The role of dorsolateral prefrontal cortex in inhibition mechanism: A study on cognitive reflection test and similar tasks through neuromodulation

Viola Oldrati a,*, Jessica Patricelli b, Barbara Colombo b, c, Alessandro Antonietti b

a Department of Brain and Behavioral Sciences, University of Pavia, Piazza A. Botta 6, Pavia, Italy
b Department of Psychology, Catholic University of the Sacred Heart, Largo A. Gemelli 1, Milan, Italy
c Division of Education and Human Studies, Champlain College, Burlington, VT, USA

A R T I C L E   I N F O

Article history:
Received 19 January 2016
Received in revised form 9 September 2016
Accepted 11 September 2016
Available online 17 September 2016

Keywords:
Problem solving
Cognitive Reflection Test
Mathematics
Dorsolateral prefrontal cortex (DLPFC)
Impulsivity
Neuromodulation
TDCS

A B S T R A C T

The main characteristic of the cognitive reflection test (CRT) is that it requires people to overcome a cognitive conflict. Solving this conflict requires (1) inhibitory control of prepotent but incorrect responses and (2) mental set-shifting in order to reframe the problem and reach a meaningful solution. Based on the well-known involvement of the dorsolateral prefrontal cortex (DLPFC) in inhibitory control we hypothesized that transcranial direct current stimulation (tDCS) of the DLPFC would modulate its contribution to problem-solving performance. Thirty-nine participants undergoing anodal, cathodal, or sham tDCS were asked to solve the CRT and similar mathematical problems that were structured to induce an automatic, impulsive but incorrect response. To provide a multi-dimensional picture of the processes underlying responding we assessed impulsivity traits using self-report measures and recorded physiological indices using biofeedback equipment. The results indicated that participants were more likely to provide incorrect impulsive responses after cathodal stimulation, i.e. when inhibitory control associated to the DLPFC was reduced. Baseline values of blood volume pulses predicted solution recognition, highlighting the potential role of individual physiological differences in problem solving. In conclusion, this study provides evidence supporting the role of the DLPFC in modulation of processes involved in solving CRTs and similar problems, thanks to its association to the inhibitory control mechanisms involved in suppressing impulsive responses.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

The Cognitive Reflection Test (CRT) (Frederick, 2005) is attracting growing interest from researchers investigating the cognitive mechanisms involved in decision making and problem solving (Albaity et al., 2014; Juanchich et al., 2015; Sinayev and Peters, 2015; Thoma et al., 2015; Toplak et al., 2014). The CRT presents three mathematical problems that are designed to induce an immediate, incorrect response. Consider the first item: "A bat and a ball cost €10.10 in total. The bat costs €10 more than the ball. How much does the ball cost?"1; The intuitive, automatic, yet wrong, response is 10 cents. Deeper reflection on the elements of the problem helps one to realise that the difference between €10.00 and 10 cents is €9.90 cents and not €10.00, making easier to reach the correct solution (5 cents). It is interesting to note that even individuals who provide the correct answer initially formulate the incorrect answer, before discarding it. This suggests that solving CRT problems requires the activation of a mechanism which inhibits impulsive responses (Frederick, 2005).

The CRT was first used in economics (Frederick, 2005) but was soon used to highlight the association between the ability to resist to the tendency to give an impulsive response and several reasoning biases (e.g., Baldi et al., 2013; Campitelli and Labollita, 2010; Hoppe and Kusterer, 2011; Noori, 2016; Oechssler et al., 2009). These biases appear to be elicited because the immediate impression created by the data provided suggests an incorrect or incoherent response. Furthermore, when used in conjunction with the Ultimatum Game, a task presenting participants with different ways of splitting a sum of money, the CRT was shown to be negatively correlated with the rejection of unfair offers (Calvillo and Burgeno, 2015). Rejection of unfair offers is usually attributed to failure to suppress the negative emotional reaction produced by the unequal split (Sanfey et al., 2003). High rates of solution of the CRT are associated to the suppression of the emotional influence

---

1 The original CRT item was as follows: "A bat and a ball cost $1.10. The bat costs $1.00 more than the ball. How much does the ball cost?" For this study the currency was changed to euros to make the problem more suited to the Italian context.

* Corresponding author.
E-mail address: viola.oldrati01@universitadipavia.it (V. Oldrati).

http://dx.doi.org/10.1016/j.neuropsychologia.2016.09.010
0028-3932/© 2016 Elsevier Ltd. All rights reserved.
that hinders rational evaluation of the cost-benefit balance in moral dilemmas (Baron et al., 2015). Campitelli and Cerrans (2014) conducted a fine-grained analysis of patterns of CRT responses in a large sample of participants and were able to identify a specific characteristic underlying successful performance, namely the ability to inhibit an impulsive response. Moreover, observations of immediate responses elicited by the CRT problems (recorded using a mouse-tracking system) supported the notion that in the first instance individuals are always attracted by the wrong response. It follows that to reach the correct answer, they have to inhibit the automatic processing such problems elicit (Travers et al., 2016).

