

The Influence of Skin Redness on Blinding in Transcranial Direct Current Stimulation Studies: A Crossover Trial

Fernando Ezquerro, MD*¹; Adriano H. Moffa, PE, PsyD, PhD Candidate*^{†1}; Marom Bikson, MD, BS, PhD[‡]; Niranjana Khadka, AAS, BE, PhD Candidate[‡]; Luana V. M. Aparicio, MD*[†]; Bernardo de Sampaio-Junior, MD, PhD Candidate*[†]; Felipe Fregni, MD, PhD, MMSc, MPH[§]; Isabela M. Bensenor, MD, PhD*[†]; Paulo A. Lotufo, MD, PhD*[†]; Alexandre Costa Pereira, MD, PhD^{¶1}; Andre R. Brunoni, MD, PhD*[†]

Objective: To evaluate whether and to which extent skin redness (erythema) affects investigator blinding in transcranial direct current stimulation (tDCS) trials.

Material and Methods: Twenty-six volunteers received sham and active tDCS, which was applied with saline-soaked sponges of different thicknesses. High-resolution skin images, taken before and 5, 15, and 30 min after stimulation, were randomized and presented to experienced raters who evaluated erythema intensity and judged on the likelihood of stimulation condition (sham vs. active). In addition, semi-automated image processing generated probability heatmaps and surface area coverage of erythema. Adverse events were also collected.

Results: Erythema was present, but less intense in sham compared to active groups. Erythema intensity was inversely and directly associated to correct sham and active stimulation group allocation, respectively. Our image analyses found that erythema also occurs after sham and its distribution is homogenous below electrodes. Tingling frequency was higher using thin compared to thick sponges, whereas erythema was more intense under thick sponges.

Conclusions: Optimal investigator blinding is achieved when erythema after tDCS is mild. Erythema distribution under the electrode is patchy, occurs after sham tDCS and varies according to sponge thickness. We discuss methods to address skin erythema-related tDCS unblinding.

Keywords: Adverse effects, blinding, computer modeling, erythema, randomized-controlled trials, skin redness, transcranial direct current stimulation

Conflict of Interest: The authors have no conflicts of interest to disclose.

Address correspondence to: Andre Russowsky Brunoni, MD, PhD, Interdisciplinary Center for Applied Neuromodulation, Av. Professor Lineu Prestes, 2565, 3o andar, CEP 05508-000, São Paulo, SP, Brazil. Email: brunoni@usp.br

* Interdisciplinary Center for Applied Neuromodulation (CINA), University Hospital, University of São Paulo, São Paulo, Brazil;

† Service of Interdisciplinary Neuromodulation (SIN), Laboratory of Neurosciences (LIM-27), Department and Institute of Psychiatry, University of São Paulo, São Paulo, Brazil;

‡ Department of Biomedical Engineering, The City College of City University of New York, New York, NY, USA;

§ Spaulding Neuromodulation Center, Spaulding Rehabilitation Hospital, Harvard Medical School, Boston, MA, USA; and

¶ Heart Institute (INCOR), University of São Paulo, São Paulo, Brazil

For more information on author guidelines, an explanation of our peer review process, and conflict of interest informed consent policies, please go to <http://www.wiley.com/WileyCDA/Section/id-301854.html>

¹Both authors equally contributed to the paper.

Sources of funding: This work was supported by a grant from the National Council for Scientific and Technological Development (CNPq, Grant Number 470904). Andre R. Brunoni is also supported by the following grants: 2013 NARSAD Young Investigator from the Brain & Behavior Research Foundation (Grant Number 20493) and 2013 FAPESP Young Researcher from the São Paulo State Foundation (Grant Number 20911-5). Andre R. Brunoni is recipient of a research fellowship award from CNPq (303197). Adriano H Moffa has a CNPq Master's grant.

INTRODUCTION

Transcranial direct current stimulation (tDCS) is a non-invasive brain stimulation technique that involves delivery of a weak, direct current to the brain through electrodes placed over the subject's scalp (1). TDCS has been increasingly investigated as a possible treatment for diverse neuropsychiatric disorders (2). Notwithstanding the presence of well-designed and conducted trials (2), the overall number of studies is still small, especially studies with moderate to large sample sizes. Thus, improved techniques and protocols are warranted to enhance internal validity and research quality (3,4).

Optimal blinding techniques remain a concern in tDCS trials. In a canonical study, Gandiga *et al.* (5) used a sham method that consisted of a brief period of 1mA stimulation followed by no stimulation until the end of the session, concluding that sham tDCS could be successfully used in double-blind trials, as subjects were not able to distinguish between real and sham stimulation. However, recent evidence suggested that Gandiga *et al.*'s method is inadequate in some contexts, such as rater's blinding, 2 mA current intensity, cross-over designs and non-naïve tDCS subjects (6,7).

Skin redness (erythema) after tDCS is one reason for inadequate investigator blinding. Palm and collaborators (8) found that operators, even when blinded using tDCS devices with a number code that automatically delivers active or sham stimulation, were able to differentiate between active and sham stimulation based on skin reddening after active tDCS.

Causes for tDCS erythema may include irritation by the saline, iontophoresis of substances present in skin prior to stimulation (make up, sunscreen, cleansing substances, etc.), pressure by headgear, and the stimulation itself; whereas electrode design and thickness, gender, skin type, nature of stimulation (anodal or cathodal), and amperage of stimulation may mediate its intensity (9–11).

