Anodal tDCS over primary motor cortex provides no advantage to learning motor sequences via observation

Apsvalka, Dace; Ramsey, Richard; Cross, Emily

Neural Plasticity

Published: 01/01/2018

Peer reviewed version

Anodal tDCS over primary motor cortex provides no advantage to learning motor sequences via observation

Dace Apšvalka[1,2], Richard Ramsey[1]* & Emily S. Cross[1,3]*

Affiliations:
[1] Social Brain in Action Laboratory
Wales Institute for Cognitive Neuroscience
School of Psychology
Bangor University, Wales

[2] MRC Cognition and Brain Sciences Unit
University of Cambridge
Cambridge, England

[3] Institute of Neuroscience and Psychology
University of Glasgow
Glasgow, Scotland

* These two corresponding authors contributed equally:
Richard Ramsey & Emily S. Cross
Social Brain in Action Laboratory
Wales Institute for Cognitive Neuroscience
School of Psychology
Bangor University, WALES
r.ramsey@bangor.ac.uk / emily.cross@glasgow.ac.uk

Abstract: 199 words
Main Text: 7890 words
References: 57 words
Figures: 4
Tables: 3

Abstract

When learning a new motor skill, we benefit from watching others. It has been suggested that observation of others' actions can build a motor representation in the observer, and as such, physical and observational learning might share a similar neural basis. If physical and observational learning share a similar neural basis, then motor cortex stimulation during observational practice should similarly enhance learning by observation as it does through physical practice. Here we used transcranial direct current stimulation (tDCS) to address whether anodal stimulation to M1 during observational training facilitates skill acquisition. Participants learned keypress sequences across four consecutive days of observational practice whilst receiving active or sham stimulation over M1. The results demonstrated that active stimulation provided no advantage to skill learning over sham stimulation. Further, Bayesian analyses revealed evidence in favour of the null hypothesis across our dependent measures. Our findings therefore provide no support for the hypothesis that excitatory M1 stimulation can enhance observational learning in a similar manner to physical learning. More generally, the results add to a growing literature that suggests the effects of tDCS tend to be small, inconsistent and hard to replicate. Future tDCS research should consider these factors when designing experimental procedures.

Keywords: tDCS, observational learning, primary motor cortex, motor learning, Bayesian analysis

1. Introduction

Learning new motor skills is crucial for successful interactions with one's environment. However, the neural mechanisms that underlie skill learning in the human brain are not well known. Most prior neuroscience research has investigated skill acquisition through physical practice. For example, prior studies have shown that motor skill learning can be facilitated by applying anodal transcranial direct current stimulation (tDCS) to the primary motor cortex (M1) during physical practice of new skills (for reviews, see [1–4]). These results suggest that M1 plays a functional role when learning novel motor skills through physical practice. However, motor learning also occurs when watching others perform actions in the absence of physical practice [5]. To date, the extent to which the motor system operates similarly in physical and observational learning remains unclear. In the present study, therefore, we use anodal tDCS over M1 to determine the extent to which stimulation of the motor system may also facilitate learning via observation.

Motor learning increases excitability of M1 and strengthens synaptic connections within M1 through long-term potentiation (LTP)-like mechanisms [6–8]. Similarly, applying an anodal current over M1 via tDCS increases excitability of cortical neurons under the surface area of the electrode [9,10] and the aftereffects of stimulation are believed to be related to LTP-like changes in synaptic plasticity [11]. In addition, combining anodal tDCS over M1 with a motor learning task (so-called "online" stimulation) has been shown to facilitate motor learning [1–4], which suggests that there may be additive effects of combining stimulation techniques with learning paradigms.

Physical practice of motor movements is not essential to learn new skills; motor skills can also be learned by watching others perform actions [5]. Although many studies have shown that motor skills can be learned via observation, the specific neural mechanisms that are required to translate visual input into motor programs are not well understood [12,13]. Several theories suggest that action observation engages an observer's own motor system by establishing internal representations of the motor programs required to perform the action (for a review, see [14]). Engagement of premotor and parietal cortices is consistently reported during both action execution and action

observation, and these two brain regions form the core of the so-called human mirror system [15,16].

Although M1 is not part of the premotor-parietal mirror system, accumulating evidence suggests that it plays an important role in action observation, as well as learning by observation. Electrophysiological recordings in monkeys have shown that cells in M1 exhibit mirror-like properties, meaning that they respond to both observed and executed movements [17–19]. In humans, repetitive transcranial magnetic stimulation (TMS) over M1, which temporary disrupts function, effectively inducing a short-lived "virtual lesion", reduces the benefits of motor learning by observation [20]. Further, M1 engagement during observation might be a critical determinant for the success of motor learning via observation [21]. If M1 plays a similar functional role in observational learning as it does in physical learning, increasing M1 excitability during observational learning should facilitate skill acquisition in a similar manner as that reported for learning by physical practice.

Here, we investigate whether applying anodal tDCS over M1 during observational practice facilitates acquisition and retention of a keypress sequence learning task. We hypothesise that observational practice coupled with anodal tDCS should have beneficial effects on learning compared to observational practice alone, as has been previously reported for learning by physical practice [1–4]. Such a pattern of findings would support the view that M1 plays a similar functional role in learning via observation and physical practice, thus further illuminating the functional mechanisms supporting action and perception links in motor learning.

2. Method

Participants

Fifty-five participants consented to participate in the study. Five participants did not finish all sessions, including the post-training testing sessions. The five participants were thus excluded from

analyses as they did not have post-training performance measures that were critical for testing our hypothesis. The final sample comprised 50 participants: 14 males and 36 females, 18 to 30 years old (M = 20.60 years, SD = 2.40). All participants were right-handed (based on self-report) Bangor University student volunteers with normal or corrected-to-normal vision and no history of neurological or psychiatric disorders. Participants reported no contraindications to TMS or tDCS (personal/family history of epilepsy or seizures, metal or implants in the body, frequent headaches, history of serious head injury, heart disease, possibility of being pregnant), and were not taking any medication that affects brain function (e.g., anti-epileptic medication, tranquilizers, or anti-depressants). Prior to the first stimulation session, participants were assigned to the sham (N = 24) or active stimulation (N = 26) group (see section 1.2.3 for assignment procedure). No significant differences existed between the groups in terms of demographics and baseline performance (summarised in Table 1). Participants provided their written informed consent prior to beginning all experimental procedures and either received eight course credits or were paid £30 for their participation following completion. The study was conducted in accordance with the Declaration of Helsinki and all procedures were approved by the Ethics Committee of the School of Psychology at Bangor University (protocol 2016-15675) and the UK Ministry of Defence Research Ethics Committee (protocol 735/MODREC/15).

Stimuli and procedure

A keypress sequence learning paradigm was implemented, based on the task used by Wiestler and Diedrichsen [22]. A standard QWERTY black computer keyboard had the Q 3 4 5 and Y keys covered with red tape and all surrounding keys removed. In pre- and post-training sessions, participants were required to press the red keys with the five fingers of their left hand in a specified order. During the observational training tDCS sessions, participants watched videos of the experimenter performing the keypress task. For the video recordings, a similar keyboard was used with the only difference that the sides of the five keys were covered in yellow to improve the

visibility of the key being pressed. Stimuli presentation and response recordings were performed using MATLAB 8.3.0 (The MathWorks, MA, USA) and Psychophysics Toolbox 3.0.12 [23]).

Keypress Sequences

The same set of 12 five-element keypress sequences was used previously by Wiestler and Diedrichsen [22]. Each sequence required the five fingers of the left hand to be pressed once in a sequential order, with each of the 12 sequences featuring a different order with no more than three adjacent finger-presses in a row. All sequences were matched for difficulty, based on a pilot experiment [22]. For each participant, from the set of 12 sequences, four sequences were randomly allocated to the Trained condition, and four other sequences were allocated to the Untrained condition. The remaining four sequences remained unused.

Videos

For the observational training sessions, 13-second videos were created showing the experimenter's left hand from a first-person perspective, slightly tilted to the right (see Figure 1A and Supplementary Materials (video)). Each video showed the experimenter executing one sequence five times, with naturally varying breaks between each sequence repetition to ensure a more authentic presentation of the performance. For the same reason, for each sequence five different video versions were recorded. This ensured closer to natural performance variation of the same sequence. An additional video version for each sequence was created where one of the five sequence executions was incorrect. This resulted in 72 videos in total.

Sequences were executed at an intermediate performance level, which was determined by behavioural pilot test results, where the average time to complete a correct sequence execution was 2.29 seconds (Pilot: $N = 17$, $M = 2.29$ s, $SE = 0.14$). Each original video, showing five repetitions of the same sequence, was slightly sped up or slowed down (±10%) to make it exactly 13 seconds

long. Consequently, the authenticity of movement performance was somewhat reduced, but the relative variability within the video remained intact. The average length of time for a single sequence execution in the videos was 2.3 seconds. The videos were presented on a computer monitor in full colour on a black background. The frame rate was 29 frames per second with the resolution of 600 x 526 pixels, showing approximately natural hand size.

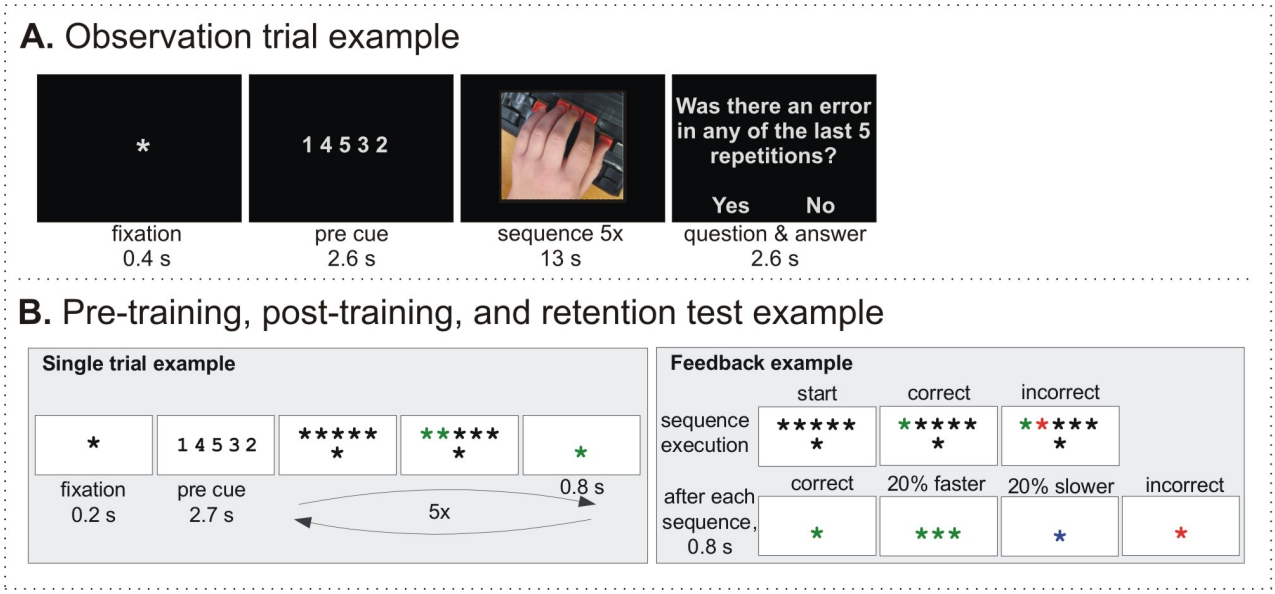


Figure 1. Sequence learning and testing elements. A. Observation trial example. A sequence cue was followed by a video showing a hand executing the sequence five times, either correctly or incorrectly. Occasionally a question was asked whether there was an error in any of the five repetitions, and a response had to be made. B. Execution trial example. A cued sequence had to be memorised and then executed five times while receiving performance feedback.

