back	1	RT	Within-participant	Yes	0.399
Ohn et al. (2008)	15	26.5	Not reported	33.3	F3	CSA	Yes	1	0.04	30	3-back	3-back	1	Accuracy	Within-participant	No	0.428
Park, Seo, Kim, and Ko (2014)	40 (20/20)	70.1/69.4	>65	35/30	F3 & F4	Nondominant arm	No	2	0.08	30	CACT	2-back	10	Accuracy	Between-participant	No	0.358
												Digit span forward					
Richmond et al. (2014)	40 (20/20)	20.7/20.7	18–30	35/35	F3	F4	Yes	1.5	0.04	15	Complex WM SPAN	OSPAN	10	Accuracy	Between-participant	No	0.089
												Symmetry span					
Teo et al. (2011)	12	Not reported	Not reported	Not reported	F3	CSA	Yes	1	0.03	20	3-back	Sternberg	1	Accuracy	Within-participant	No	0.074
									0.06								

Study	Hedges' g	Lower Limit	Upper Limit	SE
Andrews et al. (2011)	0.327	−0.502	1.155	0.423
Gill et al. (2014), Experiment 1	0.565	−0.256	1.387	0.419
Lally et al. (2013)	0.227	−0.616	1.069	0.430
Martin et al. (2014)	0.251	−0.345	0.847	0.304
Meiron and Lavidor (2012)	0.322	−0.458	1.103	0.398
Mulquiney et al. (2011)	0.399	−0.449	1.248	0.433
Ohn et al. (2008)	0.428	−0.277	1.133	0.360
Park et al. (2014)	0.358	−0.257	0.973	0.314
Richmond et al. (2014)	0.089	−0.519	0.697	0.310
Teo et al. (2011)	0.074	−0.699	0.848	0.395

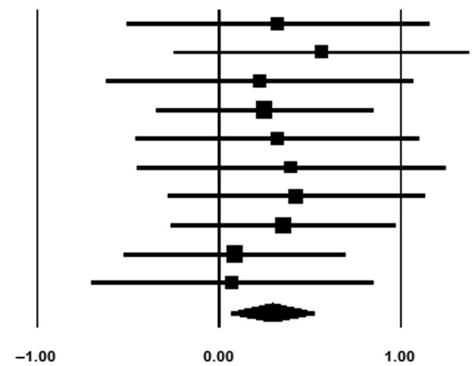


Figure 8. Forest plot: left DLPFC anodal stimulation and WM training.

The funnel plot, shown in Figure 15, appears slightly skewed. The trim-and-fill procedure trimmed five studies, reducing the estimated effect size to Cohen's $d = 0.10$, 95% $[-0.05, 0.26]$. Fail-safe *N* is not reported because the overall effect size was not significant according to our main analysis.

Summary of Findings

The clearest effect of tDCS on WM comes from its use in WM training. Left DLPFC anodal stimulation during training improved subsequent WM performance to a small but significant degree, and the effect remained significant after correction for publication bias. That the clearest support for tDCS in WM enhancement comes from its use with training makes sense in light of its known effects on cellular and synaptic physiology (Stagg, 2014), and recent discussions of cognitive enhancement with tDCS have emphasized its potential to enhance learning (e.g., Santarnecchi et al., 2015). However, it should be borne in mind that this conclusion comes from a relatively small number of studies (10), and an even smaller number of unreported null results (7) would eliminate the effect.

It also seems possible that left DLPFC anodal stimulation enhances WM performance independent of training, although the evidence is somewhat equivocal. The effect was small but significant, with evidence of publication bias. After attempting to correct for this bias using the trim-and-fill procedure, the effect became nonsignificant. Although a relatively large number of studies went into the analysis of this effect (23), the number of null results needed to eliminate the effect is disconcertingly small in proportion (14). When the selection criteria of three recent meta-analyses were replicated and analyzed with the methods used here, left DLPFC anodal stimulation was found to produce a small but significant effect when adopting Brunoni and Vanderhasselt's (2014) focus on the *n*-back task and small, near-significant effects using Horvath et al.'s (2015) and Hill et al.'s (2016) approaches to study selection.

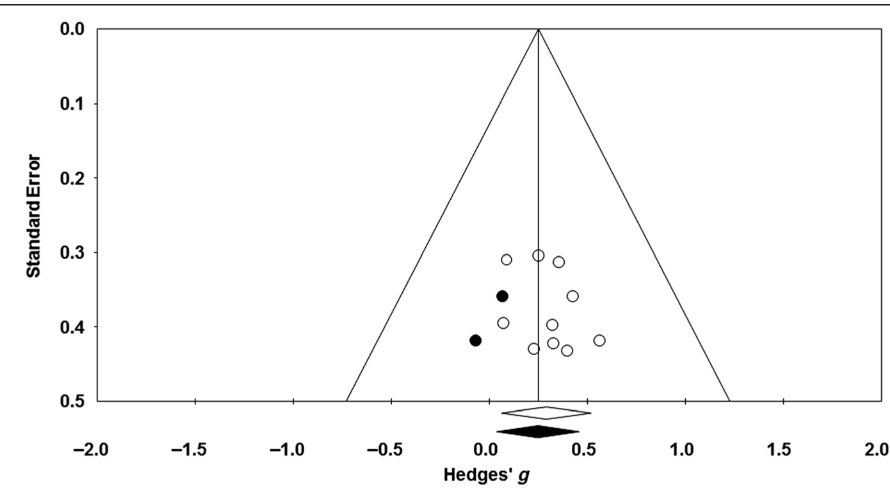


Figure 9. Funnel plot of publication bias: left DLPFC anodal stimulation and WM training. Dark data points represent studies imputed by the trim-and-fill procedure.