Although subjects' blinding can be managed by avoiding self-inspection of the forehead immediately after stimulation, this can be particularly troublesome for raters who are assessing outcomes immediately after the end of stimulation. For instance, O'Connell *et al.* (6) reported that erythema was noted after 60% of active stimulation sessions, compared to 1% after sham; moreover, 98% of the investigators associated noticeable skin redness with active stimulation. The authors also noticed that some skin redness persisted for several minutes beyond the end of stimulation.

Recent studies have been conducted to characterize and control tDCS-induced erythema. We previously reported that skin pretreatment with ketoprofen reduces tDCS-induced erythema (11), although such approach inconveniently increases the preparation time. In addition, electrode-sponge geometry (rectangular vs. round-shaped) was explored as a method for improving bias (12); however, no difference was observed on the potential for blinding. Larger electrode size was also found to be associated with cutaneous discomfort (13).

However, no study hitherto has objectively evaluated the influence of tDCS-induced erythema on investigator blinding, or whether it is dependent on specific brands of available electrodes. Therefore, this issue was investigated in the present study. To this end, we used high-definition skin photographs of tDCS-induced erythema, presented at random to investigators. We also used semi-automated image processing to determine redness and simulated a probability skin heatmap, and surface area coverage of redness using image processing software. Finally, we examined adverse effects and subjects' blinding.

Table 1. Clinical and Demographic Characteristics of the Sample.

Sample	Data
Subjects, <i>n</i>	26
Age, mean (SD)	26.2 (4.7)
Gender, <i>n</i> (%)	
Male	5 (19.2%)
Female	21 (80.8%)
Ethnicity (Self-Reported)	
Caucasian	21 (80.8%)
Yellow	0 (0%)
Mixed	4 (15.4%)
Black	1 (3.8%)
Years of schooling, mean (SD)	16.6 (2.4)
Skin type	
Dry, <i>n</i> (%)	4 (15.4%)
Mixed, <i>n</i> (%)	8 (30.8%)
Normal, <i>n</i> (%)	3 (11.5%)
Oily, <i>n</i> (%)	11 (42.3%)
Smoking habits	
No smoker, <i>n</i> (%)	22 (84.6%)
Smoker, <i>n</i> (%)	4 (5.4%)
Women on estrogen contraceptives, <i>n</i> (%)	16 (76.2%)
Fitzpatrick scale	
I, <i>n</i> (%)	4 (15.4%)
II, <i>n</i> (%)	4 (15.4%)
III, <i>n</i> (%)	8 (30.8%)
IV, <i>n</i> (%)	5 (19.2%)
V, <i>n</i> (%)	5 (19.2%)

Type I: white, very fair, red, or blond hair, blue eyes, freckles, always burns, never tans; type II: white, fair skin, red or blond hair, blue, hazel or green eyes, usually burns, tans with difficulty; type III: cream white, fair with any eye or hair color, sometimes mild burn, gradually tans; type IV: brown, typical Mediterranean Caucasian skin, rarely burns, tans with ease; type V: dark brown, Middle Eastern skin type, very rarely burns, tans easily; type VI: black, never burns, tans very easily.

MATERIALS AND METHODS

Subjects

We recruited 26 healthy volunteers (21 women) aged between 18 and 45 years ($M = 26.2$; $SD = 4.7$) who were naïve to tDCS applications and presented no active dermatosis, skin allergy, skin marks, recent exposure to intense sunlight or artificial tanning, systemic skin treatment or topical skin treatment in the region where the electrodes were placed. Clinical and demographic data were collected from each participant (Table 1). The local and national ethics committees approved the study and all participants provided written informed consent.

Materials

We used two different sets of sponges in this study: "thick" sponges (5.6 mm thickness), manufactured by Soterix Medical (Soterix Medical Inc., NY, USA) and "thin" sponges (1.5 mm thickness), manufactured by Neuroconn (Neuroconn GmbH, Munich, Germany). These sponges are commonly used in tDCS trials, both being cellulose-based, 5×5 cm wide, and behave similarly to absorb saline.

Design

Our independent variable was stimulation condition (3 levels: thick-active, thin-active, and sham). Dependent variables included the Draize erythema scale, stimulation group allocation, Likert scale,

Table 2. Investigators Evaluations.

Stimulation condition	Descriptive Stat	Time			
		Baseline	After 5 min	After 15 min	After 30 min
Sham	Mean	0.1944444	0.6533334	0.4492754	0.3484848
	Standard Deviation	0.2585105	0.5567764	0.4450479	0.3329724
	Median	0	0.6666667	0.3333333	0.3333333
	Interquartile Range	0.3333333	0.6666666	1	0.6666667
Thin	Mean	0.1923077	1.397436	1.179487	0.9166667
	Standard Deviation	0.3151448	0.6464895	0.5268841	0.3838528
	Median	0	1.333333	1.333333	1
	Interquartile Range	0.3333333	1	0.3333334	0.3333333
Thick	Mean	0.1333333	1.041667	0.8055556	0.6521739
	Standard Deviation	0.1666667	0.6240935	0.5553139	0.4765009
	Median	0	1	0.6666667	0.6666667
	Interquartile Range	0.3333333	0.8333334	0.8333334	0.6666666

adverse effects scales, comfort visual analog scale (VAS), and the region-of-interest probability heatmap for erythema.

Following a within-subject design, participants received three stimulation sessions separated by one-week intervals: active tDCS with thick sponge, active tDCS with thin sponge and sham stimulation (in this case, in half of sham sessions we used thin sponges and, in the other half, thick sponges). The sequence was “active thick” on week 1, then “sham” on week 2, and finally “active thin” on week 3.

