Brain Stimulation 8 (2015) 1093–1100

Contents lists available at ScienceDirect

Brain Stimulation

journal homepage: www.brainstimjrnl.com

A Meta-analysis of Transcranial Direct Current Stimulation Studies Examining the Reliability of Effects on Language Measures

Amy R. Price [a], Harrison McAdams [b], Murray Grossman [a], Roy H. Hamilton [a,b,*]

[a] Center for Cognitive Neuroscience, Department of Neurology, University of Pennsylvania, Philadelphia, PA 19104, USA
[b] Perelman School of Medicine, University of Pennsylvania, Philadelphia PA 19104, USA

ARTICLE INFO

Article history:
Received 16 April 2015
Received in revised form 19 June 2015
Accepted 22 June 2015
Available online 23 July 2015

Keywords:
Transcranial direct current stimulation
tDCS
Language
Meta-analysis
Reliability
Verbal fluency

ABSTRACT

Background: Transcranial direct current stimulation (tDCS) is a brain stimulation technique used to examine causal relationships between brain regions and cognitive functions. The effects from tDCS are complex, and the extent to which stimulation reliably affects different cognitive domains is not fully understood and continues to be debated.

Objective/hypothesis: To conduct a meta-analysis of studies examining the effects of single-session anodal tDCS on language.

Methods: The meta-analysis examined the behavioral results from eleven experiments of single-session anodal tDCS and language processing in healthy adults. The means and standard deviations of the outcome measures were extracted from each experiment and entered into the meta-analyses. In the first analysis, we examined the effects of single-session tDCS across all language studies. Next, a series of sub-analyses examined the effects of tDCS on specific tasks and stimulation protocols.

Results: There was a significant effect from anodal single-session tDCS in healthy adults compared to sham ($P = 0.001$) across all language measures. Next, we found significant effects on specific stimulation protocols (e.g., offline measures, $P = 0.002$), as well as specific tasks and electrode montages (e.g., verbal fluency measures and left prefrontal cortex, $P = 0.035$).

Conclusions: The results indicate that single-session tDCS produces significant and reliable effects on language measures in healthy adults.

Introduction

Transcranial direct current stimulation (tDCS) is a relatively new technique in neuroscience research and its use has rapidly grown over the last decade [1–4]. In healthy adults, tDCS offers a tool for testing causal relationships between brain regions and their underlying cognitive functions (e.g., motor control, working memory, language) [5–8]. This technique has also been tested in clinical populations for a wide variety of uses, ranging from psychiatric conditions to neurological disorders caused by stroke and neurodegenerative diseases [9–15]. Based on research in animals and humans, it is generally thought that applying anodal tDCS to a brain area leads to increased neural excitability in that region, while cathodal stimulation leads to decreased neural excitability [1,2,16–21]. However the effects from tDCS are complex and appear to be affected by a number of stimulation parameters, including intensity and duration [1,2,22,23], the types of cognitive processes engaged during stimulation [24,25], stimulation polarity [1,19–22,26,27], the underlying levels of cortical neurochemicals [28,29], and genetics [30]. Research examining the extent and reliability of tDCS effects is ongoing, and it remains an important challenge to determine the parameters under which tDCS can affect cognition and neurophysiology.

The use of tDCS is widespread in basic research and in clinical settings. Hundreds of researchers use the technique, and there have been over a thousand publications involving tDCS in the last decade. However, a recent meta-analysis examined the reliability of single-session tDCS on cognition and reported null findings for the effects of tDCS across a variety of cognitive domains [31]. Given the widespread use of direct current stimulation and the broad claims that have been made based on these recent null findings, it is important that the results from this meta-analysis are able to be

This work was supported by funding from National Institutes of Health (AG017586 and 5T32AG000255).

* Corresponding author. Department of Neurology, University of Pennsylvania, 3400 Spruce St., 3 West Gates, Philadelphia, PA 19104, USA.
E-mail address: roy.hamilton@uphs.upenn.edu (R.H. Hamilton).

http://dx.doi.org/10.1016/j.brs.2015.06.013
1935-861X/

validated. If single-session tDCS has no effect on cognition, this finding would have far-reaching implications for the field and for the future of this technique.

A detailed examination of the meta-analysis by Horvath and colleagues has raised a number of problematic issues in the methodological approach, as reviewed by a number of researchers [32–34]. It therefore remains unclear whether there is truly no effect of single-session tDCS on cognition. The goal of the analyses presented in this paper was to examine this question in the cognitive domain of language. To do this, we analyzed all of the behavioral data from the tDCS papers included in the previous review of language studies [31]. Our meta-analysis was structured into two levels. The first level consisted of a main meta-analysis that examined data from across all language studies using the same behavioral measure (e.g., accuracy measures), but that differed in task (e.g., verbal fluency or novel language learning) and the time-point of data collection (e.g., online or offline). This allowed us to generalize across studies in order to achieve maximum power to detect whether effects on language processes were present. The second level of our investigation consisted of sub-analyses that entailed more narrowly focused examinations of the effects from tDCS for individual tasks (e.g., verbal fluency), electrode montages (e.g., left PFC), and stimulation conditions (e.g., offline). This allowed us to examine more specific effects relating to particular tasks and stimulation conditions. The findings from this meta-analysis reveal significant effects across many behavioral measures in the language studies and provide important implications for future research using tDCS.