To our knowledge there has so far been no investigation into the neurobiological correlates of CRT performance, but investigations of the cerebral correlates of insight problems may constitute a useful reference point (e.g., Aziz-Zadeh et al., 2009; Bowden and Jung-Beeman, 2003; Dandan et al., 2013; Dietrich and Kanso, 2010; Jung-Beeman et al., 2004; Kounios et al., 2006; Luo and Knoblich, 2007; Luo et al., 2004). In fact the CRT shares some of the features that are usually attributed to insight problems (Macchi and Bagassi, 2012). Both require suppression of the spontaneous tendency to follow a faulty line of reasoning and a consequent switch to a different mental set (Gilhooly and Fioratou, 2009; Kounios and Beeman, 2014). Electroencephalography (EEG) and functional magnetic resonance imaging studies have shown activation of the right anterior superior temporal gyrus (Jung-Beeman et al., 2004; Kounios and Beeman, 2009) and anterior cingulate cortex (Subramaniam and Kounios, 2009) during insight problem solving. Another brain region that may be critically involved in solving insight problems is the prefrontal cortex (Antoniotti and Balconi, 2010) which is linked to cognitive control and executive functions. More specifically, it has been suggested that lateral prefrontal cortex exerts influence over various brain regions involved in information processing when it receives inputs from anterior cingulate cortex that signal the presence of a cognitive conflict (Miller and Cohen, 2001). Consistent with this interpretation, lateral prefrontal cortex activation during insight problem solving has been associated with the processes involved in regulation of attentional set, which are required to overcome cognitive conflicts between prepotent responses and resolution of the problem (Luo et al., 2004).

Neuropsychological research has clarified the specific role of the dorsolateral prefrontal cortex (DLPFC) in cognitive processes. It appears to be associated to controlling executive functions (Barbey et al., 2013), mental set shifting, i.e. shifting between different attributes of elements or rules (Wager et al., 2004), monitoring of ongoing operations, inhibition of prepotent responses (Miyake et al., 2000), and maintaining task-relevant information in working memory when faced with distractors (Colombo et al., 2015; Kane and Engle, 2002).

Clinical studies of the relationship between the prefrontal cortex and inhibitory control, based on the use of delayed response tasks in samples of patients with DLPFC damage, have shown that DLPFC is involved in inhibitory stimulus control (Floden and Stuss, 2006; Shimamura et al., 1995). Damage to this area results in an inability to inhibit internal representations of inappropriate responses or responses that are no longer appropriate and a corresponding perseverative tendency, i.e. a tendency to maintain a pre-activated mental set; this impairment contributes to poor performance on tasks involving executive functions (Aron et al., 2004; Vendrell et al., 1995).

The specific role of the DLPFC in insight problem solving was also highlighted by a brain imaging study which found activation of left DLPFC during the presentation of insight problems (Qiu et al., 2010). This activation was associated with elimination of inappropriate cognitive constraints and the subsequent overcoming of mental impasses. The left DLPFC was also found to be preferentially activated by creative tasks rather than control non-creative tasks (Aziz-Zadeh et al., 2013; Huang et al., 2013), probably due to its involvement in top-down organisation of the creative process, which involves a variety of functions mediated by a distributed network in which the left DLPFC seems to play a key role.

The neuroimaging results are corroborated by research showing that stimulation of DLPFC enhances recognition of correct solutions to a verbal insight problem solving task (Metuki et al., 2012). Evidence of a dissociation between the effects of DLPFC stimulation on performance of verbal insight problems and simple verbal fluency tasks would confirm that the involvement of the DLPFC in complex problem solving is due to the higher demand such tasks place on executive functions (Cerruti and Schlaug, 2009) and suggest that insight problem solving is primarily mediated by cognitive control processes rather than by semantic processes. Stimulation of the left DLPFC has been shown to influence the focusing and defocusing of attention, as well as shifting the focus of attention (Colombo et al., 2015). Other studies have also provided evidence that transcranial direct current stimulation (tDCS) over the left DLPFC modulates performance of reasoning tasks by decreasing response times in a probabilistic guessing task (Hecht et al., 2010) and increasing accuracy in an implicit learning task (Kincses et al., 2004).

The aim of this study was to investigate the contribution of DLPFC to performance on the CRT and similar problems by using tDCS to modulate brain activation. This technique creates a continuous, low intensity electric current on the scalp and is used to increase or decrease cortical excitability by depolarising or hyperpolarising, respectively, cortical neurons at a sub-threshold level (Jauk et al., 2013). Modulation of membrane potentials using tDCS has been linked to both cognitive facilitation and cognitive inhibition (Jacobson et al., 2012). In general, anodal stimulation, which increases the spontaneous firing frequency of cortical neurons, enhances cognitive performance (Fregni et al., 2005; Javadi et al., 2012; Metuki et al., 2012; Straube et al., 2011; Wirth et al., 2011), whereas inhibition or impairment of cognitive processes, due to decreasing spontaneous cells firing, is observed after cathodal stimulation (Bohringer et al., 2013; Pope and Miell, 2012; Straube et al., 2011). Brain stimulation techniques have the important advantage of allowing researchers to test causal relationships between neural structures and cognitive performance rather than just correlations. Non-invasive brain stimulation techniques, such as tDCS and transcranial magnetic stimulation (TMS), have been used in research on decision making to modulate cognitive processes underpinning choices base on intuition, by temporarily interfering with the functioning of cortical areas implicated in rational reasoning and deliberation (Iannello et al., 2014).