Procedures

The study was conducted in rooms with controlled temperature and humidity (temperature $20^{\circ}\text{C} \pm 2^{\circ}\text{C}$, relative humidity $50\% \pm 5$ relative humidity) as to ensure standardization of the dermatologic evaluation. As participants presented to the study, they were interviewed for exclusion criteria, demographics, medical history, use of contraceptives and other medications, sun exposure and other life habits since childhood.

High-resolution skin images were acquired before tDCS sessions, then again 5, 15, and 30 min after. Each day, stimulations were performed after gentle cleansing of forehead with ethanol. In case the participant wore make-up, a thorough face wash was requested with a 30 min interval for the beginning of procedures (including the baseline photograph). Immediately following each session, participants answered the adverse effects questionnaire, the comfort VAS, as well as whether they judged they had received true or sham stimulation.

tDCS Protocol

The anode was placed in the right supraorbital (SO) region. The electrode location was standardized with positioning according to the following parameters: the uppermost limit of the right eyebrow, in the line of the pupil, was the downmost inferior limit to place the inferior border of the electrode; transversally, the medial border of the electrode corresponded to the medial limit of the eyebrow. The anode was held by plastic straps in such a way to assure an even amount of pressure across the whole area of the electrode.

The cathode was placed over the vertex area and its position was fixed by plastic straps.

The rationale for this was that we aimed for optimal image acquisition and the most even electrode placement over de skin with littlest inter-subject anatomical bone surface variation; the right side was an arbitrary standard. Only anodal stimulation was tested in this study, based on previous finding of our group showing that anodic

tDCS generates more intense erythema (11) and to keep consistent across conditions.

Stimulations were carried out using 1×1 tDCS devices (Soterix Medical Inc., New York, NY, USA). The active stimulation consisted of 30 min of a 2 mA current intensity plateau, with ramp-up and ramp-down periods of 45 sec and 15 sec, respectively. Sham stimulation consisted of a 30-min interval with no current, with a brief period (60 sec) of 2 mA stimulation at the beginning of the session, with ramp-up and ramp-down periods of 45 sec and 15 sec, respectively. The sponges were soaked with saline solution (NaCl 0.9%).

Image Data Protocol

Investigator-Based Image Rating

A VISIA® Imaging System (Canfield Scientific, Fairfield, NJ, USA) was used to photograph the forehead of the subjects using a combination of regular white light, 365 nm wavelength UV light and cross-polarized lightning flashes. A total of 292 photos from 26 participants were obtained (4 per session per subject except for 20 (<6.5%) images that were missed due to technical reasons). The photographed images are of very high definition (21 Megapixel resolution, 3433×4171 pixels size, image DPI 96 pixels/inch, file size 8.4 MB); 145mm (width) per 175 mm (height). They were then presented in a random order to three investigators that were naive to study design and aim. Investigators were asked to sit alone in an office, observe the photograph and then give a score of skin redness and to make judgment whether the subject had received active or sham stimulation based solely on that image. The investigators were physicians participating in other tDCS trials for at least two years. Their inter-rater evaluation agreement was excellent ($\kappa = 0.82$, $p < 0.01$).

Software-Based Image Rating

The images were also analyzed for erythema distribution using customized MATLAB (MathWorks, MA, USA) based image processing script that included graphical user interface (GUI). Images corresponding to each group were first randomized and were loaded in the GUI, which was designed to define a 5×5 cm region of interest (ROI) (corresponds to the length of the SO (Fig # A1a)). Erythema beyond the ROI was not included in the analysis. Images were then filtered using Lab color space; the most accurate means of representing color, is device independent, and includes all colors in the visible spectrum, as well as colors outside human perception. Using a freehand tool enabled in the GUI, erythema inside the ROI was