Methods

Individual study selection and analysis

Our goal was to analyze behavioral data from language studies that had applied tDCS in healthy adults. Previously, Horvath and colleagues [31] had analyzed data from these studies and found no effect from stimulation. However, a number of investigators revealed substantial weaknesses in these analyses, stemming from the use of inconsistent data-selection criteria and the lack of methodological details explaining these decisions [32–34]. To address these issues, we provide a detailed outline of how data were selected for our study and for the corresponding analyses performed by Horvath and colleagues (Table 1).

Because tDCS is a relatively recent technique, it is difficult to find a large number of studies using the same electrode montages, tasks, and stimuli. However small sample sizes in meta-analyses severely limit the power and reliability of summary statistics. To address this, we structured the meta-analysis to first perform a large-sample main analysis across the data from all of the language papers, followed by smaller sub-analyses that aimed to address more specific questions about particular tasks and stimulation protocols.

We examined eleven language experiments from nine manuscripts that reported the effects of anodal stimulation relative to sham stimulation [35–43] (two manuscripts reported two separate experiments on independent groups of subjects). The manuscripts were read in detail by two independent reviewers, and the data for the mean and variance for each behavioral measure were recorded from each experiment. In some cases, the mean and/or variance were not reported in the manuscript but included as images in the figures of the manuscripts. In these cases, the images were exported and the size of each measure and/or error bar was assessed using a metric overlay in Adobe Illustrator. In the analysis performed by Horvath and colleagues [31], there were cases in which multiple measures were selected from a single experiment but counted as independent data in the meta-analysis. This produces dependent data samples (i.e., multiple measures from a single group of subjects), which violates the assumption that the data represent independent experiments. For experiments that reported multiple measures of the same type of task (e.g., two post-stimulation measures of verbal fluency, obtained close in time after the offset of stimulation), we used an averaged score of these measures in the meta-analysis, rather than counting each measure as an independent data set. For studies that reported the effects of multiple montages, the electrode montage that was the most consistent with the other studies in the analysis was selected. The data were labeled according to whether the effect represented a measure taken during stimulation (i.e., online) or after the end of stimulation (i.e., offline).

Using the data from each paper, we calculated standardized mean difference (SMD) effect sizes as follows [44,45]:

$$SMD = (X_1 - X_2)\Big/\left(SD_{pooled}\right)$$

Where X_1 is the mean of the behavioral measurement from the anode condition, $X_2$is the mean of the behavioral measurement from the sham condition, and SD_{pooled} is the pooled standard deviation from both the anode and sham conditions

A common rule of thumb for interpreting SMD effect sizes is that a value of ~0.2 indicates a small effect, a value of ~0.5 indicates a medium effect and a value of ~0.8 or higher indicates a large effect [46].

Many published tDCS studies include small sample sizes (as low as 10 subjects in the studies examined here). Small sample sizes can produce a positive bias in meta-analyses. We therefore performed a correction to adjust for small sample sizes [45]. This provides a more conservative estimate of the effect size for each experiment (i.e., results in a slight reduction of SMD). The correction was performed using the formula:

$$\mathrm{SMD_{adjusted}} = d \times \mathrm{J}(\mathrm{n}_1 + \mathrm{n}_2 - 2)$$

$$\mathrm{J}(\mathrm{n}) = \Gamma(\mathrm{n}/2)\Big/\left(\left(\sqrt{n/2}\right)\Gamma((\mathrm{n}-1)/2)\right)$$

Where Γ is the gamma function, n_1 is sample size for anodal group, n_2 is sample size for sham group, and $SMD_{adjusted}$ = the standardized mean difference corrected for small sample size.

For one experiment [35], the t-statistic was provided without the individual condition means and standard deviations. Using the t-statistic, the standardized mean difference could be calculated with the following relationship:

$$\mathrm{SMD} = \mathrm{t} \times \sqrt{((\mathrm{n}_1 + \mathrm{n}_2)/(\mathrm{n}_1 \times \mathrm{n}_2))}$$

Where n_1 is the sample size for anodal group, n_2 is sample size for sham group, and t is t-statistic.

Meta-analysis

There were two levels to this meta-analysis. The first level of analyses examined all language studies with comparable behavioral measures (e.g., accuracy), in order to identify whether single-session tDCS has a general effect on language-related behaviors. This approach benefited from having a larger number of studies [47,48], but provided less specificity with respect to the type of behavior examined. The second level of analyses aimed to identify more specific effects from tDCS. This approach had lower power but had the benefit that it included specificity with regard to the type of language task performed and the location of electrode placement.