Training and Testing Procedures

Participants were required to watch and learn four different 5-element keypress sequences performed by a model with the left (non-dominant) hand. Participants underwent six testing sessions (Figure 2). Consecutive multiple-day stimulation sessions were administered because they generally produce higher tDCS effects compared to single stimulation sessions [1], showing a

cumulative increase in cortical excitability [24] and improved motor skill consolidation and retention [25,26]. On the first day of testing (day 1), participants' left-hand motor area was localised with TMS (see below for details). After the localisation procedure, participants received task instructions and completed three single sequence execution trials to ensure they understand the task. The familiarisation procedure was followed by a pre-test, which was followed immediately by the first observational practice session. The observational practice sessions continued the next three consecutive days (day 2 to day 4). For most participants, sessions were arranged at the same time of the day as the first practice session (with 1.5 to 2.5-hour difference for three participants in the sham group, and 0.5 to 1.5-hour difference for four participants in the active stimulation group). The day after completing the final observational practice session, participants performed a post-test to assess learning (day 5). One week later, they returned to the lab one final time to perform a retention test to assess memory for the different sequences (day 12).

Stimuli presentation and response recordings were performed using MATLAB 8.3.0 (The MathWorks, MA, USA) and Psychophysics Toolbox 3.0.12 [23]). All scripts are available at Github (https://github.com/dcdace/2017_tDCS).

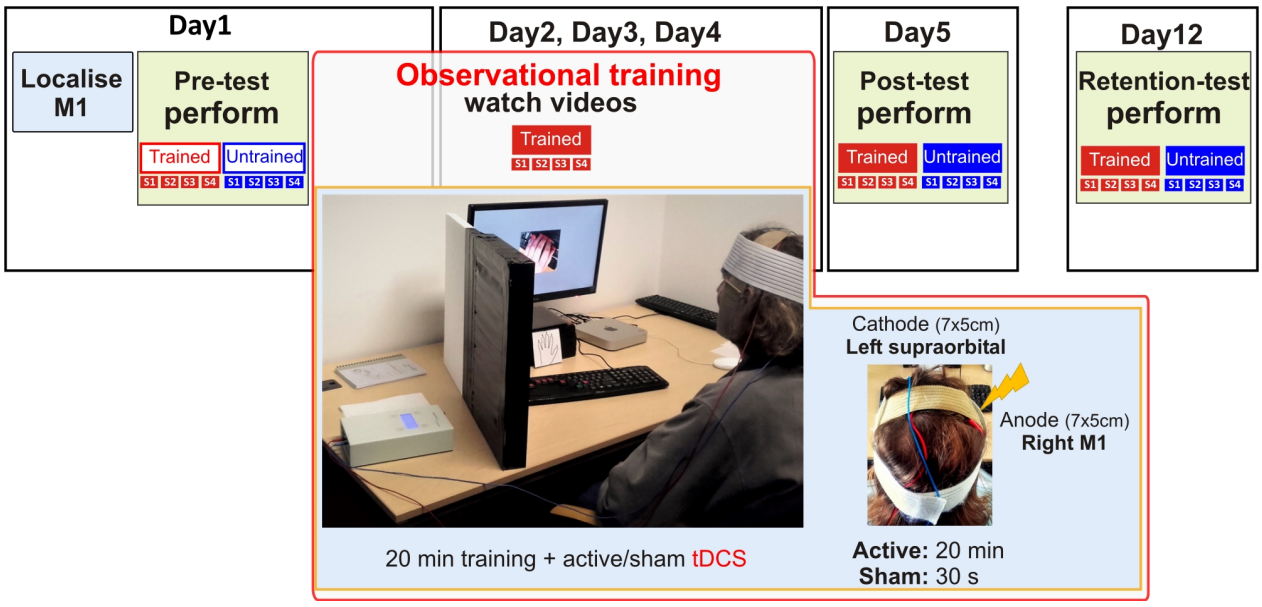


Figure 2. Experimental procedure. The experiment involved pre-test, four 20-minute-long training sessions coupled with tDCS, post-test, and retention-test. In the pre-, post- and

retention-tests, participants executed eight keypress sequences (four of them to be trained, the other four untrained) with the left (non-dominant) hand. In the training sessions, participants watched videos of a model's left hand executing four of the eight sequences. During training, participants received either sham or active (1 mA) 20-minute stimulation over the right motor cortex (35cm2 large area centred on the left-hand motor area M1).

Testing sessions

In the pre- and post- and retention performance sessions, participants performed four Trained and four Untrained sequence execution trials in a random order with the left hand. Each trial consisted of five repetitions of the same sequence. All trial-related information was presented centrally at the bottom of the screen against a grey background. A trial started with a black fixation cross (0.2 s), followed by the sequence cue presented as five digits (2.7 s) that indicated from right to left which key to press: "1" – the right-most key pressed with the thumb; "5" – the left-most key pressed with the little finger. After the cue, the digits were replaced by the fixation cross and five black asterisks above it. This served as a "go" signal to execute the memorised sequence five times as quickly and accurately as possible. If the correct key was pressed, the corresponding asterisk on the screen turned green, if a wrong key was pressed, the asterisk turned red (see Figure 1B).

After executing a single sequence, the central fixation cross changed colour to provide feedback on the performance (0.8 s): green – correct sequence execution; red – incorrect sequence execution; blue – correct, but executed 20% slower than the median execution time (ET) in the previous trials; three green asterisks – correct and executed 20% faster than the median ET in the previous trials. After this short feedback, all asterisks turned black signalling the start of the next execution trial. After five executions of the same sequence, the trial ended and the next sequence was cued.

Participants' performance was assessed as the average sequence initiation time, execution time, and error rate for the four trained (to-be-trained) and the four untrained sequences. The error rate was measured as the percentage of incorrect sequence executions. Incorrectly executed trials were excluded from initiation time and execution time measurements. The initiation time was measured as the duration between the "go" signal and the first keypress. The execution time was measured as the duration between the first and fifth keypresses.

Observational training sessions

During the observational training sessions, participants received either sham or active brain stimulation while watching videos of the model's left hand executing four sequences. Each video showed five repetitions of the same sequence. A trial started with a 5-digit cue (for 2.6 s), indicating the sequence to be executed, followed by a video (13 s) showing five executions of the cued sequence. Participants were instructed to watch whether the hand executed the correct sequence all five times. Occasionally participants were asked whether there was an error in any of the five executions – the error question.

Each practice session was divided into three blocks, separated by a one-minute rest period. Within each block, 20 videos were presented in a random order: each sequence video four times, and one 'error video' (with at least one incorrect sequence execution) for each sequence. The error question was asked randomly 5-7 times per block. At the end of each block, participants received feedback on how accurately they spotted the incorrect sequence executions. During each session, participants saw a correct execution of each sequence at least 60 times (3 blocks, 4 videos per block, 5 repetitions per video, plus some correct repetitions in the 'error video'). The whole training session lasted approximately 20 minutes and was coupled with 20-minutes of sham or active tDCS.

Motor cortex stimulation

Right M1 localisation

Single-pulse TMS was used to localise the left-hand motor area. The TMS coil was positioned on the right hemisphere, slightly anterior and ventral to the vertex of the skull to induce a muscle twitch in the relaxed fingers of the left-hand. The stimulator output was started at 45% and increased in steps of 2-5% until a visible twitch was observed. The stimulator output never exceeded 80% and participants received no more than 20 total pulses in total, with an inter-pulse interval kept to at least 5 seconds. The optimal location at which TMS evoked a just-noticeable finger twitch was marked on the participant's scalp with a surgical marker. For nine participants, a visible twitch was not observed following this procedure and the motor hand area was instead marked per position C4 of the EEG 10-20 system (after [27]). The localisation procedure was performed only on the first testing session and the marked M1 location was renewed with the surgical marker before each stimulation session.

The nine participants whose M1 area could not be localised using TMS were assigned to the sham group as the precise location of the stimulated area was not critical for sham stimulation. We acknowledge that random assignment, independent of localisation procedure, would have been a better approach. The reasons why we could not evoke a visible twitch in some participants may include extent of representation of the hand area and/or its accessibility via the cortical surface. To ensure that any group differences are not driven by the non-random assignment to groups, we repeated the main analyses of observational training and stimulation effects with the nine non-TMS localised participants excluded. The results of this analysis (see Supplementary Materials 1) suggest that non-random group assignment did not systematically bias our findings.

Stimulation parameters

We performed a single-blinded protocol. Participants were semi-randomly assigned to the sham or active stimulation group, keeping gender balanced between the groups and ensuring that the motor hand area of the active group was localised using TMS procedure described above. Participants were told that they would receive stimulation for up to 20 minutes, not specifying the exact length of the stimulation and not revealing the existence of two stimulation groups. During

each practice session, the sham group received 30 seconds and the active group received 20 minutes of tDCS (c.f. [28]).

A 1 mA constant current was delivered using a battery-driven DC-stimulator Plus (NeuroConn GmbH, Ilmenau, Germany) via a pair of conductive-rubber electrodes placed into saline-soaked sponges (7 x 5 cm; 0.029 mA/cm^2 current density). The electrodes were secured with elastic bands. The contact impedance was monitored throughout the session to ensure it stays below 15 kΩ.

The anode was centred over the previously marked right M1. Due to the electrode size, the stimulation likely extended into premotor and anterior parietal cortices as well. The cathode was placed on the left supraorbital ridge (see photographs in Figure 2). The current was ramped up to 1 mA over 10 seconds, held constant for either 30 seconds (sham) or 20 minutes (active), and then ramped down over 10 seconds. This method is recommended to reliably blind participants to stimulation condition and ensure similar sensations for sham and active stimulation groups [28].