Table 5. Left DLPFC Stimulation on WM, Modeled on Brunoni and Vanderhasselt (2014): Study Characteristics and Effect Sizes

Study	*N (Anode/ Sham)*	*Mean Age (y)*	*Age Range (y)*	*% Male*	*Stimulation Site (Anode)*	*Reference Site (Cathode)*	*Reference Site (Cephalic)*	*Current (mA)*	*Current Density (mA/cm²)*	*Duration (min)*	*Test Timing*	*Test*	*Measure Coded*	*Design*	*Ceiling Effects?*	*Hedges' g*
Berryhill and Jones (2012)	25	63.7	56–80	Not reported	F3	Right cheek	No	1.5	0.04	10	Offline	2-back	Accuracy	Within-participant	No	0.031
Fregni et al. (2005)	15	20.2	19–22	26.7	F3	CSA	Yes	1	0.03	10	Online	3-back	Accuracy	Within-participant	No	0.335
Gill et al. (2015), Experiment 1	11	21.8	18–25	72.7	F3	CSA	Yes	2	0.08	20	Online	3-back	Accuracy	Within-participant	Yes	0.139
Hoy et al. (2013)	18	24.71	Not reported	38.9	F3	CSA	Yes	1	0.03	20	Offline	2-back	Accuracy	Within-participant	No	0.123
					0.06				3-back							
Keeser et al. (2011)	10	28.89	Not reported	50	F3	CSA	Yes	2	0.06	20	Offline	2-back	Accuracy	Within-participant	No	0.442
Keshvari et al. (2013)	30	22.3	Not reported	50	F3	F4	Yes	2	0.08	20	Offline	2-back	Accuracy	Within-participant	No	0.148
Lally et al. (2013)	21 (10/ 11)	23.1	Not reported	33.3	F3	Right cheek	No	1	0.03	10	Online	3-back	Accuracy	Between-participant	No	0.292
											Offline					
Martin et al. (2013)	42 (21/ 21)	23.1/ 23.2	Not reported	57.1/61.9	F3	Right deltoid	No	2	0.06	30	Online	Adaptive dual *n*-back	Accuracy	Between-participant	No	0.159
Meiron and Lavidor (2013)	25 (14/ 11)	24.9/ 23.1	18–36 (*N* = 41)	50/36.4	F3	Cz	Yes	2	0.13	15	Online	Modified *n*-back	Accuracy	Between-participant	Yes	0.319
											Offline	Poststimulation modified *n*-back				

Table 5. (*continued*)

Study	*N (Anode/Sham)*	*Mean Age (y)*	*Age Range (y)*	*% Male*	*Stimulation Site (Anode)*	*Reference Site (Cathode)*	*Reference Site (Cephalic)*	*Current (mA)*	*Current Density (mA/cm²)*	*Duration (min)*	*Test Timing*	*Test*	*Measure Coded*	*Design*	*Ceiling Effects?*	*Hedges' g*
Mulquiney et al. (2011)	10	29.4	Not reported	40	F3	CSA	Yes	1	0.03	10	Offline	2-back	RT	Within-participant	No	0.399
Mylius et al. (2012)	12	25.1	20–25 (N = 24)	50	F3	CSA	Yes	2	0.06	20	Offline	2-back	RT	Within-participant	Yes	0.110
Nozari and Thompson-Schill (2013)	24	Not reported	19–30	45.8	F3	F4	Yes	1.5	0.06	20	Online	2-back 3-back	Accuracy	Within-participant	No	0.248
Ohn et al. (2008)	15	26.5	Not reported	33.3	F3	CSA	Yes	1	0.04	30	Online Offline	3-back	Accuracy	Within-participant	No	0.302
Teo et al. (2011)	12	Not reported	Not reported	Not reported	F3	CSA	Yes	1	0.03 0.06	20	Online	3-back	Accuracy	Within-participant	No	−0.003

Study	Hedges' g	Lower limit	Upper Limit	SE
Berryhill and Jones (2012)	0.031	−0.516	0.578	0.279
Fregni et al. (2005)	0.335	−0.367	1.036	0.358
Gill et al. (2014), Experiment 1	0.139	−0.666	0.944	0.411
Hoy et al. (2013)	0.123	−0.518	0.764	0.327
Keeser et al. (2011)	0.442	−0.409	1.293	0.434
Keshvari et al. (2013)	0.148	−0.353	0.648	0.255
Lally et al. (2013)	0.292	−0.541	1.125	0.425
Martin et al. (2013)	0.159	−0.435	0.754	0.303
Meiron and Lavidor (2012)	0.319	−0.456	1.093	0.395
Mulquiney et al. (2011)	0.399	−0.449	1.248	0.433
Mylius et al. (2011)	0.110	−0.663	0.883	0.394
Nozari and Thompson-Schill (2013)	0.248	−0.312	0.807	0.286
Ohn et al. (2008)	0.302	−0.401	1.005	0.359
Teo et al. (2011)	−0.003	−0.775	0.770	0.394

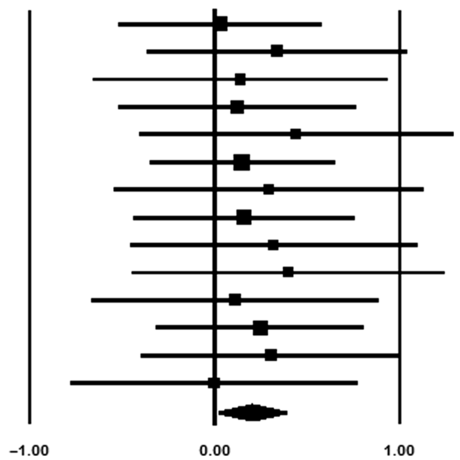


Figure 10. Forest plot: reanalysis modeled on Brunoni and Vanderhasselt (2014).

Neither right DLPFC nor right parietal anodal stimulation appeared to enhance WM, at least according to the relatively small set of studies analyzed here (eight and seven, respectively).