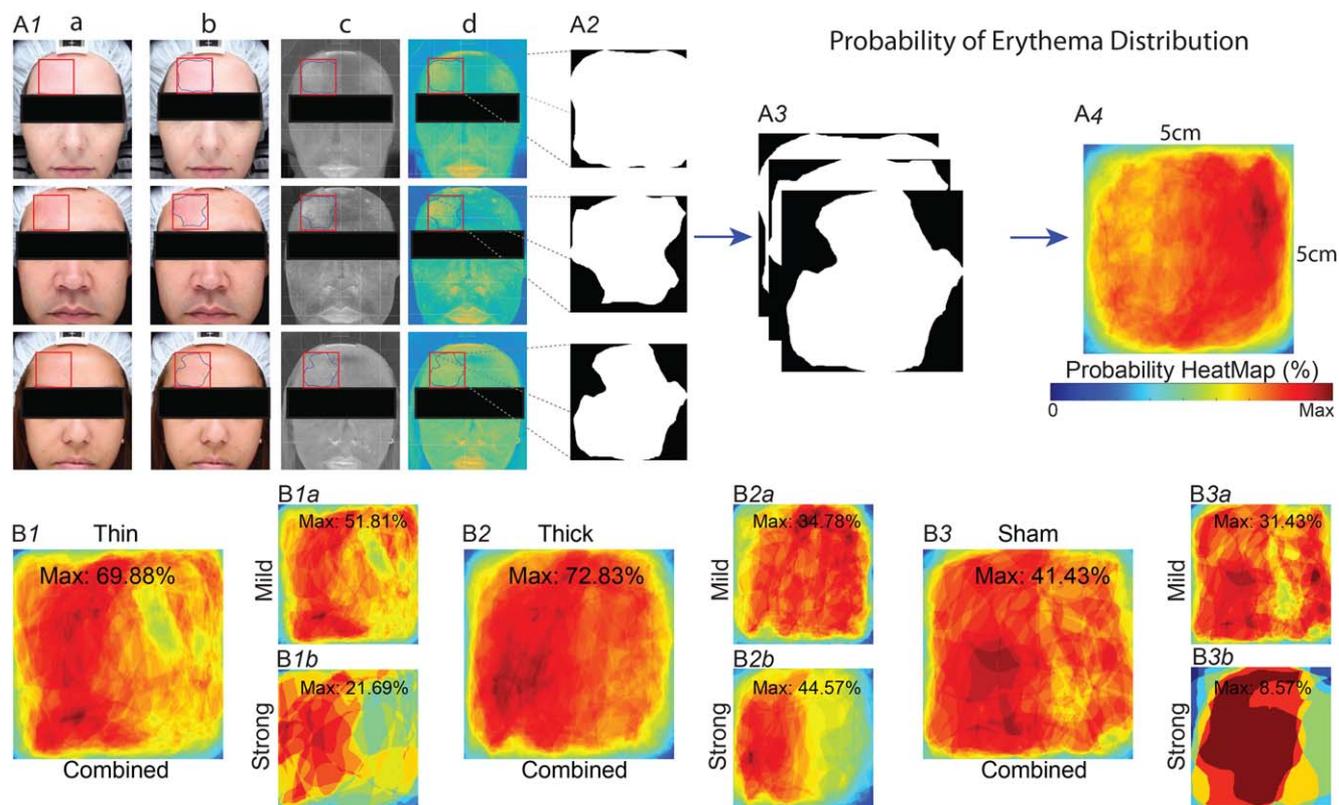


Figure 1. Erythema distribution analysis in the ROI (site of stimulation) for active (using thin and thick sponges) and sham stimulation. (A1a) Illustration of high definition images of subject photographed before and after the stimulation. (A1b) represents ROI and traced erythema distribution. (A1c and A1d) Representation of filtered Images using Lab color space to isolate erythema from regular skin color tone. (A2) illustrates binary image of erythema traces in the ROI. (A3) Illustration of the probability of erythema distribution calculation via stacking equidimensional binary images. (A4) Probability heatmap of erythema distribution (in percentage) across the ROI. Max represents 100% probability in the color bar. (B1) Graphical illustration of the calculated maximum combined probability of erythema distribution for the “thin” sponge. Distribution was diffused across the ROI. Mild erythema distribution (B1a) had higher probability than strong (B1b). (B2) represents a maximum combined probability of erythema distribution for “thick” sponge. Strong erythema distribution (B2b) was slightly higher than mild (B2a). (B3) Illustration of the maximum combined probability of erythema distribution for “sham” stimulation. Probability of the erythema distribution compared to the active stimulation sponge types was significantly lower. Mild erythema (B3a) had higher probability than strong erythema (B3b).

traced (Fig. 1A1b–d). The same rater (NK) traced erythema in all images for both active and sham stimulation cases. Erythema was separately traced as “mild” and “strong” inside the ROI. Assignment of the traces as “mild” and “strong” was based on comparison of natural tone of the facial skin (non-stimulated area) to the skin tone under the ROI. Traces of erythema were then binarized and

normalized within the ROI (Fig. 1A2). Binary images were re-categorized to their respective sections: thin, thick, and sham, as mild and strong erythema. Surface area of erythema trace inside the binary ROI was estimated first by finding the perimeter of the erythema distribution. Pixels that were part of the perimeter were only nonzero (1s) and were at least connected to one zero-valued pixel (0s). The default connectivity was 4 for a given 2D binary ROI. All the white pixels representing the erythema traces inside the ROI were enumerated and summed up to obtain the surface area in pixels. Finally, using a calibration factor, the area in pixels was then converted to area in cm². The percentage erythema was calculated by dividing the ROI by erythema area, although the ROI is the same in each case. The mean of the combined, mild, and strong erythema distribution for active and sham stimulations was calculated (Fig. 1A3) and a probability heatmap of the distribution was generated (Fig. 1A4).

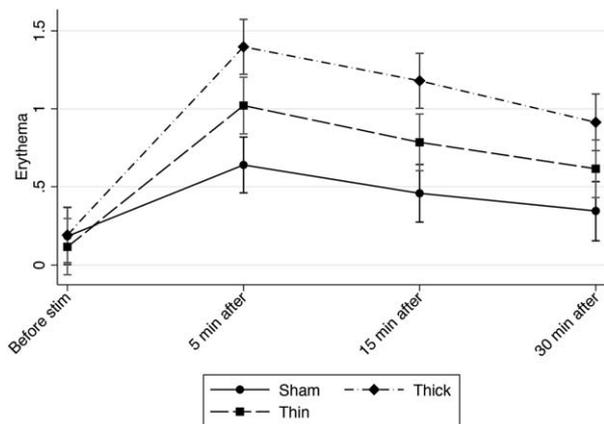


Figure 2. Erythema intensity and correct stimulation group allocation between groups.