The observational training task started one minute after stimulation onset, to allow time for participants to adapt to the stimulation sensations and to ensure they felt comfortable with carrying on with the task. The stimulation ended about one minute before the end of the task.

Sensations questionnaire

After each training session, participants provided information on the intensity of experienced sensations (itching, pain, burning, heat, pinching, metallic taste, fatigue), the timing of any discomfort (when did the discomfort begin and how long did it last?), and the perceived impact of the stimulation on their performance (adapted from [29]). At the end of the experiment (day 12) participants were debriefed and asked whether they think they received sham or active stimulation.

Data analysis

All statistical analysis was performed using R (v3.3.2, 2016-10-31) in RStudio (v1.0.136, 2016-12-21, RStudio, Inc, Boston, MA). Graphs were produced in MS Excel 2016 (Microsoft, Redmond,

289 WA, USA). The Excel files, raw data and scripts with all analysis procedures and for reproducing
290 results are available at https://github.com/dcdace/2017_tDCS.

291 Given the total sample size of 50, the study had 80% power to detect effects of tDCS that are
292 conventionally considered large (Cohen's d = 0.71; the effect size was estimated with a *power.t.test*
293 function in R for a two-sample, one-sided t-test with 25 observations per group). Three previous
294 multiple stimulation session (3-5 consecutive days, 20-25 min per day, 1-2 mA, ~12.5 participants
295 per group) M1 anodal-tDCS physical training studies reported large tDCS effects ranging from 0.95
296 to 1.33 Cohen's d [25,26,30].

297 The effect of observational training on sequence-specific learning was assessed as a post-training
298 difference between the trained and untrained sequence initiation time, execution time, and error
299 rate. For the sequence initiation time and execution time, we measured a percentage difference
300 ([(untrained/trained)-1]*100), but for the error rate (to avoid dividing by zero), we calculated an
301 absolute difference (untrained-trained) between the trained and untrained sequences. Results for all
302 of these measures are plotted in Figure 4A-C (raw performance measures are provided in
303 Supplementary Materials 2). To correct for possible pre-training differences, we performed a linear
304 regression between the pre-training difference (predictor) and the post-training difference (outcome;
305 see Figure 4E for an example plot). The intercept of the regression line was used as a measure of the
306 post-training difference between trained and untrained sequences, controlling for possible pre-
307 training differences. This method reduces the noise of unwanted differences in the difficulty of
308 trained and untrained sequences and thus allows a more accurate measurement of the training effect.

309 For the assessment of tDCS effects, we complemented null hypothesis significance testing with a
310 Bayesian analysis to provide evidence for the null result. We used the *generalTestBF* function of
311 the R package BayesFactor v0.9.12-2 [31] with its default parameters. The Bayesian test produced a
312 Bayes factor to allow quantification of evidence in favour of either the alternative (BF_{10}) or null
313 (BF_{01}) hypothesis based on prior beliefs and the present data. To describe the Bayes factor results
314 we used Jeffreys' [32] classification scheme and reported both BF_{10} and BF_{01}. Jeffreys proposed

benchmarks for evaluating the strength of evidence as anecdotal (BF_{10} 0-3), substantial (BF_{10} 3-10), strong (BF_{10} 10-30). These Bayes Factors can be readily interpreted as a ratio of evidence in favour of the experimental effect compared to the null effect. For example, a BF_{10} of 3 would represent that the experimental effect is three times more likely than the null, given the data.

The significance threshold for all statistical comparisons was $p < 0.05$. If not specified otherwise, all sample means are reported with their 95% confidence intervals in squared brackets. Confidence intervals for two-tailed tests were calculated as SE*2.07 for the sham group (df 23) and SE*2.06 for the active group (df 25), whereas confidence intervals for one-sided tests were calculated as SE*1.71 for df 23 and df 25 [33].

Results

Group characteristics and sensations during training sessions

Gender proportion between sham and active stimulation groups was compared using a Chi-square test. Mann-Whitney U tests were used to compare group age and experienced sensations during the training sessions. Participants' baseline performance (pre-training average of trained and untrained sequences) was compared using a two-tailed independent measures t-test. Results are summarised in Table 1. The reported sensations for each training day are summarised in Table 2 and averages of all training days plotted in Figure 3.

There were no differences in gender, age, and baseline performance between the groups. On average, both groups reported mild to moderate levels of discomfort during stimulation with no significant difference between the groups (Table 1; Figure 3A). Although the active stimulation group did report a small but significantly larger impact of stimulation on performance than the sham group, the perceived impact for both groups was closest to zero ("no impact") (Table 1; Figure 3B). Finally, sensations lasted significantly longer for the active compared to the sham group (Figure

339 3C), with average sensations stopping between “quickly” and “in the middle of the block” across
340 both groups.

341 The reported sensation data, therefore, shows that there were small but significant sensation
342 differences between the sham and active stimulation groups. The sham protocol should provide
343 comparable sensations to the active stimulation protocol [28]. However, small but significant
344 sensation differences between the stimulation groups, using comparable protocols to ours, have
345 been reported before [29], raising an issue that the widely accepted sham stimulation procedure may
346 not be sufficiently effective.

347 Following the recommendation of Fertonani et al. [29], at the end of the experiment, we asked
348 participants whether they think they received sham or active stimulation. In total, 54% thought they
349 received active stimulation, 32% thought they received sham stimulation and 14% did not know.
350 There was no significant difference between the two groups in terms of which kind of stimulation
351 they thought they received ($\chi^2 = 1.24$, $p = 0.538$), thus confirming the success of the blinding
352 procedure.

353 Table 1. Group characteristics and self-reported sensations during training sessions.

	Sham (N = 24)	**Active (N = 26)**	**Group difference (p-value, effect size)**
Demographics			
Gender (male:female)	8:16	6:20	0.623
Age (years; M ±SD)	20.96 ±2.97	20.27 ±1.71	0.446, d = 0.217
Baseline performance			
Pre-test initiation time (s; M ±SD)	0.77 ±0.25	0.89 ±.30	0.117, d = 0.455
Pre-test execution time (s; M ±SD)	1.92 ±0.57	2.02 ±0.68	0.590, d = 0.153
Pre-test error rate (%; M ±SD)	25 ±13	30 ±15	0.203, d = 0.366
Sensations			
Strongest (M ±SD)	1.23 ±0.49	1.46 ±0.79	0.478, d = 0.202
Affected (M ±SD)	**0.16 ±0.32**	**0.30 ±0.36**	**0.037, d = 0.618**
Lasted (M ±SD)	**1.14 ±0.48**	**1.79 ±0.71**	**0.001, d = 1.04**

Shaded fields highlight variables that significantly differed between the sham and active stimulation groups. Strongest: the strongest reported sensation intensity level (0-4); Affected: how much did sensations affect performance (0-4); Lasted: when did the discomfort stop (0-3)

354 Table 2. Frequencies of self-reported sensations during the training sessions.

The strongest intensity of discomforting sensations

	Day 1					**Day 2**					**Day 3**					**Day 4**				
	0: none, 1: mild, 2: moderate, 3: considerable, 4: strong																			
	0	**1**	**2**	**3**	**4**	**0**	**1**	**2**	**3**	**4**	**0**	**1**	**2**	**3**	**4**	**0**	**1**	**2**	**3**	**4**
Sham	1	12	10	1	-	4	11	8	1	-	2	15	7	-	-	5	14	5	-	-
Active	2	11	8	3	2	2	18	2	3	1	2	15	4	2	3	3	16	5	2	-

How much did the sensations affect performance?

	Day 1					Day 2					Day 3					Day 4				
	0: not at all, 1: slightly, 2: considerably, 3: much, 4: very much																			
	0	**1**	**2**	**3**	**4**	**0**	**1**	**2**	**3**	**4**	**0**	**1**	**2**	**3**	**4**	**0**	**1**	**2**	**3**	**4**
Sham	19	5	-	-	-	20	4	-	-	-	22	2	-	-	-	21	2	1	-	-
Active	18	7	-	1	-	20	6	-	-	-	18	7	1	-	-	20	6	-	-	-

When did the discomfort stop?

	Day 1				Day 2				Day 3				Day 4			
	ns: no sensations, 1: quickly, 2: middle of the block, 3: end of the block															
	ns	**1**	**2**	**3**	**ns**	**1**	**2**	**3**	**ns**	**1**	**2**	**3**	**ns**	**1**	**2**	**3**
Sham	1	15	4	4	4	14	4	2	2	19	3	-	5	18	-	1
Active	2	6	9	9	2	11	7	6	2	8	7	9	3	9	8	6