DISCUSSION

Despite the substantial literature on tDCS and WM, of which data from 31 published reports were meta-analyzed here, the true enhancement potential of tDCS for WM remains somewhat uncertain. At present, the best provisional conclusion is that anodal stimulation of the left DLPFC can boost the effectiveness of WM training and may possibly also be helpful when applied before or during tests of WM. However, the effects appear to be small.

The issue of whether, and by how much, tDCS enhances WM would seem to be a straightforward empirical question, which more than a few published studies have addressed. Why then is it so difficult to derive a clear answer from the literature? Several aspects of research practice in this field contribute to the persisting uncertainty.

First, the study designs used in this literature generally include small samples of participants. For example, in the largest set of studies, those assessing the effect of left DLPFC stimulation on WM, over a half of the studies included fewer than 16 participants receiving stimulation. Across a range of assumptions concerning variability in both within- versus between-participant designs, these small samples will render the experiments badly underpowered for detecting small effects such as those found here.

Figure 11. Funnel plot of publication bias: left DLPFC anodal stimulation reanalysis modeled on Brunoni and Vanderhasselt (2014). Dark data points represent studies imputed by the trim-and-fill procedure.

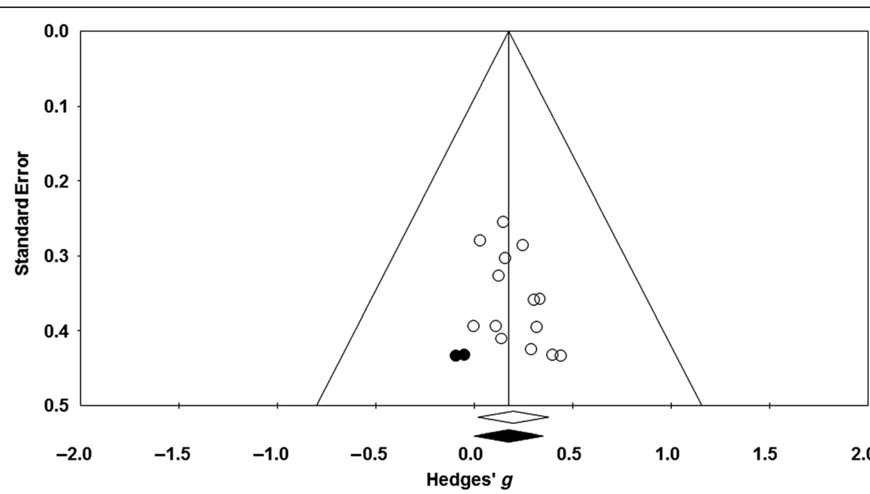


Table 6. Left DLPFC Stimulation on WM, Modeled on Horvath et al. (2015): Study Characteristics and Effect Sizes

Study	*N (Anode/Sham)*	*Mean Age (y)*	*Age Range (y)*	*% Male*	*Stimulation Site (Anode)*	*Reference Site (Cathode)*	*Reference Site (Cephalic)*	*Current (mA)*	*Current Density (mA/cm²)*	*Duration (min)*	*Test Timing*	*Test*	*Measure Coded*	*Design*	*Ceiling Effects?*	*Hedges' g*
Andrews et al. (2010)	11 (11/10)	28.1	20–51	40	F3	CSA	Yes	1	0.03	10	Offline	Digit span forward Digit span backward	Accuracy	Within-participant	No	0.327
Berryhill and Jones (2012)	25	63.7	56–80	Not reported	F3	Right cheek	No	1.5	0.04	10	Offline	2-back	Accuracy	Within-participant	No	0.031
Fregni et al. (2005)	15	20.2	19–22	26.7	F3	CSA	Yes	1	0.03	10	Online	3-back	Accuracy	Within-participant	No	0.335
Gill et al. (2015), Experiment 1	11	21.8	18–25	72.7	F3	CSA	Yes	2	0.08	20	Online	3-back	Accuracy	Within-participant	No	0.139
Gladwin et al. (2012)	14	22	Not reported	42.9	F3	CSA	Yes	1	0.03	10	Offline	Sternberg	Accuracy	Within-participant	Yes	0.305
Hoy et al. (2013)	18	24.7	Not reported	38.9	F3	CSA	Yes	1	0.03 0.06	20	Offline	2-back 3-back	Accuracy	Within-participant	No	0.123
Jeon and Han (2012)	16 (8/8)	39.5/36.5	20–59 (*N* = 32)	37.5/37.5	F3	CSA	Yes	1	0.03	20	Offline	Digit span forward Digit span backward	Accuracy	Between-participant	No	0.185
Keeser et al. (2011)	10	28.9	Not reported	50	F3	CSA	Yes	2	0.06	20	Offline	2-back	Accuracy	Within-participant	No	0.442
Keshvari et al. (2013)	30	22.3	Not reported	50	F3	F4	Yes	2	0.08	20	Offline	2-back	Accuracy	Within-participant	No	0.148

Lally et al. (2013)	21 (10/11)	23.1	Not reported	33.3	F3	Right cheek	No	1	0.03	10	Online	3-back	Accuracy	Between-participant	No	0.292
											Offline					
Mulquiney et al. (2011)	10	29.4	Not reported	40	F3	CSA	Yes	1	0.03	10	Online	Sternberg	Accuracy	Within-participant	No	0.165
											Offline	2-back	RT		Yes	
Mylius et al. (2012)	12	25.1	20–25 ($N = 24$)	50	F3	CSA	Yes	2	0.06	20	Offline	2-back	RT	Within-participant	Yes	0.110
Nozari and Thompson-Schill (2013)	24	Not reported	19–30	45.8	F3	F4	Yes	1.5	0.06	20	Online	2-back	Accuracy	Within-participant	No	0.248
												3-back				
Ohn et al. (2008)	15	26.5	Not reported	33.3	F3	CSA	Yes	1	0.04	30	Online	3-back	Accuracy	Within-participant	No	0.302
											Offline					
Tadini et al. (2011)	22	Not reported	18–64	63.6	F3	CSA	Yes	1.3	0.04	30	Offline	Digit span forward	Accuracy	Within-participant	No	−0.196
												Digit span backward				
Teo et al. (2011)	12	Not reported	Not reported	Not reported	F3	CSA	Yes	1	0.03	20	Online	3-back	Accuracy	Within-participant	No	0.036
									0.06		Offline	Sternberg				