Given that CRT-like problems require the person facing them to switch away from his or her pre-existing mental set in order to reframe the problem and reach the correct solution, and given the well-known involvement of the DLPFC in control and executive functions (discussed above), one might hypothesise that the DLPFC plays a key role in enabling people to solve CRT problems. Inhibitory processes are crucial to solving the CRT and analogous problems, since, as we have seen, they enable individuals to overcome the immediate, but inappropriate, representation of the problem.

This paper reports an experiment in which the CRT and similar numerical problems were used to determine the effects of neuromodulation on mathematical problem solving. Based on the assumption that the DLPFC is involved in cognitive control processes and in specific aspects of task performance that require inhibition and mindset switching, we hypothesised that anodal stimulation over the target area would improve performance by
decreasing impulsive responding, whereas cathodal stimulation would have the opposite effect.

We assessed participants’ impulsivity in attention, motor, and non-planning domains (Patton et al., 1995), as well as functional and dysfunctional impulsivity traits (Claes et al., 2000) in order to control for the possible moderating effect of these variables on task responses. Impulsivity is often associated with ‘cognitive control’, an umbrella term for a multidimensional, heterogeneous construct. It has been proposed that we should distinguish between cognitive inhibition, defined as an active and voluntary supervisory system which regulates lower-order processes and is able to totally or partially stop an ongoing mental process with or without intention (MacLeod, 2007), and inhibition on the behavioural level, such as response inhibition, resisting temptations, and delaying of gratification (Aron, 2007; Bari and Robbins, 2013; Bjorklund and Harnishfeger, 1995). The assessment of different aspects of impulsivity, which would lead individuals to wrong responses in the CRT and similar problems, may contribute to deepen the role of such personality traits in modulating the lack of inhibition of misleading reasoning tendencies.

Finally, biofeedback indices were recorded to enable us to examine whether different problem solving strategies would be reflected in individuals’ physiological responses. This analysis was exploratory because, to the best of our knowledge, no-one has yet compared physiological indicators in subjects using different strategies to tackle a given problem.

2. Hypotheses

The aim of this study was to achieve a multi-dimensional understanding of the various aspects of problem solving, such as the modulation of the cognitive processes underpinning CRT by DLPFC and cognitive impulsivity traits and the physiological correlates of CRT performance. Three main hypotheses were formulated.

Anodal stimulation of the left DLPFC would improve recognition of correct solutions, whereas cathodal stimulation would increase the number of incorrect, impulsive responses.

Self-report measures of cognitive impulsivity would moderate problem-solving performance. In particular, participants reporting higher impulsivity would show worse problem-solving performance, providing more wrong, impulsive answers.

Physiological parameters would vary according to response type (right vs. wrong impulsive vs. wrong not impulsive) and would reflect problem solving strategy.

3. Methods

3.1. Design

The study adopted a double-blind, between-subjects, single factor (one way) design. The independent, between-subjects variable was stimulation polarity (anodal; cathodal; sham). Scores on the Dickman Impulsivity Inventory (DII) and the Barratt Impulsivity Scale (BIS-11) and various physiological parameters were used as covariates.

3.2. Participants

Thirty-nine healthy participants volunteered to take part in this study (15 men; mean age = 25.28 yrs, SD = 8.04, range 20–52; left-handed n = 6). Gender and handedness were homogenously distributed across conditions (gender: \(\chi^2 (2, N=39) = .65, p = .72\); handedness: \(\chi^2 (2, N=39) = 1.82, p = .55\). All participants were native Italian speakers (one participant was bilingual) and had normal or corrected-to-normal vision.

Prior to the experiment all participant filled in a questionnaire to evaluate their suitability for tDCS. None of the volunteers had a history of neurological disorders or brain trauma, or a family history of epilepsy. Participants provided written informed consent according to a protocol approved by the local ethics committee and compliant with the Code of Ethics of the World Medical Association (Declaration of Helsinki).

3.3. Materials

3.3.1. Cognitive reflection test and similar tasks

The CRT consists of 3 items, each one presenting a mathematical problem that elicits an immediate, but erroneous response. The other 10 problems (see Appendix) used in the study share the main characteristics of the CRT items in that they all induce solvers to provide an impulsive, erroneous response (often based on simple mathematical operations) and the salience of the erroneous response prevents solvers from realising that the immediate representation of the problem has to be restructured if they are to reach the correct answer.

The 13 problems were presented to participants in a fixed order: first the three CRT items, then the 10 similar problems. Four response options were presented: the correct answer, the automatic/impulsive one, and two other incorrect answers which were not impulsive choices. Including non-impulsive incorrect answers (as proposed by Pennycook et al., 2016) allows researchers to determine whether a participant failed to solve the problems because of a specific impulsive tendency to select the apparently satisfactory, but actually incorrect, option. The reasoning underlying the correct and impulsive responses for each problem is given in the Appendix. The order of presentation of the four options was randomly set for each problem.

3.3.2. Assessment of impulsivity – Dickman Impulsivity Inventory and Barratt Impulsivity Scale

The DII and BIS-11 are widely used self-report measures of the personality/behavioural construct of impulsivity. The DII is composed of 23 dichotomous items organised into two subscales: the dysfunctional impulsivity subscale, which evaluates tendency to act with little forethought even when this is unproductive, and the functional impulsivity subscale, which assesses tendency to act with relatively little forethought in contexts where this may be appropriate (Dickman, 1990). Reliability analyses run of the Dutch version of the questionnaire showed that both subscales had adequate internal consistency, and that the correlation between them was low, which suggests that they measure different personality traits (Claes et al., 2000).