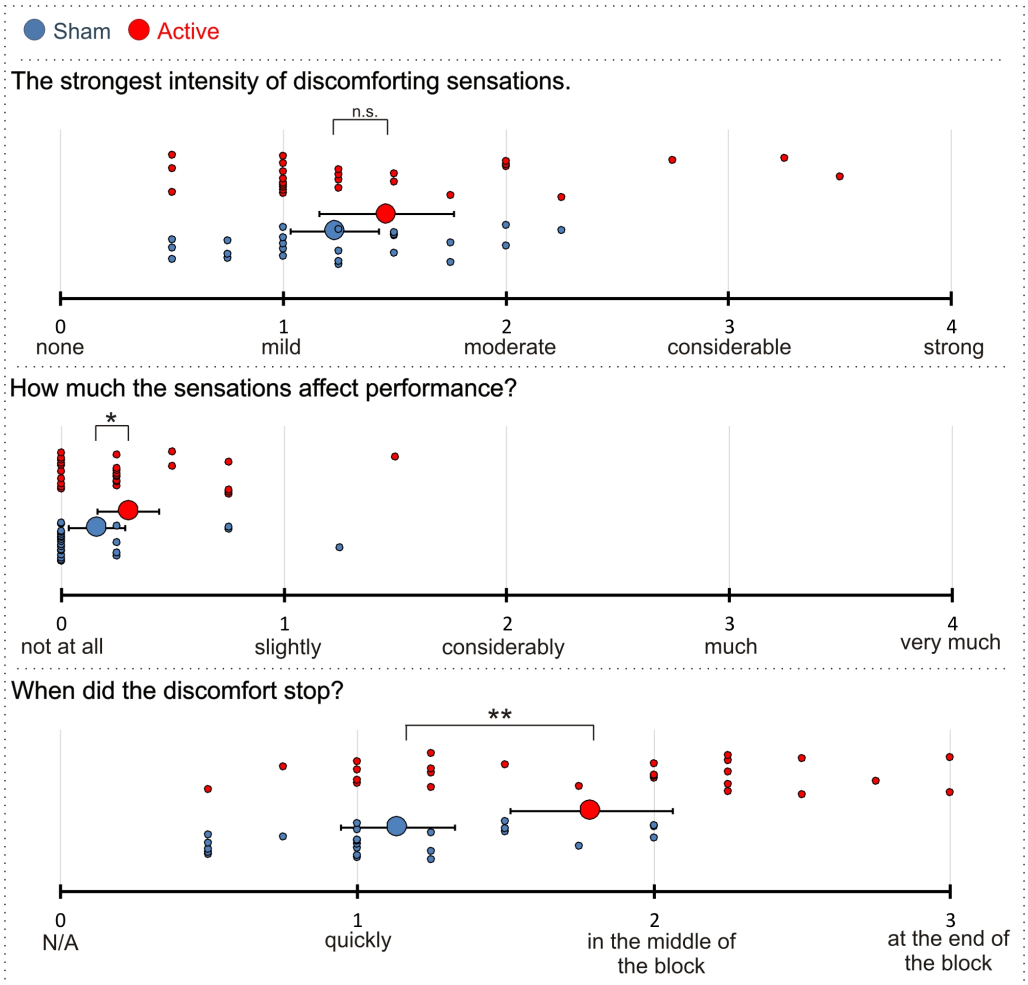


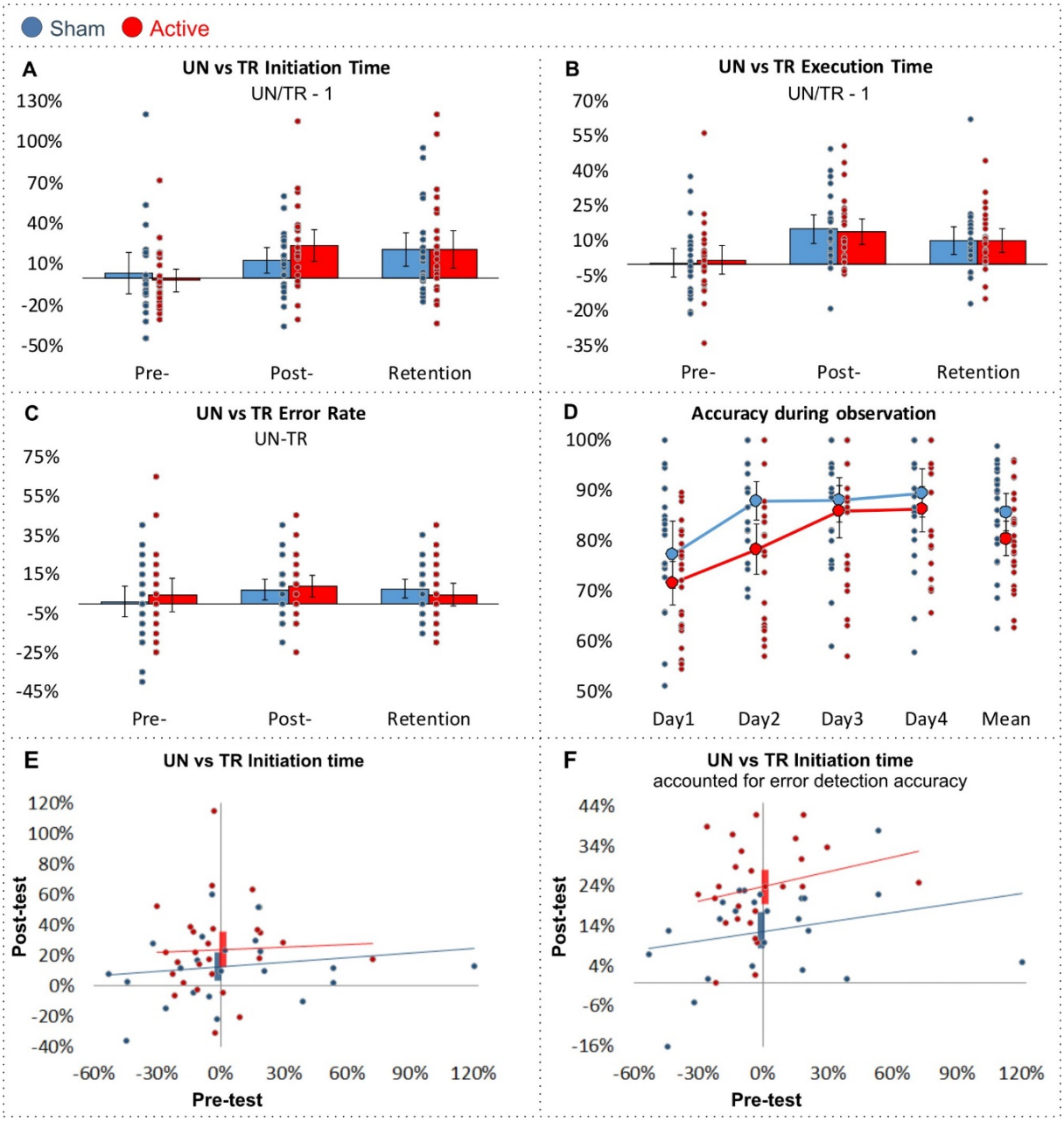
Figure 3. The 4-day average values of self-reported sensations during the training sessions. Large dots: group averages; small dots: individual participant values; red: active; blue: sham; error bars: 95% CI, two-tailed; * $p < 0.05$, ** $p < 0.01$, two-tailed.

Accuracy during training sessions

During the observational practice sessions, attention to the task was assessed by accurate responses to the error question (spotting incorrectly executed sequences). The overall accuracy was 83%, significantly higher than a 50% chance level (yes/no answers; $t_{49} = 24.61$, $p < 0.001$, two-tailed), confirming that participants paid attention to the task. The average accuracies for each group and day are plotted in Figure 4D. On average, across the four training days, the sham group performed better (M = 86% [82%, 90%]) than the active group (M = 81% [77%, 85%]), with a marginally significant difference between the two groups (Welch two sample t-test for non-equal variance: $t_{47.27} = 1.99$, $p = 0.052$, two-tailed, $d = 0.56$).

The small difference in error detection accuracy between the groups was an unexpected finding. It cannot be ruled out that anodal tDCS of M1 had some negative effects on the error detection accuracy. However, we do not have any *a priori* or theoretical grounds to support this suggestion. Another possibility is that the error detection accuracy was influenced by the discomforting sensations during the training sessions that, as reported above, affected the stimulation group more than the sham group. This possibility is supported by a significant negative correlation between the average error detection accuracy and the average self-report on how much performance was affected by the discomforting sensations (Kendall's tau-b = -0.296, $p = 0.008$, two-tailed; across both groups).

The lower error detection accuracy for the active stimulation group raises a possibility that the active group may not have been able to learn from the videos as well as the sham group due to stimulation-related discomfort and consequent impact on attention. To account for this possibility, we complement the planned analysis with an exploratory analysis that includes mean error detection accuracy as a covariate when assessing the stimulation effect.



385

386 **Figure 4. Performance results. Pre-, post-, and retention-test difference in initiation time**
387 **(A), execution time (B), and error rate (C) between trained (TR) and untrained (UN)**
388 **sequences for sham (blue) and active (red) stimulation groups. D. Error detection accuracy**
389 **during observational practice sessions. A-D. Bars and large dots: group averages; small dots:**
390 **individual participant values; error bars: 95% CI (one-tailed for A, B, and C; two-tailed for**
391 **D). E. Regression lines of pre-test (predictor) and the post-test difference between trained and**
392 **untrained sequence initiation times for sham (blue) and active (red) stimulation groups.**
393 **Intercepts of the regression lines represent the predicted post-test difference if the pre-test**

difference is zero. Vertical bars represent 95% CIs (one-tailed) of intercepts F. Same as E, but post-test difference corrected for error detection accuracy during training sessions.

Observational training effects on sequence-specific learning

Both groups showed significant observational training effects at both post-test and retention-test on all three performance measures, with medium to large effect sizes for the performance difference between trained and untrained sequences ($d_z = 0.52 – 1.02$; comparable to previous reports on keypress sequence learning by observation, e.g., [34–36]). The only exception to this pattern of results was that the active stimulation group demonstrated no effect on error rates at the retention-test. Detailed results are provided in Table 3, columns I and II, where B_0 represents the percentage performance improvement from pre-test. All tests in Table 3 are one-tailed as we were testing a directional prediction for the difference between trained and untrained sequences. Furthermore, Supplementary Materials 3 document the extent to which the training manipulation generalised to the untrained sequences, comparing the active and sham stimulation groups.

tDCS effects on sequence-specific learning by observation

Primary analysis

The effect of stimulation on sequence-specific learning was assessed by comparing observational training effects (the post-training ~ pre-training regression line intercepts) between the sham and active stimulation groups. The performed analysis of covariance (ANCOVA) did not reveal any significant difference between the two groups on any of the three measures either at post-test or retention-test (Figure 4E plots post-test initiation time results; **see Supplementary Materials 4 for ANCOVA results of the raw means**). The Bayes factor analyses yielded anecdotal to substantial evidence against the stimulation effect. Detailed results are provided in Table 3, column III (reporting significance of the group as a predictor variable for the training effect).

Secondary analysis: accounting for error detection accuracy

Due to error detection differences between the groups, in an exploratory analysis we added mean error detection accuracy as a covariate to the previous ANCOVA model and repeated the group comparison analysis. This exploratory analysis revealed evidence for the stimulation effect on the percentage difference between trained and untrained sequence initiation times at post-test. Compared to the sham group, the active stimulation group showed greater difference on this measure (see Figure 4F). The error detection accuracy significantly predicted the outcome (β = 0.431, p = 0.003; the better the accuracy during training, the faster initiation time of trained relative to untrained sequences at post-test). All other measures showed substantial to strong evidence against the stimulation effect when accounting for the error detection accuracy. Detailed results are provided in Table 3, column IV (reporting significance of the group as a predictor variable for the training effect accounting for the error detection accuracy).

Table 3. Observational practice effects and tDCS effects on sequence-specific learning.