Study	Hedges' g	Lower Limit	Upper Limit	SE
Andrews et al. (2011)	0.327	−0.502	1.155	0.423
Berryhill and Jones (2012)	0.031	−0.516	0.578	0.279
Fregni et al. (2005)	0.335	−0.367	1.036	0.358
Gill et al. (2014), Experiment 1	0.139	−0.666	0.944	0.411
Gladwin et al. (2012)	0.305	−0.420	1.029	0.369
Hoy et al. (2013)	0.123	−0.518	0.764	0.327
Jeon and Han (2012)	0.185	−0.744	1.114	0.474
Keeser et al. (2011)	0.442	−0.409	1.293	0.434
Keshvari et al. (2013)	0.148	−0.353	0.648	0.255
Lally et al. (2013)	0.292	−0.541	1.125	0.425
Mulquiney et al. (2011)	0.165	−0.678	1.008	0.430
Mylius et al. (2011)	0.110	−0.663	0.883	0.394
Nozari and Thompson-Schill (2013)	0.248	−0.312	0.807	0.286
Ohn et al. (2008)	0.302	−0.401	1.005	0.359
Tadini et al. (2011)	−0.196	−0.786	0.395	0.301
Teo et al. (2011)	0.036	−0.737	0.809	0.395

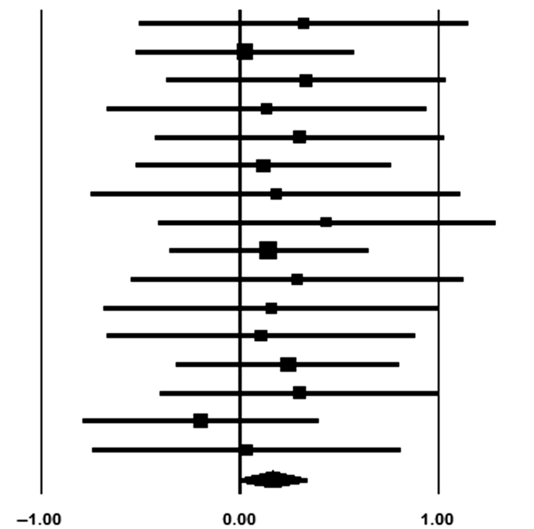


Figure 12. Forest plot: reanalysis modeled on Horvath et al. (2015).

Second, the tasks used to assess WM offer researchers a choice of specific measures to analyze. To take as an example the most frequently used WM task, *n*-back, one can focus on total accuracy, misses (a measure of accuracy on *n*-back matches), overall RT, or RT to *n*-back matches. If decisions about which measure to prioritize in analysis are made after the data have been examined, researchers can be lured by chance differences to focus on the measure with the biggest effect, believing that it shows the enhancement effect "most clearly."

Third, tDCS studies of WM in healthy participants are relatively easy and inexpensive to carry out, compared with fMRI studies or studies of unusual participants. This minimal investment would be expected to increase the willingness of researchers to simply scrap null or inconclusive results rather than to report them so that they can instead move on to a new study. All three of these characteristics increase the risk of spurious findings and publication bias and make it difficult to assess the true enhancing potential of tDCS.

Research laboratories are not the only contexts in which tDCS is used in an attempt to enhance cognition. A growing number of people are using tDCS to perform better at work or in online gaming, with online communities offering advice concerning the purchase, fabrication, and use of tDCS devices (Batuman, 2015). Subscribers to the largest Web site number in thousands (Jwa, 2015). How can this practice be reconciled with the weak effects found here? Several possibilities exist. Users may be experiencing a placebo effect. Alternatively, a subset of individuals may

Figure 13. Funnel plot of publication bias: left DLPFC anodal stimulation reanalysis modeled on Horvath et al. (2015). Dark data points represent studies imputed by the trim-and-fill procedure.

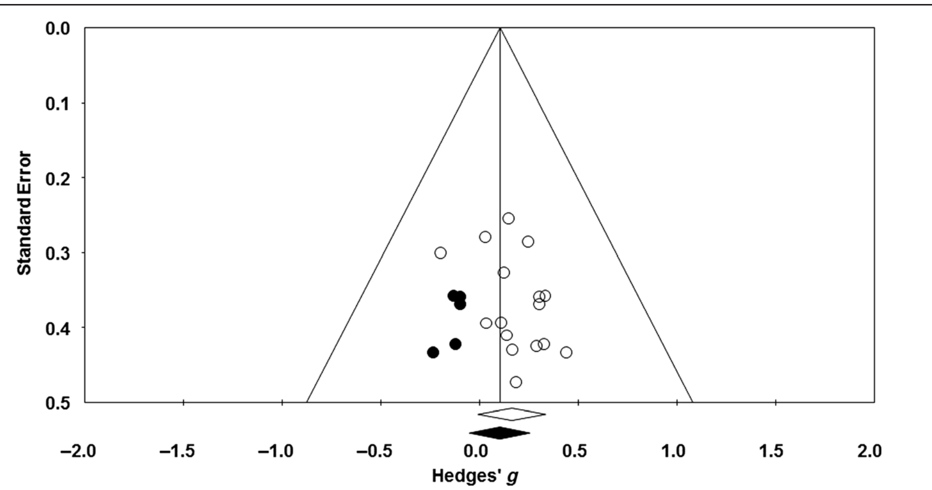


Table 7. Left DLPFC Stimulation on WM, Modeled on Hill et al. (2016): Study Characteristics and Effect Sizes