The BIS-11 consists of 30 items to which responses are given using a Likert scale. The items are structured into six factors and three subscales: (1) attention impulsivity and cognitive instability (attentional domain), (2) motor impulsivity and perseverance (motor domain) and (3) self-control and cognitive complexity (non-planning domain) (Patton et al., 1995). The Italian translation of the BIS-11 has been shown to be internally consistent (Fossati et al., 2001).

3.3.3. Transcranial direct current stimulation

A constant, direct current stimulation of 1.5 mA was induced by two saline-soaked surface sponge electrodes (25 cm²) and delivered by a battery-driven, constant-current stimulator (HDC Series by Newronika S.r.l, Milan) for 20 min. In the unilateral anodal condition the active anode electrode was positioned over the left DLPFC and the reference cathode electrode was placed over the right deltoid muscle (monoephalic montage). In the cathodal condition the electrode positions were reversed: active cathode...
electrode was placed over the left DLPFC and the reference anode electrode over the right deltoïd muscle. (A picture of the montage is available as Supplementary Material). This specific montage has been shown to provide effective stimulation in similar research contexts (e.g., Filmer et al., 2014; Im et al., 2012; Nasseri et al., 2015).

In the sham condition electrodes were placed as in the unilateral anodal stimulation, but the stimulation was automatically turned off 10 seconds after the start of the session, thus participants felt the characteristic tingling sensations in the vicinity of the electrodes for a brief period of time, which enhanced the plausibility of the sham condition.

The DLPFC was localised using a 10–20 system EEG technique; F3, defined as the intersection between F7, Fz and P7, was identified as the target area (Cerruti and Schlau, 2009). Landmarks on the scalp were taken with a measuring tape and marked using a skin marker.

3.3.4. Procedure

The experiment was carried out in a single, one-hour session at the Laboratory of Cognitive Psychology of the Catholic University of the Sacred Heart in Milan, Italy. Participants were randomly assigned to the conditions: 13 participants underwent anodal stimulation, 13 cathodal stimulation and 13 sham stimulation; in all cases stimulation lasted 20 min (Fig. 1). Stimuli were presented immediately after the stimulation period, on a desktop computer screen using E-prime 2.0 software (Psychology Software Tools, Pittsburgh, PA) in accordance with the protocol applied.

Participants were shown 13 slides, each presenting a different problem. The slides were displayed for a maximum of 45 seconds each and participants were asked to indicate their response choice using the keyboard (letters ‘z’, ‘x’, ‘c’, ‘v’). The four response options for each problem were presented on the same slide as the problem. After submitting their response to a problem participants were presented with the next problem.

Response choice and response time in milliseconds were recorded. We also recorded several psychophysiological measures using biofeedback equipment (Biofeedback 2000<sup>®</sup>-pertz by Schuhfried); skin conductance level (SCL), skin temperature, blood volume pulse (BVP), pulse volume amplitude (PVA), and pulse frequency. A Velcro<sup>®</sup> electrode was attached to the left index finger of the non-dominant hand to record psychophysiological data throughout task. A two-minute baseline was recorded before stimulus presentation began.

After completing the experimental problem solving task participants completed the DII and BIS-11 in paper-and-pencil format.

4. Results

4.1. Assessment of impulsivity

Multivariate analyses of the self-report impulsivity data did not show group differences. A MANOVA on BIS factors indicated that the only factor which differed between groups was cognitive complexity (F<sub>2,36</sub> = 5.46, p < .009, η<sup>2</sup> = .23). Pairwise comparisons indicated that the sham group (mean difference = −1.77, SE = 1.03, p = .09) and cathodal group (mean difference = −3.39, SE = 1.03, p = .002) had higher cognitive complexity scores than the anodal group. Analyses of other BIS factors did not reveal group differences: attention (F<sub>2,36</sub> = 3.0, p = .07, η<sup>2</sup> = .04); motor impulsivity (F<sub>2,36</sub> = .05, p = .89, η<sup>2</sup> = .03); self-control (F<sub>2,36</sub> = .12, p = .89, η<sup>2</sup> = .04); perseverance (F<sub>2,36</sub> = 2.07, p = .23, η<sup>2</sup> = .08); cognitive instability (F<sub>2,36</sub> = 2.08, p = .14, η<sup>2</sup> = .03). An additional MANOVA showed that scores on all BIS subscales were similar in all conditions: attentional impulsivity (F<sub>2,36</sub> = .92, p = .41, η<sup>2</sup> = .05); motor impulsivity (F<sub>2,36</sub> = .47, p = .63, η<sup>2</sup> = .03), and non-planning impulsivity (F<sub>2,36</sub> = 1.84, p = .17, η<sup>2</sup> = .09). Lastly, a multivariate analysis of DII subscale scores revealed that the levels of functional impulsivity (F<sub>2,36</sub> = .45, p = .64, η<sup>2</sup> = .03) and dysfunctional impulsivity (F<sub>2,36</sub> = .46, p = .64, η<sup>2</sup> = .03) were similar in the three conditions.