		I	II	III	IV
		Observational training effect (trained vs. untrained performance)		**Primary results**	**Secondary results**
		Sham	**Active**	**tDCS effect (group difference)**	**tDCS effect (group difference), accounting for accuracy during training sessions**
Initiation time	**Post**	$t_{(22)} = 2.65$, $p = 0.008$, $B_0 = 13\%$, $d_z = 0.54$.	$t_{(24)} = 4.02$, $p < 0.001$, $B_0 = 24\%$, $d_z = 0.79$.	$t_{(47)} = 1.50$, $p = 0.072$, $d = 0.44$, anecdotal evidence against the effect ($BF_{10}/BF_{01} = 0.70/1.43$).	$t_{(46)} = 2.48$, $p = 0.008$, $d = 0.73$, anecdotal evidence for the effect ($BF_{10}/BF_{01} = 2.41/0.41$).
	Ret.	$t_{(22)} = 3.21$, $p = 0.002$, $B_0 = 21\%$, $d_z = 0.66$.	$t_{(24)} = 2.87$, $p = 0.004$, $B_0 = 21\%$, $d_z = 0.56$.	$t_{(47)} = 0.05$, $p = 0.480$, $d = 0.01$, substantial evidence against the effect ($BF_{10}/BF_{01} = 0.29/3.49$).	$t_{(46)} = 0.01$, $p = 0.496$, $d = 0$, substantial evidence against the effect ($BF_{10}/BF_{01} = 0.29/3.45$).
Execution time	**Post**	$t_{(22)} = 5.02$, $p < 0.001$, $B_0 = 15\%$, $d_z = 1.02$.	$t_{(24)} = 4.75$, $p < 0.001$, $B_0 = 14\%$, $d_z = 0.93$.	$t_{(47)} = -0.37$, $p = 0.355$, $d = 0.11$, substantial evidence against the effect ($BF_{10}/BF_{01} = 0.30/3.31$).	$t_{(46)} = -0.49$, $p = 0.312$, $d = 0.15$, substantial evidence against the effect ($BF_{10}/BF_{01} = 0.31/3.20$).
	Ret.	$t_{(22)} = 4.02$, $p < 0.001$, $B_0 = 10\%$, $d_z = 0.82$.	$t_{(24)} = 3.99$, $p < 0.001$, $B_0 = 10\%$, $d_z = 0.78$.	$t_{(47)} = -0.06$, $p = 0.475$, $d = 0.02$, substantial evidence against the effect ($BF_{10}/BF_{01} = 0.28/3.55$).	$t_{(46)} = -0.02$, $p = 0.492$, $d = 0.01$, substantial evidence against the effect ($BF_{10}/BF_{01} = 0.29/3.43$).
Error rate	**Post**	$t_{(22)} = 2.56$, $p = 0.009$, $B_0 = 7\%$, $d_z = 0.52$.	$t_{(24)} = 2.89$, $p = 0.004$, $B_0 = 9\%$, $d_z = 0.57$.	$t_{(47)} = 0.47$, $p = 0.322$, $d = 0.14$, substantial evidence against the effect ($BF_{10}/BF_{01} = 0.31/3.20$).	$t_{(46)} = 0.20$, $p = 0.422$, $d = 0.06$, substantial evidence against the effect ($BF_{10}/BF_{01} = 0.31/3.28$).
	Ret.	$t_{(22)} = 2.99$, $p = 0.004$, $B_0 = 7\%$, $d_z = 0.61$.	$t_{(24)} = 1.45$, $p = 0.08$, $B_0 = 4\%$, $d_z = 0.28$.	$t_{(47)} = -0.81$, $p = 0.210$, $d = 0.24$, anecdotal evidence against the effect ($BF_{10}/BF_{01} = 0.37/2.71$).	$t_{(46)} = -1.05$, $p = 0.149$, $d = 0.31$, anecdotal evidence against the effect ($BF_{10}/BF_{01} = 0.44/2.27$).

Shaded fields highlight non-significant effects. All p-values reported reflect one-tailed tests as we had directional predictions for inflence of training and stimulation on our performance measures. Results are uncorrected for multiple comparisons.

Discussion

We investigated the extent to which anodal tDCS over M1 facilitates motor sequence learning by observation, as previously reported for learning by physical practice [1–4]. Both active and sham stimulation groups benefited from observational practice, replicating previous findings that motor skills can be learned by observation without overt physical practice [5,34–39]. However, active stimulation over M1 did not provide an advantage to learning the motor sequences through observation over and above sham stimulation. Furthermore, Bayesian analyses revealed anecdotal to substantial evidence in favour of the null hypothesis across our dependent measures. Our findings therefore do not provide strong support for the hypothesis that excitatory M1 stimulation can enhance observational learning in a similar manner to physical learning.

Understanding the role of the motor system during observational learning

Although there is a consensus that shared mechanisms exist between action observation and execution [14], the role played by the motor system in observational learning is not clear [12,13]. Indeed, several studies have questioned the notion of motor-driven learning by observation, arguing instead that it is driven by perceptual and cognitive processes [40–42]. It is possible, therefore, that primary motor areas might be engaged during action observation [43–45], but their involvement might not be *critical* in shaping observational learning.

Alternatively, it is possible that the effect of anodal tDCS over M1 during observational learning is smaller than during physical learning and subtler than we could detect in the current study. The current study had 80% power to detect an effect size that is typically considered large (0.71 Cohen's d). Therefore, we have reasonable confidence that we could detect large effects of stimulation, similar to which were reported previously during physical learning, should they exist. In addition, we followed recommended stimulation protocols by stimulating on consecutive days to enhance

457 effects of stimulation [1] and skill learning [25,26] (although see work by Monte-Silva and
458 colleagues [46] that demonstrates the abolishment of LTP-like plasticity in motor cortex when
459 follow-up stimulation occurs 24 hours after initial stimulation). As such, we designed the
460 experiment to increase the likely impact of tDCS on skill learning, but nonetheless report a null
461 result. We suggest that future studies wishing to further explore the role of M1 in observational
462 learning use of a similar protocol with larger sample sizes, in order to increase statistical power to
463 detect smaller effects.

464 The null result we report here adds to a growing set of null results in tasks ranging from
465 working memory [47,48] to language [49,50]. In addition, several recent meta-analyses document
466 conflicting evidence regarding the efficacy of tDCS in a variety of paradigms where effects have
467 previously been reported, as well as growing scepticism regarding a causal role of tDCS in
468 performance enhancement [48,50]. Given concerns over publication bias in general [51] and in the
469 domain of tDCS in particular [52], it is important to report null results in order to provide a less
470 biased estimate of the likely effect sizes that tDCS may have on behaviour. Therefore, balanced
471 reporting of null results (in addition to positive results, such as those observed with tDCS over
472 premotor cortex facilitating observational learning of a motor sequence [53]) will help to build a
473 cumulative science of observational learning and tDCS. For instance, based on the details of the
474 current study, researchers who wish to further explore the relationship between primary motor
475 cortex activity and observational learning will have a more accurate estimate of the likely effect
476 sizes that they might be targeting, which will directly inform power calculations and study design
477 decisions.

478 The current study also provides a platform for future tDCS studies to build upon in other
479 ways. Indeed, there are many avenues that future work could pursue in order to probe the
480 relationship between the motor system and observational learning. For example, the effects of tDCS
481 on observational learning may be task-dependent. Aridan and Mukamel [21] reported a positive
482 relationship between M1 activity during action observation and the success of motor skill learning

via observation only if the observed model's performance was faster than the observer's performance at baseline. The current study used an intermediate model, which may not have been challenging enough to engage the motor system sufficiently. Future studies could use an expert model whose performance consistently exceeds the observer's baseline performance to test this possibility directly.

Follow-up work could also investigate the impact of different stimulation protocols. For example, several reports demonstrate a powerful effect of dual-M1 stimulation on motor learning [30,54], which outperforms unilateral M1 stimulation montages [55–58]. Another possibility to explore concerns the impact of tDCS intensity on motor learning effects. Recent work demonstrates that 1.5 mA, but not 1.0 mA, anodal tDCS over M1 reliably facilitates motor learning [59], which raises the possibility that our stimulation intensity was not optimised to induce reliable results. A further consideration is that small differences were observed in the sensations associated with active compared to sham stimulation, which is consistent with prior research [29]. The impact that such sensation differences have on task performance are worth studying in order to more effectively design sham protocols. Moreover, due to the electrode size (7 x 5 cm), the focality of tDCS stimulation is necessarily imprecise, and stimulation in our study may have extended beyond M1 into nearby premotor and anterior parietal brain regions as well. The modulation of cortical excitability under and between the electrodes is still under debate and investigation [10,60]**.** As these suggestions demonstrate, many different lines of inquiry will be needed to better understand the relationship between motor system engagement and observational learning.

Conclusions

Our results do not support the hypothesis that anodal tDCS over M1 facilitates skill learning through observation to a large degree. The null finding does not necessarily imply that the motor system is not involved in sequence learning by observation. Rather, the results suggest that using

the parameters employed in the current study, anodal tDCS over M1 does not reliably enhance observational learning. Given that no prior study has used tDCS over M1 in an attempt to enhance observational learning, this finding makes an important contribution to the literature by informing future brain stimulation studies and offering a platform upon which to base further investigation into the role of primary motor cortex in observational learning.

Acknowledgements: This work was supported by the Ministry of Defence of the United Kingdom Defence Science and Technology Laboratory [grant number DSTLX-1000083177 to ESC and RR], the Economic and Social Research Council [grant numbers ES/K001884/1 to RR and ES/K001892/1 to ESC], and a Marie Curie Actions/FP7 [CIG11-2012-322256 to ESC]. The authors would like to thank Paul Carter for his help with participant testing, Joshua Payne for the tDCS training, and Kenneth Valyear and Cosimo Urgesi for the TMS training and support. The authors declare no conflicts of interest regarding the publication of this paper.

2 References

[1] F. Hashemirad, M. Zoghi, P.B. Fitzgerald, S. Jaberzadeh, The effect of anodal transcranial direct current stimulation on motor sequence learning in healthy individuals: A systematic review and meta-analysis, Brain Cogn. 102 (2016) 1–12. doi:10.1016/j.bandc.2015.11.005.

[2] J. Reis, B. Fritsch, Modulation of motor performance and motor learning by transcranial direct current stimulation, Curr. Opin. Neurol. 24 (2011) 590–596. doi:10.1097/WCO.0b013e32834c3db0.

[3] E.R. Buch, E. Santarnecchi, A. Antal, J. Born, P.A. Celnik, J. Classen, C. Gerloff, M. Hallett, F.C. Hummel, M.A. Nitsche, A. Pascual-Leone, W.J. Paulus, J. Reis, E.M. Robertson, J.C. Rothwell, M. Sandrini, H.M. Schambra, E.M. Wassermann, U. Ziemann, L.G. Cohen,

533 Effects of tDCS on motor learning and memory formation: a consensus and critical position
534 paper, 2016. doi:10.1101/064204.

535 [4] C. Ammann, D. Spampinato, J. Márquez-Ruiz, Modulating motor learning through
536 transcranial direct-current stimulation: an integrative view, Front. Psychol. 7 (2016) 3–6.
537 doi:10.3389/fpsyg.2016.01981.

538 [5] S. Vogt, R. Thomaschke, From visuo-motor interactions to imitation learning: behavioural
539 and brain imaging studies., J. Sports Sci. 25 (2007) 497–517.
540 doi:10.1080/02640410600946779.

541 [6] D. Spampinato, P. Celnik, Temporal dynamics of cerebellar and motor cortex physiological
542 processes during motor skill learning, Sci. Rep. 7 (2017) 40715. doi:10.1038/srep40715.