Study	*N (Anode/Sham)*	*Mean Age (y)*	*Age Range (y)*	*% Male*	*Stimulation Site (Anode)*	*Reference Site (Cathode)*	*Reference Site (Cephalic)*	*Current (mA)*	*Current Density (mA/cm^2)*	*Duration (min)*	*Test Timing*	*Test*	*Measure Coded*	*Design*	*Ceiling Effects?*	*Hedges' g*
Andrews et al. (2010)	11 (11/10)	28.1	20–51	40	F3	CSA	Yes	1	0.03	10	Offline	Digit span forward Digit span backward	Accuracy	Within-participant	No	0.327
Berryhill and Jones (2012)	25	63.7	56–80	Not reported	F3	Right cheek	No	1.5	0.04	10	Offline	2-back	Accuracy	Within-participant	No	0.031
Fregni et al. (2005)	15	20.2	19–22	26.7	F3	CSA	Yes	1	0.03	10	Online	3-back	Accuracy	Within-participant	No	0.335
Gill et al. (2015), Experiment 1	11	21.8	18–25	72.7	F3	CSA	Yes	2	0.08	20	Online	3-back	Accuracy	Within-participant	No	0.139
Gladwin et al. (2012)	14	22	Not reported	42.9	F3	CSA	Yes	1	0.03	10	Offline	Sternberg	Accuracy	Within-participant	Yes	0.305
Hoy et al. (2013)	18	24.7	Not reported	38.9	F3	CSA	Yes	1	0.03 0.06	20	Offline	2-back 3-back	Accuracy	Within-participant	No	0.123
Jeon and Han (2012)	16 (8/8)	39.5/36.5	20–59 (N = 32)	37.5/37.5	F3	CSA	Yes	1	0.03	20	Offline	Digit span forward Digit span backward	Accuracy	Between-participant	No	0.185
Keeser et al. (2011)	10	28.9	Not reported	50	F3	CSA	Yes	2	0.06	20	Offline	2-back	Accuracy	Within-participant	No	0.442
Keshvari et al. (2013)	30	22.3	Not reported	50	F3	F4	Yes	2	0.08	20	Offline	2-back	Accuracy	Within-participant	No	0.148
Lally et al. (2013)	21 (10/11)	23.1	Not reported	33.3	F3	Right cheek	No	1	0.03	10	Online Offline	3-back	Accuracy	Between-participant	No	0.292

Table 7. (*continued*)

Study	*N (Anode/Sham)*	*Mean Age (y)*	*Age Range (y)*	*% Male*	*Stimulation Site (Anode)*	*Reference Site (Cathode)*	*Reference Site (Cephalic)*	*Current (mA)*	*Current Density (mA/cm²)*	*Duration (min)*	*Test Timing*	*Test*	*Measure Coded*	*Design*	*Ceiling Effects?*	*Hedges' g*
Mulquiney et al. (2011)	10	29.4	Not reported	40	F3	CSA	Yes	1	0.03	10	Online	Sternberg	Accuracy	Within-participant	No	0.165
											Offline	2-back	RT		Yes	
Mylius et al. (2012)	12	25.1	20–25 (N = 24)	50	F3	CSA	Yes	2	0.06	20	Offline	2-back	RT	Within-participant	Yes	0.110
Nozari and Thompson-Schill (2013)	24	Not reported	19–30	45.8	F3	F4	Yes	1.5	0.06	20	Online	2-back	Accuracy	Within-participant	No	0.248
												3-back				
Ohn et al. (2008)	15	26.5	Not reported	33.3	F3	CSA	Yes	1	0.04	30	Online	3-back	Accuracy	Within-participant	No	0.302
											Offline					
Tadini et al. (2011)	22	Not reported	18–64	63.6	F3	CSA	Yes	1.3	0.04	30	Offline	Digit span forward	Accuracy	Within-participant	No	−0.196
												Digit span backward				
Teo et al. (2011)	12	Not reported	Not reported	Not reported	F3	CSA	Yes	1	0.03	20	Online	3-back	Accuracy	Within-participant	No	0.036
									0.06		Offline	Sternberg				

Study	Hedges' g	Lower Limit	Upper Limit	SE
Andrews et al. (2011)	0.327	−0.502	1.155	0.423
Berryhill and Jones (2012)	0.031	−0.516	0.578	0.279
Fregni et al. (2005)	0.335	−0.367	1.036	0.358
Gill et al. (2014), Experiment 1	0.139	−0.666	0.944	0.411
Gladwin et al. (2012)	0.305	−0.420	1.029	0.369
Hoy et al. (2013)	0.123	−0.518	0.764	0.327
Jeon and Han (2012)	0.185	−0.744	1.114	0.474
Keeser et al. (2011)	0.442	−0.409	1.293	0.434
Keshvari et al. (2013)	0.148	−0.353	0.648	0.255
Lally et al. (2013)	0.292	−0.541	1.125	0.425
Mulquiney et al. (2011)	0.165	−0.678	1.008	0.430
Mylius et al. (2011)	0.110	−0.663	0.883	0.394
Nozari and Thompson-Schill (2013)	0.248	−0.312	0.807	0.286
Ohn et al. (2008)	0.302	−0.401	1.005	0.359
Tadini et al. (2011)	−0.196	−0.786	0.395	0.301
Teo et al. (2011)	0.036	−0.737	0.809	0.395