Assessments

Immediately after stimulation, adverse effects were assessed using a standard questionnaire (14) and subjects were asked whether they received real or sham stimulation. At the end of all sessions participants were asked again to identify which session was the sham one. Comfort experience after stimulation was measured using a visual analogue scale (VAS) ruler by 10 centimeters (cm), which was

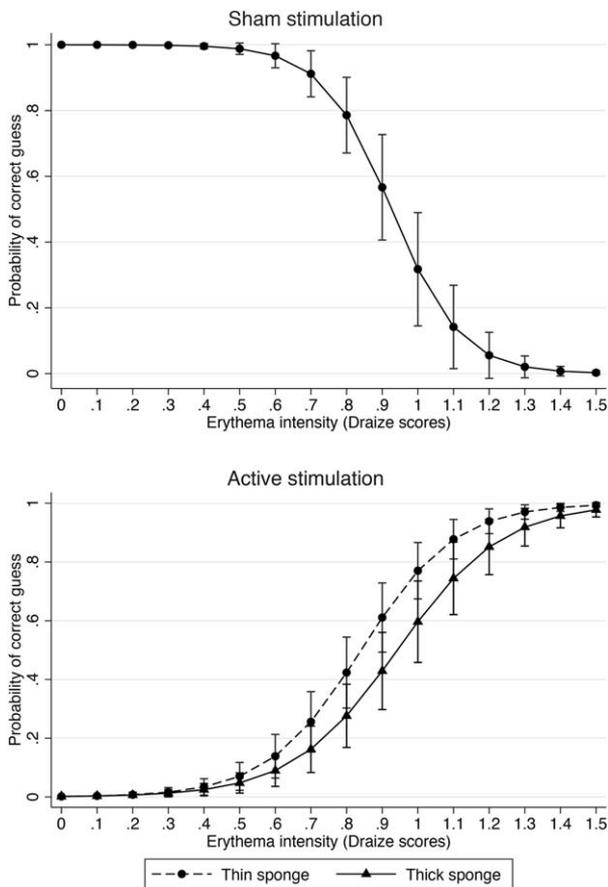


Figure 3. Probability of correctly judging stimulation condition according to erythema intensity. The figure shows the predicted probability of correctly judging the stimulation condition in (A) sham stimulation and (B) active stimulations according to erythema intensity (Draize scores). As depicted, erythema scores between 0.8 and 1.1 (mild intensity) seem optimal for a satisfactory blinding (i.e., around 50%).

marked from “not comfortable at all” to “very comfortable,” which yielded a value between 0 and 100.

Investigators were asked to score the erythema of skin photographs using the Draize scoring system scale, grading as “0” if no erythema, “1” if slight erythema, “2” if well defined erythema, “3” if moderate to severe erythema, or “4” if severe erythema. They were also asked to judge for the stimulation condition through a Likert scale: “–2” if high confidence of sham stimulation; “–1” if low confidence of sham stimulation; “0” if unsure/cannot judge; “1” if low confidence of active stimulation or “2” if high confidence of active stimulation.

The “erythema score” of a given picture was the mean of the Draize scores according to the investigators’ evaluation (e.g., if, for a given picture, two evaluators scored “1” and one evaluator scored “0,” then this picture had an erythema score of 0.67; Table 2).

Statistical Analysis

Analyses were performed using Stata 12 (Statacorp, College Station, TX, USA) and considered significant at a two-sided $p \leq 0.05$. The mean scores of investigators’ evaluation were used in the analyses.

Power calculation was performed in Stata (fpower package), using 5% and 20% as the probabilities of error type I and II, respectively, and 0.9 as the estimated effect size. This yielded a sample size of 25

participants. The estimation of effect size was based on our previous study (11) that observed mean (SD) values for anodal stimulation of 1.25 (0.9). However, that study did not evaluate erythema after sham stimulation and to the best of our knowledge, no previous studies used the Draize score system to evaluate erythema as well. Therefore, based on our previous clinical observations of a discrete visible erythema after sham tDCS, we presumed a score of 0.45 (0.9) for the sham group.

Although the variable “erythema score” was moderately skewed (skewness = 0.82) and the test for normality (Shapiro-Wilk) was marginally significant ($p = 0.04$), we applied parametric tests for this variable according to the Central Limit Theorem that authorizes this approach when the number of observations is high, particularly for more than 30 observations (15)—in the present study, we have 292 observations. To further validate our findings, we also performed a non-parametric analysis for our main outcome.

The variable “judgment” (of stimulation condition) was categorized in two conditions: sham (values of –2 to –0.5) and active (0.5–2) group judgment. Values between –0.5 and 0.5 were disregarded as indicative of very low confidence on judging the stimulation condition. The variable “correct stimulation group allocation” was formulated according to “judgment” and “group” – “yes” if the estimation matched the stimulation condition and “no” if otherwise.

To compare erythema intensity between groups at different time points, we used a repeated-measures ANOVA with two independent variables: sponge (three levels: sham, thin, and thick) and time (four levels: t_1 , t_2 , t_3 , t_4). The Kruskal-Wallis one-way analysis of variance is a non-parametric test that was also used to compare erythema intensity between groups – when significant, pairwise comparisons were performed using the non-parametric Dunn’s test. To explore whether different erythema intensities impacted on judgment and for the correct stimulation group allocation analysis, we used logit regression models. To compare the frequency of adverse effects and the frequency of correct stimulation group allocation between groups, we used χ^2 -tests. To compare comfort perception between groups, we used a one-way ANOVA.