4.2. Effects of neuromodulation

To assess the effect of tDCS we ran two ANOVAs with stimulation condition (anodal; cathodal; sham) as the independent variable and the total numbers of correct responses and impulsive incorrect responses as dependent variables. Mean scores and SDs are reported in Table 1 of the Supplementary Materials.

A univariate test on the number of correct response revealed a main effect of tDCS (F<sub>2,36</sub> = 7.15, p = .002, η<sup>2</sup> = .28). A post hoc comparison using the Bonferroni correction revealed that there was a difference between the cathodal and anodal conditions (mean difference = −1.92, SE = .51, p = .002) but not between the cathodal and sham conditions (mean difference = −1.15, SE = .51, p = .09; without correction: p < .03) or the anodal and sham condition (mean difference = .77, SE = .51, p = .43; without correction: p = .14).

ANOVA with the total number of impulsive incorrect responses as the dependent variable revealed a main effect of neuromodulation (F<sub>2,36</sub> = 7.97, p = .001, η<sup>2</sup> = .31). A post hoc comparison using the Bonferroni correction revealed differences between the cathodal and sham conditions (mean difference = −2.54, SE = .75, p = .005) and between the cathodal and anodal conditions (mean difference = 2.62, SE = .75, p = .004), but not between the sham and

![Fig. 1. Phases of testing procedure.](Image 113x61 to 473x215)
stimulation conditions, both for correct responses (F_{2,32} = .30, p = .76) and impulsive incorrect responses. Responses times were similar in all three conditions. However, anodal stimulation did not improve recognition of correct solutions (F_{2,32} = .14, p = .89), whereas cathodal stimulation provided fewest correct responses and that enhanced participants’ tendency to respond impulsively.

It is worth mentioning that similar trends emerged when only CRT scores were considered (F_{2,36} = .94, p = .40; CRT correct incorrect responses: F_{2,36} = 1.32, p = .28; CRT incorrect incorrect responses: F_{2,36} = .07) (Fig. 3; mean scores and SDs are reported in Table 2 of the Supplementary Materials). This result is interesting given that most participants perceived the CRT problems as more challenging. The greater perceived difficulty of the CRT problems is consistent with analyses indicating that the proportion of correct responses was lower for CRT problems than for non-CRT problems (Paired-sample t-tests, CRT-other correct: t_{38} = −7.234, p < .001; CRT-other incorrect: t_{38} = −11.734, p < .001). Yet, the trend after the stimulation was the same.

As a last step, ANOVA was used to test whether participants in the cathodal condition responded faster, which would provide corroborative evidence of an enhanced tendency to provide impulsive incorrect responses. Responses times were similar in all stimulation conditions, both for correct responses (F_{2,33} = 1.24, p = .30, η^2 = .07) and impulsive incorrect responses (F_{2,33} = .04, p = .97, η^2 = .002).

### 4.3. Effects of impulsivity

We were also interested in exploring the possible role of individual levels of impulsivity in moderating the effects of neuro-modulation. We ran two GLM ANCOVAs using the same dependent and independent variables as in the previous model, but with impulsivity variables (as DII and BIS scores) as covariates.

Impulsivity as measured by the DII did not have an effect on number of correct responses (functional impulsivity: F_{1,34} = .22, p = .64, η^2 = .01; dysfunctional impulsivity: F_{1,34} = .01, p = .99, η^2 < .001) or number of impulsive incorrect responses (functional impulsivity: F_{1,34} = .03, p = .87, η^2 = .001; dysfunctional impulsivity F_{1,34} = .76, p = .39, η^2 = .02).

A second ANCOVA showed that impulsivity as assessed by the BIS had an effect on number of correct answers (F_{1,35} = 5.89, p = .02, η^2 = .14) but not number of impulsive incorrect responses (F_{1,35} = 1.54, p = .22, η^2 = .04). Additional ANCOVAs with specific subscale scores as covariates showed that only the motor impulsivity subscale had an effect, and only on the number of correct responses (F_{1,35} = 5.69, p = .02, η^2 = .14), not the number of impulsive incorrect responses (F_{1,35} = .61, p = .44, η^2 = .02). The parameter estimate indicated that motor impulsivity had a negative association with number of correct responses (β = −.130, SE = .05, t = −2.39, p = .02). It may also be relevant that, although the association was not significant, the valence of the association between motor impulsivity and impulsive incorrect responses was instead positive (β = .07, SE = .08, t = .78, p = .44).

### 4.4. Effects on physiological activation

We also investigated whether physiological activation varied according to participants’ cognitive behaviour. As stated in the Introduction, this analysis was exploratory and it did not reveal any differences between the stimulation conditions in any of the physiological indices.

In further analysis we used GLM ANOVA with the baseline values of physiological parameters as covariates in order to assess whether differences in physiological activation prior to the task might have influenced performance. There was a positive effect (β = .07, SE = .03, t = 2.50, p = .01) of BVP on the number of correct responses (F_{1,31} = 6.23, p = .01, η^2 = .17). There was a similar positive effect (β = .03, SE = .01, t = 2.33, p = .02) of PVA on number of correct responses (F_{1,31} = 5.45, p = .02, η^2 = .15).