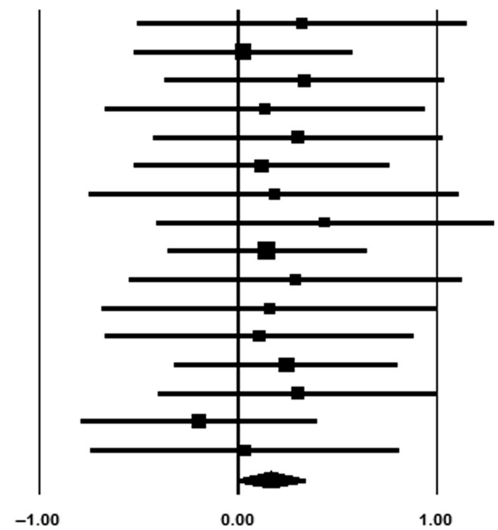
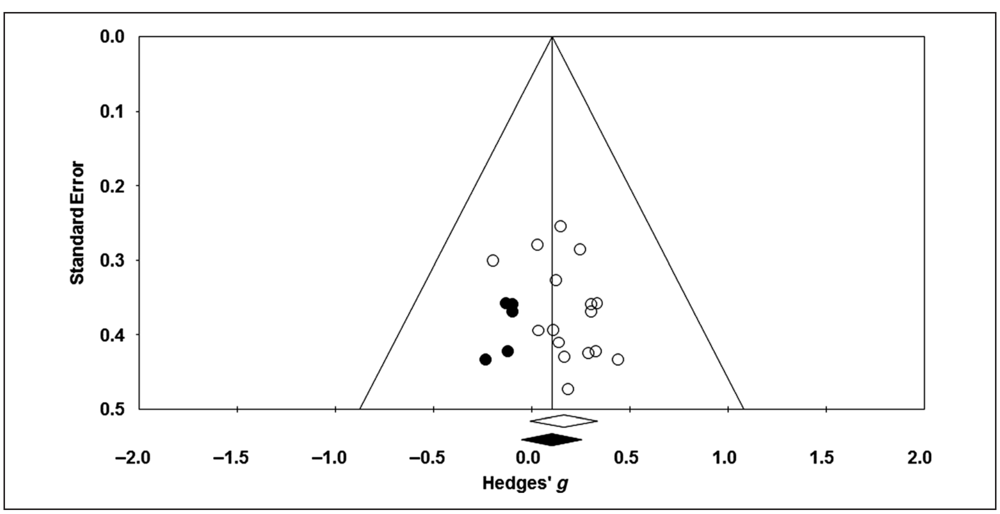


Figure 14. Forest plot: reanalysis modeled on Hill et al. (2016).

experience real enhancement effects (Berryhill & Jones, 2012) and they may be overrepresented among users, even though their response is not typical. It is also possible that small improvements to WM can lead to larger improvements in other tasks or that small improvements are themselves worthwhile. Finally, it is of course possible that tDCS enhances cognitive systems other than WM.

New stimulation protocols using transcranial current are now being explored, raising new possibilities for cognitive enhancement as well as new questions regarding efficacy. High-definition tDCS, in which multiple small electrodes are used to create more focal current flow, may more effectively modulate cognitive abilities including WM (e.g., Nikolin, Loo, Bai, Dokos, & Martin, 2015). By varying current over time, with alternating current stimulation, random noise stimulation, and pulsed current stimulation, potentially different psychological effects may be obtained. The effectiveness of these new stimulation protocols for cognitive enhancement remains to be determined as the literature grows. It stands to reason that tDCS and newer forms of transcranial current stimulation could modulate cognitive performance and learning, given their effects on neuronal excitability. Discovering which of these methods can enhance cognition requires adequately powered studies, a priori selection of outcome measures, and reporting of null results.

Figure 15. Funnel plot of publication bias: left DLPFC anodal stimulation reanalysis modeled on Hill et al. (2016). Dark data points represent studies imputed by the trim-and-fill procedure.



Reprint requests should be sent to Lauren Elizabeth Mancuso, Center for Neuroscience and Society, University of Pennsylvania, 3720 Walnut St., Philadelphia, PA 19104, or via e-mail: itslaurenmancuso@gmail.com.

REFERENCES

Andrews, S. C., Hoy, K. E., Enticott, P. G., Daskalakis, Z. J., & Fitzgerald, P. B. (2010). Improving working memory: The effect of combining cognitive activity and anodal transcranial direct current stimulation to the left dorsolateral prefrontal cortex. *Brain Stimulation, 4,* 84–89.

Batsikadze, G., Moliadze, V., Paulus, W., Kuo, M. F., & Nitsche, M. A. (2013). Partially non linear stimulation intensity dependent effects of direct current stimulation on motor cortex excitability in humans. *Journal of Physiology, 591,* 1987–2000.

Batuman, E. (2015, April). Electrified: Adventures in transcranial direct-current stimulation. *The New Yorker*. Retrieved from www.newyorker.com/magazine/2015/04/06/electrified.

Berryhill, M. E., & Jones, K. T. (2012). tDCS selectively improves working memory in older adults with more education. *Neuroscience Letters, 521,* 148–151.

Berryhill, M. E., Peterson, D. J., Jones, K. T., & Stephens, J. A. (2014). Hits and misses: Leveraging tDCS to advance cognitive research. *Frontiers in Psychology, 5,* 800.

Berryhill, M. E., Wencil, E. B., Coslett, B. H., & Olson, I. R. (2010). A selective working memory impairment after transcranial direct current stimulation to the right parietal lobe. *Neuroscience Letters, 479,* 312–316.

Bona, S., & Silvanto, J. (2014). Accuracy and confidence of visual short-term memory do not go hand-in-hand: Behavioral and neural dissociations. *PLoS One, 9,* e90808.

Braver, T. S., Cohen, J. D., Nystrom, L. E., Jonides, J., Smith, E. E., & Noll, D. C. (1997). A parametric study of prefrontal cortex involvement in human working memory. *Neuroimage, 5,* 49–62.