Finally, we explored the influence of some variables on erythema and on correct stimulation group allocation. These variables were: age, sex and Fitzpatrick scale. Here it is important to notice that Brazil is a country of highly mixed ethnical heritage, with Amerindian, African and European roots. In our demographic data, following country’s census standard practice, ethnicity is self-reported and may not reflect reliably ancestry. That is why we chose Fitzpatrick scale, not “race,” for analysis. The Fitzpatrick scale is a simple method to clinically evaluate (16) human skin pigmentation based on skin, hair, and eye color and skin response to solar exposition. In our sample, roughly 1/3 of the subjects presented types I and II, 1/3 presented type III and 1/3 presented type IV or V (Table 1), therefore, we reclassified these categories in three levels.

RESULTS

For the results section, “sham group” indicates sham tDCS, “thin sponge” indicates active tDCS with thin sponge and “thick sponge” indicates active tDCS with thick sponge.

Main Findings

Erythema was significantly lower in the sham compared to the active groups and in the thin sponge compared to the thick sponge at all timepoints, except at baseline ($p < 0.01$ for all comparisons, Bonferroni correction [< 0.004 adjusted]) (Fig. 2).

The Kruskal-Wallis test also showed that groups were significantly different ($\chi^2 = 49.3$, $p < 0.01$) and the Dunn's tests showed that erythema was lower in the sham vs. thick ($z = 7.12$, $p < 0.01$), vs. thin ($z = 3.65$, $p < 0.01$) and in the thin vs. thick ($z = 3.42$, $p < 0.01$) groups. In other words, non-parametric and parametric tests presented similar results.

In our exploratory analyses, we found a non-significant trend for main effects for Fitzpatrick ($F_{2,209} = 2.71$, $p = 0.07$). We also found a significant interaction between sex and sponge ($F_{2,209} = 13.74$, $p < 0.01$), but no main effects of sex. Further analyses revealed that, in sham stimulation, there was no difference in erythema intensity according to gender; however, for active stimulations (i.e., thin and thick sponges) erythema was more intense in men ($ps < 0.01$). Age was not associated to erythema ($p = 0.34$).

Regarding stimulation condition judgment, raters were moderately confident to judge sham stimulation in the sham group ($M = -0.77$; $SD = 0.9$) and moderately confident to judge active stimulation in the thick sponge group ($M = 0.65$; $SD = 1.67$). Raters were uncertain when judging the stimulation condition for the thin sponge group ($M = 0.01$; $SD = 1.18$). These findings were statistically significant ($F_{2,214} = 32.6$, $p < 0.01$). Bonferroni analyses showed that all groups were statistically different between them ($p < 0.01$ [< 0.016 adjusted] for all comparisons). We found no influence of gender, age or Fitzpatrick scale on correct stimulation group allocation.

Regarding correct stimulation group allocation, the raters' percentages of correctly judging the stimulation condition were 85% ($SD = 35$), 75% ($SD = 43$), and 40% ($SD = 49$), respectively for sham, thick sponge, and thin sponge groups. These values were significantly different between thin sponge stimulation and the other groups ($ps < 0.01$), but not between sham and thick groups ($p = 0.16$). Considering a 50% probability of correctly judging the stimulation condition, raters judged beyond chance for the sham and thick groups ($ps < 0.01$) but not for the thin sponge group ($p = 0.25$).

We further explored the frequency of correct stimulation group allocation according to erythema intensity using logit regressions. The probability of correctly judging in the sham group was inversely associated to erythema intensity ($\beta = -10.33$, $p < 0.01$) whereas the probability of correctly judging in the active groups was directly associated to erythema intensity ($\beta = 7.3$, $p < 0.01$) (Fig. 3).

Probability of Erythema Distribution

The calculated probability heatmap across active stimulation (using thin and thick sponges) and sham condition indicated that erythema was diffused across the ROI. For "thin" sponge, the maximum combined probability of erythema distribution was 69.88% (51.81% for mild and 21.69% for strong) (Fig. 1B1). "Thick" sponge had the maximum probability of 72.83% erythema distribution (34.78% for mild and 44.57% for strong) (Fig. 1B2). In case of "sham" stimulation, the max combined probability of erythema distribution was 41.43% (31.43% for mild and 8.57% for strong) (Fig. 1B3). Mean probability of erythema distribution yielded by thick sponge was slightly higher than that of the thin sponge, whereas for the sham stimulation, it was significantly lower (Fig. 1). Accordingly, a one-way ANOVA of erythema surface area coverage on thin, thick and sham groups was performed. There was a significant difference in the erythema surface area between groups ($F_{2,241} = 13.38$, $p < 0.01$). Further analyses indicated that both thin and thick have significantly larger erythema surface area compared to the sham ($p < 0.01$ for both

comparisons). However, the active groups (thin and thick sponges) did not differ significantly ($p = 0.4$).

We also investigated the influence of sex and Fitzpatrick scale on the mean of the erythema surface area coverage. No main effect for the Fitzpatrick scale was observed ($F = 0.92$, $p = 0.40$). This analysis also showed a main effect of sex ($F = 9.04$, $p < 0.01$) and no interaction between sex and sponge.

Other Findings

There was no difference in subject's perceived comfort during stimulation across groups ($F_{2,78} = 0.53$, $p = 0.59$).

The frequency of tingling was significantly higher after thin sponge stimulation compared to the other groups (88% vs. 64% and 40% for thin, thick sponges and sham, respectively, $\chi^2 = 12.5$, $p < 0.01$). The frequency of other adverse effects was similar in all groups (Supporting Information Figure).

Finally, statistical tests were performed to explore possible correlations between erythema scores, comfort data and adverse effects. However, no significant associations were observed (data not shown).