### 5. Discussion and conclusions

The main aim of this study was to investigate the role of the DLPFC in inhibitory control processes involved in problem solving, using tDCS. Some problems are hard to solve as they require solvers to inhibit their spontaneous tendency to rely on the response that comes immediately to mind, which although it seems to be adequate is actually wrong, and to switch from this automatically elicited mental set, reframing the problem in order to find the correct answer. It has been suggested that as the DLPFC is involved in executive functions, stimulating it would modulate its contribution to this specific kind of problem solving. We tested this hypothesis by requiring individuals who had undergone tDCS of the DLPFC to solve the CRT and other similar mathematical problems designed to induce solvers to provide an impulsive incorrect response which, therefore, require inhibitory control to solve. Specifically, we hypothesised that anodal stimulation would increase the proportion of correct responses and decrease the proportion of impulsive responses, whereas cathodal stimulation was expected to produce the opposite effect. The results were partially in line with our predictions; participants who underwent anodal conditions (mean difference = .07, SE = .75, p = 1.00).

Overall, the analyses showed that participants who underwent cathodal stimulation provided fewest correct responses and that anodal stimulation did not improve recognition of correct solutions (Fig. 2). These results indicate that cathodal stimulation caused a decline in performance, by enhancing participants’ tendency to respond impulsively.

### 5.1. Effects on performance

The analysis of number of correct and impulsive incorrect responses to CRT problems by stimulation condition. Error bars represent ± 1 SEM.
stimulation provided more impulsive incorrect responses than both the sham stimulation and anodal stimulation groups, but anodal stimulation did not improve problem solving.

In line with the hypothesis about cathodal effects, a tDCS-induced decrease in cortical excitability seems to have interfered with inhibitory control, reducing participants’ ability to suppress impulsive incorrect responses. These results are consistent with previous results demonstrating that the DLPFC is involved in inhibitory control (Aron et al., 2004; Floden and Stuss, 2006; Shimamura et al., 1995; Vendrell et al., 1995). Applying tDCS over the DLPFC has been shown to affect the recognition of the incorrect-ness of impulsive solutions to CRT problems and the ability to inhibit them, depending on the polarity of the stimulation (Juan and Muggleton, 2012). Hsu et al. (2011) reported that participants made more errors in performing a stop signal task after they had received cathodal tDCS stimulation of the pre-supplemental motor area, whereas anodal stimulation facilitated inhibition of impulsive responses. Similarly, in a study of participants exposed to a virtual roller coaster scenario reducing DLPFC activity by cathodal stimulation modulated both feeling of presence in the virtual environment, as indicated by enhanced skin conductance during the virtual experience, and performance of a stop signal task – participants who had received cathodal stimulation made more false alarm responses than participants in the sham and anodal stimulation conditions (Beeli et al., 2008).

Nevertheless, another recently published tDCS study found that stimulation over the right DLPFC did not modulate delayed response inhibition performance (Stramaccia et al., 2015). The authors argued that the result might have been due to features of the experimental protocol, namely that both anodal and cathodal stimulation-induced engagement of the right DLPFC might have been too brief to influence performance on the selected stop signal task. Discrepancies between study outcomes may reflect differences in the nature of the inhibitory control required in the tasks and hence in the involvement of the DLPFC, or they may reflect differences between the roles of the left and right DLPFC in inhibitory control. If the DLPFC contributes to inhibitory control in both the motor and cognitive domains it is likely to depend on the type of task and on the specific processes underpinning the performance.

The relationship between impulsivity trait measures and inhibitory mechanisms was also investigated. It is interesting to note that of all the sub-types of impulsivity assessed by the BIS-11, only motor impulsivity was related to cognitive performance. Analyses revealed a negative effect of motor impulsivity on solution recognition. A recent meta-analysis of rodent behavioural studies revealed that motor impulsivity is associated with poor decision-making; highly motor impulsive rats (distinguished from less motor impulsive rats on the basis of premature responses at baseline) were slower to adopt advantageous choice strategies on the rodent version of Iowa Gambling Test (Barrus et al., 2015). Together with the negative association between motor impulsivity and decision-making this result suggests that, on a behavioural level, impulsivity is related to poor deliberation. Similar evidence from a human study highlighted the potential link between personality traits of impulsivity – assessed through the BIS-11 - and response inhibition, reporting a significant positive correlation between the subscale of motor impulsivity and the error rate in a stop signal task (Caswell et al., 2015). Electrophysiological studies of the neural correlates of impulsive behaviour on go-no go and stop signal tasks have reported that trait impulsivity is associated with enhanced P3 activation (Shen et al., 2014). These data provide support for the notion that there is a difference between the cognitive performance of individuals low and high on self-report measures of trait impulsivity. The negative association between motor impulsivity and recognition of the correct answer on the CRT that we have described here corroborates the earlier research and provides further evidence that trait impulsivity influences inhibitory control.