543 [7] M.S. Rioult-Pedotti, D. Friedman, J.P. Donoghue, Learning-induced LTP in neocortex.,
544 Science. 290 (2000) 533–536. doi:10.1126/science.290.5491.533.

545 [8] J.N. Sanes, J.P. Donoghue, Plasticity and primary motor cortex., Annu. Rev. Neurosci. 23
546 (2000) 393–415. doi:10.1146/annurev.neuro.23.1.393.

547 [9] M.A. Nitsche, W. Paulus, Excitability changes induced in the human motor cortex by weak
548 transcranial direct current stimulation., J. Physiol. 527 Pt 3 (2000) 633–9. doi:PHY_1055
549 [pii].

550 [10] M.A. Nitsche, L.G. Cohen, E.M. Wassermann, A. Priori, N. Lang, A. Antal, W. Paulus, F.
551 Hummel, P.S. Boggio, F. Fregni, A. Pascual-Leone, Transcranial direct current stimulation:
552 State of the art 2008., Brain Stimul. 1 (2008) 206–23. doi:10.1016/j.brs.2008.06.004.

553 [11] C.J. Stagg, M.A. Nitsche, Physiological basis of transcranial direct current stimulation.,
554 Neuroscientist. 17 (2011) 37–53. doi:10.1177/1073858410386614.

555 [12] D.J. Ostry, P.L. Gribble, Sensory Plasticity in Human Motor Learning, Trends Neurosci. 39
556 (2016) 114–123. doi:10.1016/j.tins.2015.12.006.

557 [13] H.R. McGregor, J.G.A. Cashaback, P.L. Gribble, Functional Plasticity in Somatosensory
558 Cortex Supports Motor Learning by Observing, Curr. Biol. 26 (2016) 921–927.
559 doi:10.1016/j.cub.2016.01.064.

560 [14] A. Gentsch, A. Weber, M. Synofzik, G. Vosgerau, S. Schütz-Bosbach, Towards a common
561 framework of grounded action cognition: Relating motor control, perception and cognition,
562 Cognition. 146 (2016) 81–89. doi:10.1016/j.cognition.2015.09.010.

563 [15] M. Iacoboni, I. Molnar-Szakacs, V. Gallese, G. Buccino, J.C. Mazziotta, Grasping the
564 intentions of others with one's own mirror neuron system, PLoS Biol. 3 (2005) 0529–0535.
565 doi:10.1371/journal.pbio.0030079.

566 [16] G. Rizzolatti, C. Sinigaglia, The functional role of the parieto-frontal mirror circuit:
567 interpretations and misinterpretations, Nat. Rev. Neurosci. 11 (2010) 264–274.
568 doi:10.1038/nrn2805.

569 [17] D. Tkach, J. Reimer, N.G. Hatsopoulos, Congruent activity during action and action
570 observation in motor cortex, J. Neurosci. 27 (2007) 13241–13250.
571 doi:10.1523/JNEUROSCI.2895-07.2007.

572 [18] J. Dushanova, J. Donoghue, Neurons in primary motor cortex engaged during action
573 observation, Eur. J. Neurosci. 31 (2010) 386–398. doi:10.1111/j.1460-9568.2009.07067.x.

574 [19] R. Wahnoun, J. He, S.I.H. Tillery, Selection and parameterization of cortical neurons for
575 neuroprosthetic control, J. Neural Eng. 3 (2006) 162–171. doi:10.1088/1741-2560/3/2/010.

576 [20] L.E. Brown, E.T. Wilson, P.L. Gribble, Repetitive transcranial magnetic stimulation to the
577 primary motor cortex interferes with motor learning by observing., J. Cogn. Neurosci. 21
578 (2009) 1013–22. doi:10.1162/jocn.2009.21079.

579 [21] N. Aridan, R. Mukamel, Activity in primary motor cortex during action observation covaries
580 with subsequent behavioral changes in execution, Brain Behav. 6 (2016) e00550.

581 doi:10.1002/brb3.550.

582 [22] T. Wiestler, J. Diedrichsen, Skill learning strengthens cortical representations of motor
583 sequences., Elife. 2 (2013) e00801. doi:10.7554/eLife.00801.

584 [23] D.H. Brainard, The Psychophysics Toolbox., Spat. Vis. 10 (1997) 433–6.

585 [24] A. Alonzo, J. Brassil, J.L. Taylor, D. Martin, C.K. Loo, Daily transcranial direct current
586 stimulation (tDCS) leads to greater increases in cortical excitability than second daily
587 transcranial direct current stimulation, Brain Stimul. 5 (2012) 208–213.
588 doi:10.1016/j.brs.2011.04.006.

589 [25] J. Reis, H.M. Schambra, L.G. Cohen, E.R. Buch, B. Fritsch, E. Zarahn, P. a Celnik, J.W.
590 Krakauer, Noninvasive cortical stimulation enhances motor skill acquisition over multiple
591 days through an effect on consolidation., Proc. Natl. Acad. Sci. U. S. A. 106 (2009) 1590–
592 1595. doi:10.1073/pnas.0805413106.

593 [26] C.M. Saucedo Marquez, X. Zhang, S.P. Swinnen, R. Meesen, N. Wenderoth, Task-specific
594 effect of transcranial direct current stimulation on motor learning., Front. Hum. Neurosci. 7
595 (2013) 333. doi:10.3389/fnhum.2013.00333.

596 [27] A.F. DaSilva, M.S. Volz, M. Bikson, F. Fregni, Electrode Positioning and Montage in
597 Transcranial Direct Current Stimulation, J. Vis. Exp. (2011). doi:10.3791/2744.

598 [28] A.J. Woods, A. Antal, M. Bikson, P.S. Boggio, A.R. Brunoni, P. Celnik, L.G. Cohen, F.
599 Fregni, C.S. Herrmann, E.S. Kappenman, H. Knotkova, D. Liebetanz, C. Miniussi, P.C.
600 Miranda, W. Paulus, A. Priori, D. Reato, C. Stagg, N. Wenderoth, M.A. Nitsche, A technical
601 guide to tDCS, and related non-invasive brain stimulation tools, Clin. Neurophysiol. (2015).
602 doi:10.1016/j.clinph.2015.11.012.

603 [29] A. Fertonani, C. Ferrari, C. Miniussi, What do you feel if I apply transcranial electric
604 stimulation? Safety, sensations and secondary induced effects, Clin. Neurophysiol. 126

(2015) 2181–2188. doi:10.1016/j.clinph.2015.03.015.

[30] S. Waters-Metenier, M. Husain, T. Wiestler, J. Diedrichsen, Bihemispheric transcranial direct current stimulation enhances effector-independent representations of motor synergy and sequence learning., J. Neurosci. 34 (2014) 1037–50. doi:10.1523/JNEUROSCI.2282-13.2014.

[31] R. Morey, J. Rouder, J. Love, B. Marwick, BayesFactor: 0.9.12-2 CRAN, (2015). doi:10.5281/zenodo.31202.

[32] H. Jeffreys, Theory of probability, 3rd ed., Oxford: Clarendon Press, 1961.

[33] G. Cumming, Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis., Routledge/Taylor & Francis Group, New York, NY, US, 2012.

[34] G. Bird, C. Heyes, Effector-dependent learning by observation of a finger movement sequence., J. Exp. Psychol. Hum. Percept. Perform. 31 (2005) 262–75. doi:10.1037/0096-1523.31.2.262.

[35] C.M. Heyes, C.L. Foster, Motor learning by observation: evidence from a serial reaction time task., Q. J. Exp. Psychol. A. 55 (2002) 593–607. doi:10.1080/02724980143000389.

[36] M. Osman, G. Bird, C. Heyes, Action observation supports effector-dependent learning of finger movement sequences., Exp. Brain Res. 165 (2005) 19–27. doi:10.1007/s00221-005-2275-0.

[37] A. Lago-Rodríguez, B. Cheeran, The role of mirror neurons in observational motor learning: an integrative review, Eur. J. Hum. Mov. 32 (2014) 82–103. http://www.eurjhm.com/index.php/eurjhm/article/view/320.

[38] L.P. Kirsch, E.S. Cross, Additive routes to action learning: Layering experience shapes engagement of the action observation network., Cereb. Cortex. 25 (2015) 4799–811. doi:10.1093/cercor/bhv167.

629 [39] T. Gardner, A. Aglinskas, E.S. Cross, Using guitar learning to probe the Action Observation
630 Network's response to visuomotor familiarity, Neuroimage. 156 (2017) 174–189.
631 doi:10.1016/j.neuroimage.2017.04.060.

632 [40] S.B. Lim, B.C. Larssen, N.J. Hodges, Manipulating visual-motor experience to probe for
633 observation-induced after-effects in adaptation learning., Exp. Brain Res. 232 (2014) 789–
634 802. doi:10.1007/s00221-013-3788-6.

635 [41] G. Vannuscorps, A. Caramazza, Typical action perception and interpretation without motor
636 simulation, Proc. Natl. Acad. Sci. 113 (2016) 86–91. doi:10.1073/pnas.1516978112.

637 [42] D. Maslovat, N.J. Hodges, O.E. Krigolson, T.C. Handy, Observational practice benefits are
638 limited to perceptual improvements in the acquisition of a novel coordination skill, Exp.
639 Brain Res. 204 (2010) 119–130. doi:10.1007/s00221-010-2302-7.

640 [43] S. Caspers, K. Zilles, A.R. Laird, S.B. Eickhoff, ALE meta-analysis of action observation
641 and imitation in the human brain., Neuroimage. 50 (2010) 1148–67.
642 doi:10.1016/j.neuroimage.2009.12.112.

643 [44] T.T. de Beukelaar, K. Alaerts, S.P. Swinnen, N. Wenderoth, Motor facilitation during action
644 observation: The role of M1 and PMv in grasp predictions, Cortex. 75 (2016) 180–192.
645 doi:10.1016/j.cortex.2015.11.009.

646 [45] L. Fadiga, L. Fogassi, G. Pavesi, G. Rizzolatti, Motor facilitation during action observation: a
647 magnetic stimulation study., J. Neurophysiol. 73 (1995) 2608–11.
648 http://www.ncbi.nlm.nih.gov/pubmed/7666169.

649 [46] K. Monte-Silva, M.F. Kuo, S. Hessenthaler, S. Fresnoza, D. Liebetanz, W. Paulus, M.A.
650 Nitsche, Induction of late LTP-like plasticity in the human motor cortex by repeated non-
651 invasive brain stimulation, Brain Stimul. 6 (2013) 424–432. doi:10.1016/j.brs.2012.04.011.