Brunoni, A. R., & Vanderhasselt, M. A. (2014). Working memory improvement with non-invasive brain stimulation of the dorsolateral prefrontal cortex: A systematic review and meta-analysis. *Brain and Cognition, 86,* 1–9.

Carlson, S., Martinkauppi, S., Rämä, P., Salli, E., Korvenoja, A., & Aronen, H. J. (1998). Distribution of cortical activation during visuospatial *n*-back tasks as revealed by functional magnetic resonance imaging. *Cerebral Cortex, 8,* 743–752.

Carpenter, P. A., Just, M. A., & Reichle, E. D. (2000). Working memory and executive function: Evidence from neuroimaging. *Current Opinion in Neurobiology, 10,* 195–199.

Chhatbar, P. Y., & Feng, W. (2015). Data synthesis in meta-analysis may conclude differently on cognitive effect from transcranial direct current stimulation. *Brain Stimulation, 8,* 974–976.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cooper, H. (2010). *Research synthesis and meta-analysis*. Thousand Oaks, CA: Sage Publications.

Dunlap, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods, 1,* 170–177.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General, 128,* 309–331.

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

Gill, J., Shah-Basak, P. P., & Hamilton, R. H. (2015). It's the thought that counts: Examining the task-dependent effects of transcranial direct current stimulation on executive function. *Brain Stimulation, 8,* 253–259.

Gladwin, T. E., den Uyl, T. E., Fregni, F. F., & Wiers, R. W. (2012). Enhancement of selective attention by tDCS: Interaction with interference in a Sternberg task. *Neuroscience Letters, 512,* 33–37.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational and Behavioral Statistics, 6,* 107–128.

Hill, A. T., Fitzgerald, P. B., & Hoy, K. E. (2016). Effects of anodal transcranial direct current stimulation on working memory: A systematic review and meta-analysis of findings from healthy and neuropsychiatric populations. *Brain Stimulation, 9,* 197–208.

Horvath, J. C. (2015). New quantitative analyses following Price & Hamilton's critique do not change original findings of Horvath et al. *Brain Stimulation, 8,* 665.

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,* 535–550.

Hoy, K. E., Emonson, M. R. L., Arnold, S. L., Thomson, R. H., Daskalakis, Z. J., & Fitzgerald, P. B. (2013). Testing the limits: Investigating the effect of tDCS dose on working memory enhancement in healthy controls. *Neuropsychologia, 51,* 1777–1784.

Hsu, T. Y., Tseng, P., Liang, W. K., Cheng, S. K., & Juan, C. H. (2014). Transcranial direct current stimulation over right posterior parietal cortex changes prestimulus alpha oscillation in visual short-term memory task. *Neuroimage, 98,* 306–313.

Ilieva, I. P., Hook, C. J., & Farah, M. J. (2015). Prescription stimulants' effects on healthy inhibitory control, working memory, and episodic memory: A meta-analysis. *Journal of Cognitive Neuroscience, 27,* 1069–1089.

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine, 2,* 696–701.

Jacobson, L., Koslowsky, M., & Lavidor, M. (2012). tDCS polarity effects in motor and cognitive domains: A meta-analytical review. *Experimental Brain Research, 216,* 1–10.

Jeon, S. Y., & Han, S. J. (2012). Improvement of the working memory and naming by transcranial direct current stimulation. *Annals of Rehabilitation Medicine, 36,* 585–595.

Jones, K. T., & Berryhill, M. E. (2012). Parietal contributions to visual working memory depend on task difficulty. *Frontiers in Psychiatry, 3,* 81.

Jones, K. T., Gözenman, F., & Berryhill, M. E. (2015). The strategy and motivational influences on the beneficial effect of neurostimulation: A tDCS and fNIRS study. *Neuroimage, 105,* 238–247.

Jwa, A. (2015). Early adopters of the magical thinking cap: A study on do-it-yourself (DIY) transcranial direct current stimulation (tDCS) user community. *Journal of Law and the Biosciences, 2,* 292–335.

Keeser, D., Padberg, F., Reisinger, E., Pogarell, O., Kirsch, V., Palm, U., et al. (2011). Prefrontal direct current stimulation modulates resting EEG and event-related potentials in healthy subjects: A standardized low resolution tomography (sLORETA) study. *Neuroimage, 55,* 644–657.

Keshvari, F., Pouretemad, H. R., & Ekhtiari, H. (2013). The polarity-dependent effects of the bilateral brain stimulation on working memory. *Basic and Clinical Neuroscience, 4,* 224–231.

Lally, N., Nord, C. L., Walsh, V., & Roiser, J. P. (2013). Does excitatory fronto-extracerebral tDCS lead to improved working memory performance? *F1000Research, 2,* 219.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage Publications.

Martin, D. M., Liu, R., Alonzo, A., Green, M., Player, M. J., Sachdev, P., et al. (2013). Can transcranial direct current stimulation enhance outcomes from cognitive training? A randomized controlled trial in healthy participants. *International Journal of Neuropsychopharmacology, 16,* 1927–1936.

Meiron, O., & Lavidor, M. (2013). Unilateral prefrontal direct current stimulation effects are modulated by working memory load and gender. *Brain Stimulation, 6,* 440–447.

Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology, 49,* 270–291.

Mulquiney, P. G., Hoy, K. E., Daskalakis, Z. J., & Fitzgerald, P. B. (2011). Improving working memory: Exploring the effect of transcranial random noise stimulation and transcranial direct current stimulation on the dorsolateral prefrontal cortex. *Clinical Neurophysiology, 122,* 2384–2389.

Mylius, V., Jung, M., Menzler, K., Haag, A., Khader, P. H., Oertel, W. H., et al. (2012). Effects of transcranial direct current stimulation on pain perception and working memory. *European Journal of Pain, 16,* 974–982.