Lastly, we turn to our exploration of the relationship between various physiological parameters and problem solving strategies. Although there have been reports that physiological parameters change during problem-solving activities (Klinge et al., 1973), during various cognitive tasks (Sosnowski et al., 2010, 2012) and according to the level of cognitive demand (Richter et al., 2008), we failed to find any systematic relationship between problem solving performance and the physiological parameters we measured. One possible explanation for this is that the cognitive demands of the problems were too brief to produce variations in physiological parameters during problem presentation. However, we did find a positive association between baseline BVP (recorded immediately before participants started the task) and insight solution recognition. This result is consistent with the finding that a heart-related index predicted response inhibition on an emotional stop signal task (Krypotos et al., 2011). The moderating influence of heart rate variability (HRV) indices on cognitive performance indicates that individual differences in heart-related indices are reflected in cognitive performance; in fact they predicted ability to inhibit responses. It is important to note, however, that our analyses of baseline values of physiological parameters were not corrected for multiple comparisons and so our findings should be considered preliminary. Replication and further research is required to confirm that HRV measurements predict CRT performance.

This study suffers from two main limitations. Firstly, the task items were not matched in difficulty. The CRT problems were generally perceived as more challenging and this perception was confirmed by data on the number of correct answers. In addition, the low number of correct responses in all the conditions might be read as a hint of a floor effect, possibly caused by the difficulty of the task. Secondly, the sample size was modest. Using a recognition paradigm may have considerably reduced the cognitive demands of the task and it is possible that a version of the task in which participants were asked to generate rather than recognise solutions might have produced an effect of tDCS on number of correct responses as well as suppression of impulsive incorrect responses. On the other hand, the recognition paradigm used has the advantage of increasing participants’ tendency to select the impulsive incorrect response and requiring inhibitory control. Displaying several response options, including both a misleadingly appealing incorrect response, and the correct response may induce cognitive conflict. This cognitive conflict may cause a mental impasse, which is escaped by inhibiting the prepotent but incorrect problem representation and thus finding a more appropriate representation which enables one to solve the problem. Therefore, the occurrence of solutions in this specific setting might have implied both cognitive conflicts and the action of inhibitory control. The way solutions were presented may also explain why anodal stimulation did not affect performance. Displaying the problems and possible solutions separately, with a short interval between presentation of the problem and the response options, might strengthen the effects of anodal stimulation. This should enhance the inhibitory control required to override the impulsive incorrect response that, in the case of a recognition paradigm, is primed and more difficult to inhibit.

Lastly, as we used a between-subjects design we cannot exclude the possibility that participants’ performance may have been affected by extraneous factors other than the stimulation conditions. We could not have used a within-subjects design as this would have meant presenting participants with the same set of problems three times, allowing them to become more accomplished at solving them as the experiment progressed.

In conclusion, this study provides evidence that the DLPFC is
involved in problem solving. Specifically, we have shown that a DLPFC-related inhibitory control mechanism contributes to the suppression of impulsive incorrect responses on the CRT. Reducing cortical excitability over the left DLPFC using cathodal stimulation impaired participants’ suppression of impulsive responses, presumably by interfering with executive processes. In future research it might be worth using a version of the task which requires participants to generate solutions rather than choosing between options, as a recognition paradigm may not be the best method of assessing the specific contribution of DLPFC to mental set switching functions that, together with inhibitory mechanisms, are likely to be involved in problem solving.

Appendix A

**problems**

(CRT) ITEM 1: A bat and a ball cost €10.10 in total. The bat costs €10 more than the ball. How much does the ball cost?
   a) 10 cents
   b) 5 cents
   c) 20 cents
   d) 15 cents
   **Impulsive incorrect response**= 10 cents. The overall cost (€10.10) is split into the two components (10 euros and 10 cents).
   **Correct response**= 5 cents. The difference between 10.05 euros and .05 euros is 10 euros, as stated in the text of the problem.

(CRT) ITEM 2: If it takes 5 machines 5 min to make 5 widgets, how long would it take 100 machines to make 100 widgets?
   a) 10 min
   b) 5 min
   c) 100 min
   d) 50 min
   **Impulsive incorrect response**= 100 min. If the number of machines and widgets change from 5 to 20, also the third datum (number of minutes) would change from 5 to 100.
   **Correct response**= 5 min. Each machine makes the same number of widgets per unit time. If 5 machines produce 5 widgets in 5 min then 1 machine makes 1 widget in 5 min so if the number of machines changes from 5 to 100, the number of widgets produced in a 5-min period will increase from 5 to 100.

(CRT) ITEM 3: There is a patch of lily pads in a lake. Every day the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?
   a) 1 day
   b) 24 days
   c) 36 days
   d) 47 days
   **Impulsive incorrect response**= 24 days. If the lake is covered by lily pads in 48 days half of its surface would be covered in half that number of days.
   **Correct response**= 47 days. On the 48th day the size is totally covered by lily pads. If the area of lily pads doubles in size from one day to the next then on the previous day (the 47th) the lily pads must have covered half the lake.

ITEM 4: A rope ladder hangs over the side of a boat with the bottom rung on the surface of the water. The rope ladder has 6 runs that are 30 cm apart from each other. The tide rises 70 cm. How many rungs will stick out of the water at high tide?
   a) 6 rungs
   b) 2 rungs
   c) 3 rungs
   d) no rungs
   **Impulsive incorrect response**= 3 rungs. If the water rises 70 cm it will cover the bottom rung (which was previously at water level), the second rung (which is 30 cm above the original water level) and the third rung (60 cm above the original water level). It rises further 10 cm, so at high tide it would be between the third and fourth rung. Three rungs (the fourth, fifth and sixth) would stick out of the water.
   **Correct response**= 6 rungs. As the tide rises the whole boat, including the hanging rope-ladder, is pushed up by the rising water, so the same number of rungs (6) will stick out of the water at high tide.