652 [47] J. Nilsson, A. V. Lebedev, A. Rydström, M. Lövdén, Direct-Current Stimulation Does Little

653 to Improve the Outcome of Working Memory Training in Older Adults, Psychol. Sci. 28
654 (2017) 907–920. doi:10.1177/0956797617698139.

655 [48] J. Medina, S. Cason, No evidential value in samples of transcranial direct current stimulation
656 (tDCS) studies of cognition and working memory in healthy populations, Cortex. 94 (2017)
657 131–141. doi:10.1016/j.cortex.2017.06.021.

658 [49] A.R. Price, H. McAdams, M. Grossman, R.H. Hamilton, A Meta-analysis of Transcranial
659 Direct Current Stimulation Studies Examining the Reliability of Effects on Language
660 Measures, Brain Stimul. 8 (2015) 1093–1100. doi:10.1016/j.brs.2015.06.013.

661 [50] S.J. Westwood, A. Olson, R.C. Miall, R. Nappo, C. Romani, Limits to tDCS effects in
662 language: Failures to modulate word production in healthy participants with frontal or
663 temporal tDCS, Cortex. 86 (2017) 64–82. doi:10.1016/j.cortex.2016.10.016.

664 [51] U. Simonsohn, L.D. Nelson, J.P. Simmons, P-curve: A key to the file-drawer., J. Exp.
665 Psychol. Gen. 143 (2014) 534–547. doi:10.1037/a0033242.

666 [52] M.E. Héroux, C.K. Loo, J.L. Taylor, S.C. Gandevia, Questionable science and
667 reproducibility in electrical brain stimulation research, PLoS One. 12 (2017) e0175635.
668 doi:10.1371/journal.pone.0175635.

669 [53] S. Wade, G. Hammond, Anodal transcranial direct current stimulation over premotor cortex
670 facilitates observational learning of a motor sequence, Eur. J. Neurosci. 41 (2015) 1597–
671 1602. doi:10.1111/ejn.12916.

672 [54] S. Koyama, S. Tanaka, S. Tanabe, N. Sadato, Dual-hemisphere transcranial direct current
673 stimulation over primary motor cortex enhances consolidation of a ballistic thumb
674 movement, Neurosci. Lett. 588 (2015) 49–53. doi:10.1016/j.neulet.2014.11.043.

675 [55] S. Karok, A.G. Witney, Enhanced motor learning following task-concurrent dual transcranial
676 direct current stimulation, PLoS One. 8 (2013) 1–10. doi:10.1371/journal.pone.0085693.

677 [56] S. Karok, D. Fletcher, A.G. Witney, Task-specificity of unilateral anodal and dual-M1 tDCS
678 effects on motor learning, Neuropsychologia. 94 (2017) 84–95.
679 doi:10.1016/j.neuropsychologia.2016.12.002.

680 [57] B.W. Vines, C. Cerruti, G. Schlaug, Dual-hemisphere tDCS facilitates greater improvements
681 for healthy subjects' non-dominant hand compared to uni-hemisphere stimulation., BMC
682 Neurosci. 9 (2008) 103. doi:10.1186/1471-2202-9-103.

683 [58] L. Mordillo-Mateos, L. Turpin-Fenoll, J. Millán-Pascual, N. Núñez-Pérez, I. Panyavin, J.M.
684 Gómez-Argüelles, E. Botia-Paniagua, G. Foffani, N. Lang, A. Oliviero, Effects of
685 simultaneous bilateral tDCS of the human motor cortex, Brain Stimul. 5 (2012) 214–222.
686 doi:10.1016/j.brs.2011.05.001.

687 [59] K. Cuypers, D.J.F. Leenus, F.E. van den Berg, M.A. Nitsche, H. Thijs, N. Wenderoth, R.L.J.
688 Meesen, Is motor learning mediated by tDCS intensity?, PLoS One. 8 (2013) e67344.
689 doi:10.1371/journal.pone.0067344.

690 [60] M. Kuo, R. Polanía, M. Nitsche, Physiology of Transcranial Direct and Alternating Current
691 Stimulation, in: A. Brunoni, M. Nitsche, C. Loo (Eds.), Transcranial Direct Curr. Stimul.
692 Neuropsychiatr. Disord., Springer International Publishing, Cham, 2016: pp. 29–46.
693 doi:10.1007/978-3-319-33967-2_3.

694

695

Supplementary Materials 1: Results without the nine non-TMS localised participants

Accuracy during training sessions

During the observational practice sessions, attention to the task was assessed by accurate responses to the error question (spotting incorrectly executed sequences). The overall accuracy was 82%, significantly ($p < 0.001$) higher than a 50% chance level (yes/no answers), confirming that participants paid attention to the task. The average accuracies for each group and day are plotted in Supplementary Figure 1D. On average, across the four training days, the sham group performed better (M = 85% [79%, 91%]) than the active group (M = 81% [77%, 85%]), with no significant difference between the two groups ($t_{26.44} = 1.35$, $p = 0.189$). There was a negative correlation between the average accuracy and the average self-report on how much performance was affected by the discomforting sensations (Kendall's tau-b = -0.256, $p = 0.042$; across both groups).

Observational training effects on sequence-specific learning

The effect of observational training on sequence-specific learning was assessed as a post-training (separately for the post-test and retention-test) difference between the trained and untrained sequence initiation time, execution time, and error rate. For the sequence initiation time and execution time, we measured a percentage difference ([(untrained/trained)-1]*100), but for the error rate (to avoid dividing by zero), we calculated an absolute difference (untrained-trained) between the trained and untrained sequences (results of these measures are plotted in Supplementary **Figure** 1A-C). To correct for possible pre-training differences, we performed a linear regression between the pre-training difference (predictor) and the post-training difference (outcome; see Supplementary **Figure** 1E for an example plot). The intercept of the regression line was used as a measure of the post-training difference between trained and untrained sequences, controlling for possible pre-training differences. This method reduces the noise of unwanted differences in the difficulty of trained and untrained sequences and thus allows a more accurate measurement of the training effect.

Both groups showed significant observational training effects at both post-test and retention-test on all three performance measures, with medium to large effect sizes ($d_z = 0.52 – 1.02$). Except, the active stimulation group demonstrated no effect on error rates at retention-test. Detailed results are provided in Supplementary Table 1.

tDCS effects on sequence-specific learning by observation

Primary analysis

The effect of stimulation on sequence-specific learning was assessed by comparing observational training effects (the post-training ~ pre-training regression line intercepts) between the sham and active stimulation groups. The performed analysis of covariance (ANCOVA) did not reveal any significant difference between the two groups on any of the three measures either at post-test or retention-test. (Supplementary **Figure** 1E plots post-test initiation time results). The Bayes factor analysis returned anecdotal to substantial evidence against the stimulation effect. Detailed results are provided in Supplementary Table 1.

Secondary analysis: accounting for error detection accuracy

Due to concern that the stimulation effect could be confounded by sensation and error detection differences (both of which were negatively correlated) between the sham and active stimulation groups, we added the mean error detection accuracy as a covariate to the previous ANCOVA model and repeated the group comparison analysis.

The corrected analysis revealed evidence for the stimulation effect on the percentage difference between trained and untrained sequence initiation times at post-test. Compared to the sham group, the active stimulation group showed a greater difference on this measure (see Supplementary **Figure** 1F). The error detection accuracy significantly predicted the outcome ($\beta = 0.554$, $p < 0.001$; the better the accuracy during training, the faster initiation time of trained relative to untrained sequences at post-test). All other measures showed anecdotal to substantial evidence against the

744 stimulation effect when accounting for the error detection accuracy. Detailed results are provided in
745 Supplementary Table 1.

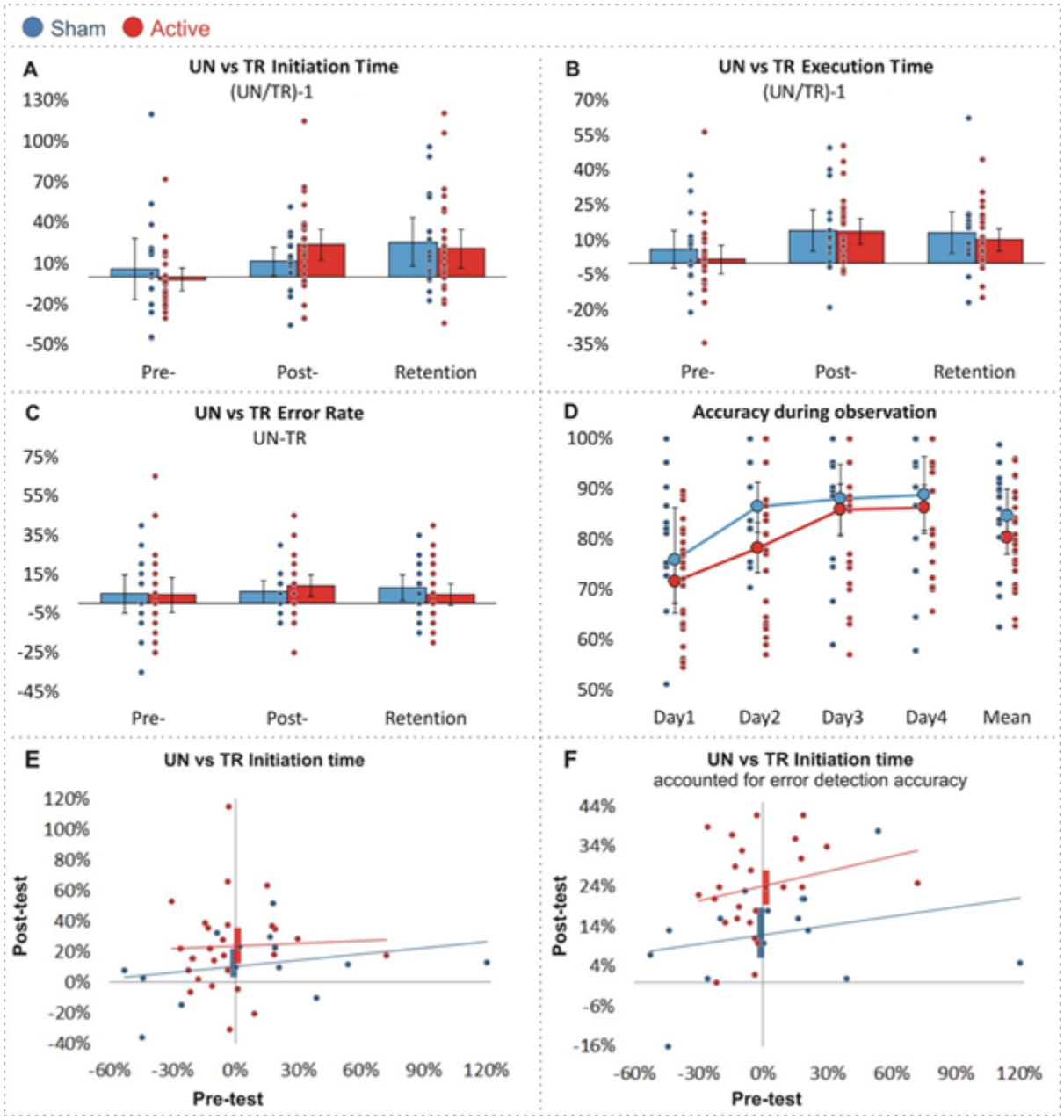
746 **Supplementary Table 1. Observational practice effects and tDCS effects on sequence-specific learning with nine**
747 **non-TMS localised participants excluded.**