DISCUSSION

In this within-subjects study enrolling 26 healthy volunteers, we found that tDCS-induced erythema is generally mild to moderate. Interestingly, both looking at rater-based and software-based data, a very mild erythema occurred after sham stimulation although it was significantly higher after active stimulation, and even higher for the thick compared to thin sponge.

Our image processing analyses confirm that sham is significantly different from active tDCS stimulation but could not confirm the differences between thick and thin observed with rater and user data. Moreover, investigators, solely based on erythema evaluation, were moderately confident in judging stimulation condition for sham and thick sponge groups, respectively; although they were unsure to identify the stimulation condition for the thin sponge. Accordingly, correct stimulation group allocation occurred beyond chance for sham and thick sponge groups, but not for the thin sponge group. We also demonstrated that erythema is associated with correct stimulation group allocation in a "S-curve" pattern.

Our study also presented two new approaches to examine erythema and investigator blinding. Previous studies asked investigators to examine subject's head to judge the stimulation allocation group and score erythema intensity (6). Even if investigators are blinded to the stimulation allocation group, they will be aware of the timing of stimulation (before vs. after stimulation). Moreover, non-verbal cues can break blinding during erythema examination (e.g., scratching during examination) and subjects can inadvertently report other adverse events to the investigators. By asking investigators to evaluate erythema through high-definition skin photographs presented at random, we were able to eliminate this source of bias. Conversely, in our study, erythema evaluation might be overestimated compared to the usual trial where investigators only do a quick visual inspection of the skin.

Another novel approach was to use the collected images for estimating a probability heatmap on the skin area, which presumably represents the erythema distribution under the electrode. This model corroborated the investigators' observation of skin redness after sham stimulation. This might have occurred for some reasons such as 1) the brief period of active stimulation at the session onset; 2) pressure of the pad, depending on how it is fixed; and 3) irritation

of the skin due to the saline solution. In the active groups, the model showed that erythema was comparable between groups. Moreover, redness does not concentrate around pad edges but it is rather diffuse under the electrode. Assuming that the electric current causes redness, it seems that current density is fairly homogeneous below the pad, and redness would be caused by an increase in blood perfusion among the tissue. This is in contrast to a previous modeling study that showed that a thin sponge would have the current concentrated in the center of the sponge and a thick sponge, on the edges (17,18). However, that model did not fully capture the inhomogeneity and anisotropy within the skin; for instance, skin/scalp was considered a combined mass of muscle, skin, fat and connective tissues. A more recent model also estimated higher current densities at the edges when conductivity is high (19). Nonetheless, our finding reinforces the need to validate modeling studies empirically.

In specific cases, we found that visual inspection and the scoring assisted by redness segmentation produced different results, presumably reflecting the method of analysis. For instance, visual inspection showed no main effects of gender, but a significant interaction between gender and sponge that revealed that for sham stimulation there was no difference in erythema intensity according to gender, although skin reddening was higher during active stimulations in men. In fact, it has been previously found that men present more cutaneous vasodilation during capsaicin compared to women (20), a similar mechanism might be involved to explain greater tDCS skin redness in males. However, this finding was not corroborated by redness segmentation, which showed a main gender effect, but no interaction effects between gender and sponge. However, it should be noticed that the sample of males was very small compared to females – therefore, our evaluation of gender differences on skin redness is limited from a statistical perspective. Although generally the severity of adverse events was low across all condition, as expected (1), the frequency of tingling was significantly higher under thin vs. thick sponge stimulation (88% vs. 64% incidence, respectively). Tingling is the most common adverse effect reported in tDCS studies, being observed in almost 3 out of 4 subjects (21,22). This finding is of interest for further trials exploring methodologies aiming to reduce the frequency of this common adverse effect. This dissociation between erythema and tingling is compelling, and may potentially be explained by the thicker sponge producing more uniform current density at the skin surface, resulting in more distributed erythema and reduced sensation.

The implications of our erythema results in informing tDCS trial design should be taken with caution. First, our results are specific to the headgear (e.g., presuming sham erythema reflects pressure), electrode technologies, electrolyte (gel/saline/cream) used, subject demographics, and waveforms tested. We in fact show a dependence on electrode design and skin type. Trial-specific considerations would determine the need and value to mitigate erythema-related sham concerns. At a minimum, researchers should be rigorous in controlling and reporting relevant headgear and electrode, as well as other factors that could induce erythema. Simple methods to conceal exposed skin areas can be implemented. If appropriate, erythema intensity can be reduced by topic ketoprofen 2% before stimulation (11). Triple-blind studies where the raters do not apply tDCS reduce confound of operator un-blinding. Importantly, our protocol involved either trained operators or quantified segmentation, with optimal lighting and image capture, and with the targeted intention to identify erythema difference across arms, something impractical for regular use in tDCS trials.

Our findings, therefore, do not necessarily contradict conventional experience in tDCS trials where sham was found effective by

operator and subject reports, but rather raise the alert for more detailed report of procedures used in future research to conceal stimulation group allocation, since it is now well documented that erythema is an independent factor for breaking investigator blinding in within-subjects design.

Authorship Statements

Drs. Fernando Ezquerro, Adriano Moffa, and Andre Brunoni designed and conducted the study, including patient recruitment, data collection and data analysis. Dr. Fernando Ezquerro prepared the manuscript draft with important intellectual input from Andre R. Brunoni, Luana VM Aparicio, and Bernardo de Sampaio-Junior. All authors approved the final manuscript. Drs. Fernando Ezquerro, Adriano H. Moffa, Marom Bikson, Niranjana Khadka, Luana VM Aparicio, Bernardo de Sampaio-Junior, Felipe Fregno, Isabela M. Bensenor, Paulo A. Lotufo, Alexandre Costa Pereira, and Andre R. Brunoni had complete access to the study data. We would like to thank Drs. Felipe Fregni, Isabela M. Bensenor, and Paulo Lotufo for their editorial support during the preparation of this manuscript.