ITEM 5: There are 12 one-cent stamps in a dozen. How many two-cent stamps are there in a dozen?
   a) 12 stamps
   b) 6 stamps
   c) 24 stamps
   d) 18 stamps
   **Impulsive incorrect response**= 6 stamps. If the value of the stamps doubles, their number is halved.
   **Correct response**= 12 stamps. There are always 12 items in a dozen.

ITEM 6: A farmer makes 4 piles of hay in one corner of a field and other 5 piles in another corner. If he merges them how many piles will he have?
   a) 20 piles
   b) 1 pile
   c) 9 piles
   d) 10 piles
   **Incorrect impulsive response**= 9 piles. Adding together the number of piles in each corner makes 9 piles (4 + 5 = 9).
   **Correct response**= 1 pile. If you merge the piles then irrespective of how many there were, you will have one (bigger) pile afterwards.

ITEM 7: You are participating in a run. You overtake the second runner in the last meters before the finish line. In what position did you finish?
   a) second position
   b) first position
   c) position cannot be determined
   d) third position
   **Impulsive incorrect response**= first position. Overtaking a competitor means gaining a position, so if you overtook the person in second you are in the next best position i.e. first.
   **Correct response**= second position. Before overtaking you were in third position. When you overtake someone you take his or her position, so when you overtake the person in second just before the line, that puts you in second position. There was still one person ahead of the person you passed and that person was still ahead of you after you overtook.

ITEM 8: 25 soldiers are standing in a row 3 m from each other. How long is the row?
   a) 70 m
   b) 73 m
   c) 72 m
   d) 75 m
   **Incorrect impulsive response**= 75 m. If there are 25 persons and the distance between each of them is 3 m, you have to multiply the distance by the number of persons (3 m x 25 = 75 m).
   **Correct response**= 72 m. There are only 24 gaps between 25
people, so the length of the row is 3 m x 24 = 72 m.

ITEM 9: A snail starts climbing up a five-meter-high wall in the morning. During day it climbs 2 m and during the night it slips back 1 m. How many days will it take the snail to reach the top of the wall?

a) 3 days
b) 5 days
c) 2 days
d) 4 days

Impulsive incorrect response = 5 days. The snail gains 1 m each day (2 m climbed during the day minus 1 m slipped down during the night) so to climb 5 m it needs 5 days.

Correct response = 4 days. At the end of the first day the snail is 2 m above the ground, but by the next morning it has slipped back to 1 m above the ground. By the end of the second night it is at +2 (it starts +1 m, reaches +3 m by the end of the day but loses 1 m during the night = +2). At the end of the third night it is at +3 m, so on the fourth day it starts from +3 m and climbs 2 m to reach the top of the wall (it won’t slip down from there as it is a flat surface).

ITEM 10: A brick weighs 1 kg plus half a brick. How much does half a brick weigh?

a) 0.5 kg
b) 1 kg
c) 1.5 kg
d) 2 kg

Impulsive incorrect response = .5 kg. The question is about the half of the brick, so the requested weight would be half the mentioned weight.

Correct response = 1 kg. If a brick is split into two halves both halves weigh the same. If the weight of the two halves of a brick is ‘1 kg + half a brick’ then the ‘1 kg’ corresponds to the weight of the first half of the brick.

ITEM 11: There are 5 white and 5 black socks in Franco’s drawer. Franco’s room is in the dark. How many socks should Franco take out of the drawer to be sure that he gets a matching pair?

a) all 10 socks
b) it cannot be determined
c) 3 socks
d) 5 socks

Impulsive incorrect response = It cannot be determined. Since the outcome of any extraction is unpredictable, this applies to the overall situation.

Correct response = 3 socks. There are only two colours of sock, so the third sock taken out will necessarily be the same colour as at least one of the two already taken.

ITEM 12: You go to bed at eight. You set your old analogue alarm clock to wake you up at nine. How many hours of sleep will you get?

a) 1 h
b) 6 h
c) 11 h
d) 13 h

Impulsive incorrect response = 13 h. The time between when you go to bed (8 pm) and when you want to wake up (9 am) is 13 h.

Correct response = 1 h. At 8 pm you set the hands of the clock so that the alarm will ring at 9, so the alarm rings one hour later at 9 pm, waking you up.

ITEM 13: One month has 28 days. How many of the 11 months left have 30 days?

a) 4 months
b) 11 months
c) 6 months
d) 10 months

Impulsive incorrect response = 4 months. April, June, September, and November have exactly (no more and no less than) 30 days. Correct response = 11 months. Every month, except February, has at least 30 days.

Appendix B. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.neuropsychologia.2016.09.010.

References


There is a popular Italian nursery rhyme which adults also often use to remind which months have exactly 30 days. The rhyme goes: ‘Trenta giorni ha no vembre, con aprile, giugno e settembre; di ventotto ce n’è uno; tutti gli altri ne ha trentuno’ [Thirty days have November, April, June, and September. Only one month has twenty-eight days; all the rest have thirty-one]. This is similar to the English rhyme ‘Thirty days have September, April, June, and November. All the rest have thirty-one, save for February, which has twenty-eight clear, and twenty-nine.

Appendix B. Supporting information