		Observational training effect (trained vs. untrained performance)		tDCS effect (group difference)	tDCS effect, accounted for the accuracy during training sessions
		Sham (N = 15)	Active (N = 26)		
Initiation time	Post	$t_{(13)} = 1.95$, $p = 0.073$, $B_0 = 11\%$, $d_z = 0.50$.	$t_{(24)} = 4.02$, $p < 0.001$, $B_0 = 24\%$, $d_z = 0.79$.	$t_{(38)} = 1.50$, $p = 0.141$, $d = 0.49$, anecdotal evidence against the effect ($BF_{10}/BF_{01} = 0.76/1.31$).	$t_{(37)} = 2.69$, $p = 0.011$, $d = 0.89$, substantial evidence for the effect ($BF_{10}/BF_{01} = 3.44/0.29$).
	Ret.	$t_{(13)} = 2.67$, $p = 0.019$, $B_0 = 25\%$, $d_z = 0.69$.	$t_{(24)} = 2.87$, $p = 0.008$, $B_0 = 21\%$, $d_z = 0.56$.	$t_{(38)} = -0.35$, $p = 0.729$, substantial evidence against the effect ($BF_{10}/BF_{01} = 0.33/3.00$).	$t_{(37)} = -0.29$, $p = 0.773$, anecdotal evidence against the effect ($BF_{10}/BF_{01} = 0.34/2.97$).
Execution time	Post	$t_{(13)} = 2.42$, $p = 0.031$, $B_0 = 10\%$, $d_z = 0.62$.	$t_{(24)} = 4.75$, $p < 0.001$, $B_0 = 14\%$, $d_z = 0.93$.	$t_{(38)} = 0.16$, $p = 0.876$, substantial evidence against the effect ($BF_{10}/BF_{01} = 0.31/3.18$).	$t_{(37)} = -0.07$, $p = 0.943$, substantial evidence against the effect ($BF_{10}/BF_{01} = 0.32/3.16$).
	Ret.	$t_{(13)} = 2.40$, $p = 0.032$, $B_0 = 9\%$, $d_z = 0.62$.	$t_{(24)} = 3.99$, $p = 0.001$, $B_0 = 10\%$, $d_z = 0.78$.	$t_{(38)} = -0.47$, $p = 0.64$, anecdotal evidence against the effect ($BF_{10}/BF_{01} = 0.35/2.84$).	$t_{(37)} = -0.42$, $p = 0.678$, anecdotal evidence against the effect ($BF_{10}/BF_{01} = 0.35/2.82$).
Error rate	Post	$t_{(13)} = 1.90$, $p = 0.079$, $B_0 = 6\%$, $d_z = 0.49$.	$t_{(24)} = 2.89$, $p = 0.008$, $B_0 = 9\%$, $d_z = 0.57$.	$t_{(38)} = 0.69$, $p = 0.497$, anecdotal evidence against the effect ($BF_{10}/BF_{01} = 0.38/2.63$).	$t_{(37)} = 0.43$, $p = 0.667$, anecdotal evidence against the effect ($BF_{10}/BF_{01} = 0.36/2.82$).
	Ret.	$t_{(13)} = 2.13$, $p = 0.053$, $B_0 = 8\%$, $d_z = 0.55$.	$t_{(24)} = 1.45$, $p = 0.161$, $B_0 = 4\%$, $d_z = 0.28$.	$t_{(38)} = -0.72$, $p = 0.476$, anecdotal evidence against the effect ($BF_{10}/BF_{01} = 0.38/2.61$).	$t_{(37)} = -1.00$, $p = 0.322$, anecdotal evidence against the effect ($BF_{10}/BF_{01} = 0.46/2.20$).

Shaded fields highlight non-significant effects.

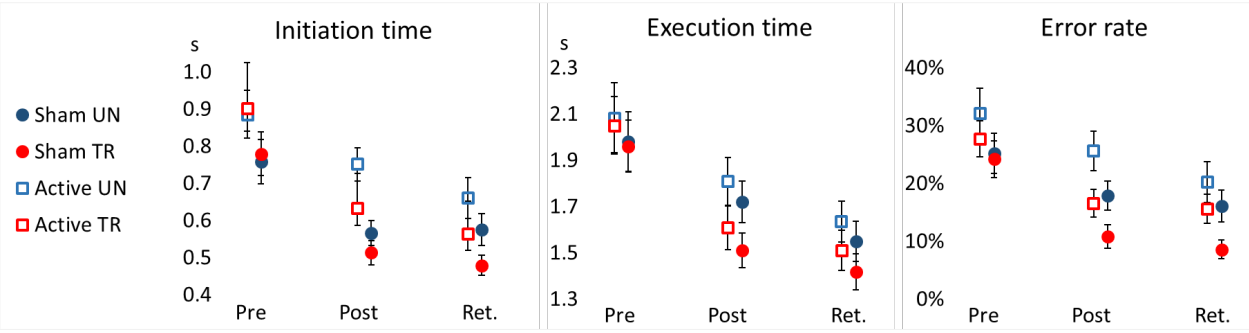
748

749



Supplementary Figure 1. Performance results. Pre-, post-, and retention-test difference in initiation time (**A**), execution time (B), and error rate (**C**) between trained (TR) and untrained (UN) sequences for sham (blue) and active (red) stimulation groups. **D**. Error detection accuracy during observational practice sessions. **A-D**. Bars and large dots: group averages; small dots: individual participant values; error bars: 95% CI. **E**. Regression lines of pre-test (predictor) and the post-test difference between trained and untrained sequence initiation times for sham (blue) and active (red) stimulation groups. Intercepts of the regression lines represent the predicted post-test difference if the pre-test difference is zero. Vertical bars represent 96% CIs of intercepts **F**. Same as E, but post-test difference corrected for error detection accuracy during training sessions.

Supplementary Materials 2: Raw performance measures



The figure above illustrates the mean and standard deviation values for participants' initiation time, execution time, and error rate (three different panels), split into Pre, Post and Retention tests (along x-axes). These data are further split into stimulation groups (active vs. sham; filled circles vs. open squares) and the training status of the sequences (blue for untrained and red for trained).

Supplementary Materials 3: Group differences in performance generalisation to the untrained sequences

The table below shows independent sample t-test significance comparing sham and active stimulation group differences in performance generalisation to the untrained sequences.

	p-value
Pre vs . Post difference of untrained sequence	
Initiation time	0.107
Execution time	0.847
Error rate	0.901
Pre vs. Retention difference of untrained sequence	
Initiation time	0.515
Execution time	0.936
Error rate	0.674
Accounted for the accuracy during training sessions	
Pre vs. Post difference of untrained sequence	
Initiation time	0.045
Execution time	0.784
Error rate	0.982
Pre vs. Retentions difference of untrained sequence	
Initiation time	0.661
Execution time	0.909
Error rate	0.596

Supplementary Materials 4: tDCS effects on sequence-specific learning using ANCOVA

Dependent variable: Post-training (separate for post-test and retention-test) trained/untrained sequence performance difference, UN/TR-1 for the IT and ET and UN-TR for Err.
Covariate: Pre-training trained/untrained sequence performance difference, UN/TR-1 for the IT and ET and UN-TR for Err.
Within-subject factor: stimulation (active/sham).

Significance reported one-tailed

Post-test

Initiation time

Tests of Between-Subjects Effects

Dependent Variable: ITPostDiff

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	.186[a]	2	.093	1.293	.142
Intercept	1.669	1	1.669	23.250	.000
ITPreDiff	.037	1	.037	.510	.240
stimulation	**.161**	**1**	**.161**	**2.238**	**.072**
Error	3.374	47	.072		
Total	5.286	50			
Corrected Total	3.560	49			

a. R Squared = .052 (Adjusted R Squared = .012)

Execution time

Tests of Between-Subjects Effects

Dependent Variable: ETPostDiff

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	.043[a]	2	.021	1.009	.187
Intercept	1.002	1	1.002	47.007	.000
ETPreDiff	.041	1	.041	1.915	.087
stimulation	**.003**	**1**	**.003**	**.139**	**.355**
Error	1.001	47	.021		
Total	2.077	50			
Corrected Total	1.044	49			

a. R Squared = .041 (Adjusted R Squared = .000)

Error rate

Tests of Between-Subjects Effects

Dependent Variable: ErrPostDiff

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	.005[a]	2	.003	.130	.439
Intercept	.315	1	.315	15.292	.000
ErrPreDiff	.001	1	.001	.029	.433
stimulation	**.004**	**1**	**.004**	**.217**	**.322**
Error	.969	47	.021		
Total	1.303	50			
Corrected Total	.974	49			

a. R Squared = .006 (Adjusted R Squared = -.037)

Retention-test

Initiation time

Tests of Between-Subjects Effects

Dependent Variable: ITRetDiff

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	.046[a]	2	.023	.198	.411
Intercept	2.197	1	2.197	18.750	.000
ITPreDiff	.046	1	.046	.395	.267
stimulation	**.000**	**1**	**.000**	**.002**	**.480**
Error	5.506	47	.117		
Total	7.771	50			
Corrected Total	5.553	49			

a. R Squared = .008 (Adjusted R Squared = -.034)

Execution time

Tests of Between-Subjects Effects

Dependent Variable: ETRetDiff

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	.062[a]	2	.031	1.755	.092
Intercept	.476	1	.476	27.028	.000
ETPreDiff	.062	1	.062	3.511	.034
stimulation	**6.958E-5**	**1**	**6.958E-5**	**.004**	**.475**
Error	.827	47	.018		
Total	1.392	50			
Corrected Total	.889	49			

a. R Squared = .070 (Adjusted R Squared = .030)

Error rate

Tests of Between-Subjects Effects

Dependent Variable: ErrRetDiff

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	.023[a]	2	.011	.615	.273
Intercept	.168	1	.168	9.105	.002
ErrPreDiff	.012	1	.012	.668	.209
stimulation	**.012**	**1**	**.012**	**.662**	**.210**
Error	.867	47	.018		
Total	1.070	50			
Corrected Total	.890	49			

a. R Squared = .026 (Adjusted R Squared = -.016)