How to Cite this Article:

Ezquerro F., Moffa A.H., Bikson M., Khadka N., Aparicio L.V.M., de Sampaio-Junior B., Fregni F., Bensenor I.M., Lotufo P.A., Pereira A.C., Brunoni A.R. 2016. The Influence of Skin Redness on Blinding in Transcranial Direct Current Stimulation Studies: A Crossover Trial. *Neuromodulation* 2016; E-pub ahead of print. DOI: 10.1111/ner.12527

REFERENCES

- Brunoni AR, Nitsche MA, Bolognini N et al. Clinical research with transcranial direct current stimulation (tDCS): challenges and future directions. *Brain Stimul* 2012;5: 175–195.
- Tortella G, Casati R, Aparicio LV et al. Transcranial direct current stimulation in psychiatric disorders. *World J Psychiatry* 2015;5:88–102.
- Khadka N, Rahman A, Sarantos C, Truong DQ, Bikson M. Methods for specific electrode resistance measurement during transcranial direct current stimulation. *Brain Stimul* 2015;8:150–159.
- Woods AJ, Antal A, Bikson M et al. A technical guide to tDCS, and related non-invasive brain stimulation tools. *Clin Neurophysiol* 2016;127:1031–1048.
- Gandiga PC, Hummel FC, Cohen LG. Transcranial DC stimulation (tDCS): a tool for double-blind sham-controlled clinical studies in brain stimulation. *Clin Neurophysiol* 2006;117:845–850.
- O'Connell NE, Cossar J, Marston L et al. Rethinking clinical trials of transcranial direct current stimulation: participant and assessor blinding is inadequate at intensities of 2mA. *PLoS One* 2012;7:e47514.
- Ambrus GG, Al-Moyed H, Chaieb L, Sarp L, Antal A, Paulus W. The fade-in - short stimulation - fade out approach to sham tDCS - reliable at 1 mA for naive and experienced subjects, but not investigators. *Brain Stimul* 2012;5:499–504.
- Palm U, Reisinger E, Keeser D et al. Evaluation of sham transcranial direct current stimulation for randomized, placebo-controlled clinical trials. *Brain Stimul* 2013;6: 690–695.
- Guleyupoglu B, Febles N, Minhas P, Hahn C, Bikson M. Reduced discomfort during high-definition transcutaneous stimulation using 6% benzocaine. *Front Neuroeng* 2014;7:28.
- Dundas JE, Thickbroom GW, Mastaglia FL. Perception of comfort during transcranial DC stimulation: effect of NaCl solution concentration applied to sponge electrodes. *Clin Neurophysiol* 2007;118:1166–1170.
- Guarienti F, Caumo W, Shiozawa P et al. Reducing transcranial direct current stimulation-induced erythema with skin pretreatment: considerations for sham-controlled clinical trials. *Neuromodulation* 2015;18:261–265.
- Ambrus GG, Antal A, Paulus W. Comparing cutaneous perception induced by electrical stimulation using rectangular and round shaped electrodes. *Clin Neurophysiol* 2011;122:803–807.
- Turi Z, Ambrus GG, Ho KA, Sengupta T, Paulus W, Antal A. When size matters: large electrodes induce greater stimulation-related cutaneous discomfort than smaller electrodes at equivalent current density. *Brain Stimul* 2014;7:460–467.

14. Brunoni AR, Amadera J, Berbel B, Volz MS, Rizzerio BG, Fregni F. A systematic review on reporting and assessment of adverse effects associated with transcranial direct current stimulation. *Int J Neuropsychopharmacol* 2011;14:1133–1145.
15. Portney LG, Watkins MP. *Foundations of clinical research: applications to practice*. Upper Saddle River, NJ: Prentice Hall Health, 2008.
16. Fitzpatrick TB. The validity and practicality of sun-reactive skin types I through VI. *Arch Dermatol* 1988;124:869–871.
17. Miranda PC, Lomarev M, Hallett M. Modeling the current distribution during transcranial direct current stimulation. *Clin Neurophysiol* 2006;117:1623–1629.
18. Wagner T, Fregni F, Fecteau S, Grodzinsky A, Zahn M, Pascual-Leone A. Transcranial direct current stimulation: a computer-based human model study. *Neuroimage* 2007;35:1113–1124.
19. Saturnino GB, Antunes A, Thielscher A. On the importance of electrode parameters for shaping electric field patterns generated by tDCS. *Neuroimage* 2015; 120:25–35.
20. Ferrell WR, Wong BB, Lockhart JC, Ramsay JE. Gender differences in regional cutaneous microcirculatory responses to capsaicin. *Fundam Clin Pharmacol* 2004;18:195–200.
21. Kessler SK, Turkeltaub PE, Benson JG, Hamilton RH. Differences in the experience of active and sham transcranial direct current stimulation. *Brain Stimul* 2012;5:155–162.
22. Poreisz C, Boros K, Antal A, Paulus W. Safety aspects of transcranial direct current stimulation concerning healthy subjects and patients. *Brain Res Bull* 2007;72:208–214.

SUPPORTING INFORMATION

Additional supporting information may be found online in the supporting information tab for this article.