Nikolin, S., Loo, C. K., Bai, S., Dokos, S., & Martin, D. M. (2015). Focalised stimulation using high definition transcranial direct current stimulation (HD-tDCS) to investigate declarative verbal learning and memory functioning. *Neuroimage, 117,* 11–19.

Nitsche, M. A., Bikson, M., & Bestmann, S. (2015). On the use of meta-analysis in neuromodulatory non-invasive brain stimulation. *Brain Stimulation, 8,* 666–667.

Nitsche, M. A., Cohen, L. G., Wassermann, E. M., Priori, A., Lang, N., Antal, A., et al. (2008). Transcranial direct current stimulation: State of the art 2008. *Brain Stimulation, 1,* 206–223.

Nitsche, M. A., & Paulus, W. (2000). Excitability changes induced in the human motor cortex by weak transcranial direct current stimulation. *Journal of Physiology, 527,* 633–639.

Nozari, N., & Thompson-Schill, S. L. (2013). More attention when speaking: Does it help or does it hurt? *Neuropsychologia, 51,* 2770–2780.

Ohn, S. H., Park, C. I., Yoo, W. K., Ko, M. H., Choi, K. P., Kim, G. M., et al. (2008). Time-dependent effect of transcranial direct current stimulation on the enhancement of working memory. *NeuroReport, 19,* 43–47.

Park, S. H., Seo, J. H., Kim, Y. H., & Ko, M. H. (2014). Long-term effects of transcranial direct current stimulation combined with computer-assisted cognitive training in healthy older adults. *NeuroReport, 25,* 122–126.

Price, A. R., & Hamilton, R. H. (2015). A re-evaluation of the cognitive effects from single-session transcranial direct current stimulation. *Brain Stimulation, 8,* 663–665.

Price, A. R., McAdams, H., Grossman, M., & Hamilton, R. H. (2015). A meta-analysis of transcranial direct current stimulation studies examining the reliability of effects on language measures. *Brain Stimulation, 8,* 1093–1100.

Priori, A. (2003). Brain polarization in humans: A reappraisal of an old tool for prolonged non-invasive modulation of brain excitability. *Clinical Neurophysiology, 114,* 589–595.

Ragland, J. D., Turetsky, B. I., Gur, R. C., Gunning-Dixon, F., Turner, T., Schroeder, L., et al. (2002). Working memory for complex figures: An fMRI comparison of letter and fractal *n*-back tasks. *Neuropsychology, 16,* 370–379.

Richmond, L. L., Wolk, D., Chein, J., & Olson, I. R. (2014). Transcranial direct current stimulation enhances verbal working memory training performance over time and near transfer outcomes. *Journal of Cognitive Neuroscience, 26,* 2443–2454.

Rosenthal, R., & Rubin, D. B. (1986). Meta-analytic procedures for combining studies with multiple effect sizes. *Psychological Bulletin, 99,* 400–406.

Rothstein, H. R., Sutton, A. J., & Borenstein, M. (2006). *Publication bias in meta-analysis: Prevention, assessment and adjustments*. Chichester, UK: Wiley.

Sandrini, M., Fertonani, A., Cohen, L. G., & Miniussi, C. (2011). Double dissociation of working memory load effects induced by bilateral parietal modulation. *Neuropsychologia, 50,* 396–402.

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.

Smith, E. E. (2001). Working memory. In R. Wilson & F. Keil (Eds.), *The MIT encyclopedia of the cognitive sciences*. Cambridge, MA: MIT Press.

Stagg, C. J. (2014). The physiological basis of brain stimulation. In R. Cohen Kadosh (Ed.), *The stimulated brain: Cognitive enhancement using non-invasive brain stimulation* (pp. 145–177). London: Academic Press.

Sternberg, S. (1998). Inferring mental operations from reaction time data. In D. Scarborough & S. Sternberg (Eds.), *Invitation to cognitive science* (2nd ed.). Cambridge, MA: MIT Press.

Tadini, L., El-Nazer, R., Brunoni, A. R., Williams, J., Carvas, M., Boggio, P., et al. (2011). Cognitive, mood, and electroencephalographic effects of noninvasive cortical stimulation with weak electrical currents. *Journal of ECT, 27,* 134–140.

Teo, F., Hoy, K. E., Daskalakis, Z. J., & Fitzgerald, P. B. (2011). Investigating the role of current strength in tDCS modulation of working memory performance in healthy controls. *Frontiers in Psychiatry, 2,* 45.

Tseng, P., Hsu, T. Y., Chang, C. F., Tzeng, O. J. L., Hung, D. L., Muggleton, N. G., et al. (2012). Unleashing potential: Transcranial direct current stimulation over the right posterior parietal cortex improves change detection in low-performing individuals. *Journal of Neuroscience, 32,* 10554–10561.

Vanderhasselt, M. A., Brunoni, A. R., Loeys, T., Boggio, P. S., & De Raedt, R. (2013). Nosce te ipsum–Socrates revisited? Controlling momentary ruminative self-referent thoughts by neuromodulation of emotional working memory. *Neuropsychologia, 51,* 2581–2589.

Wiley, J., & Jarosz, A. F. (2012). Working memory capacity, attentional focus, and problem solving. *Current Directions in Psychological Science, 21,* 258–262.

Wolkenstein, L., & Plewnia, C. (2012). Amelioration of cognitive control in depression by transcranial direct current stimulation. *Biological Psychiatry, 73,* 646–651.

Wu, Y. J., Tseng, P., Chang, C. F., Pai, M. C., Hsu, K. S., Lin, C. C., et al. (2014). Modulating the interference effect on spatial working memory by applying transcranial direct current stimulation over the right dorsolateral prefrontal cortex. *Brain and Cognition, 91,* 87–94.