Accepted for publication in *Neuropsychologia*

*Note: This is an uncorrected version of an author's manuscript accepted for publication.* Copyediting, typesetting, and review of the resulting proofs will be undertaken on this manuscript before final publication. During production and pre-press, errors may be discovered that could affect the content.

Nosce te ipsum – Socrates revisited? Controlling momentary ruminative self-referent thoughts by neuromodulation of emotional working memory

Marie-Anne Vanderhasselt¹, Andre R. Brunoni²,³, Tom Loeys⁴, Paulo S. Boggio⁵, Rudi De Raedt¹*

(1) Department of Experimental Clinical and Health Psychology, Ghent University, Belgium.
(2) Clinical Research Centre, University Hospital, University of São Paulo, Brazil.
(3) Department of Neurosciences and Behavior, Institute of Psychology, University of São Paulo, Brazil.
(4) Department of Data Analysis, Ghent University, Belgium.
(5) Cognitive Neuroscience Laboratory and Developmental Disorders Program, Centre for Health and Biological Sciences, Mackenzie Presbyterian University, Sao Paulo, Brazil.

number of words: 6669 (without abstract, references, tables and figures)

*Corresponding Author: Marie-Anne Vanderhasselt, Ph.D. Department of Experimental and Clinical Health Psychology, Ghent University, Henri Dunantlaan 2, 9000 Ghent, Belgium. Tel: 0032 9 264 64 07; Fax: 0032 9 264 64 87. E-mail: MarieAnne.Vanderhasselt@UGent.be
Abstract

It becomes ever more evident that cognitive operations serve as fundamental mechanisms underlying higher order ruminative thoughts. In this sham controlled within subjects study, we performed anodal transcranial Direct Current Stimulation (tDCS) over the left dorsolateral prefrontal cortex (DLPFC) in 32 healthy participants. We tested the causal hypothesis that the relationship between DLPFC activity and ruminative thinking is mediated by working memory operations. We used the Internal Shift Task, a paradigm in which participants have to update and shift between specific (non)emotional representations in working memory. Subsequently, during an unguided rest period approximately twenty minutes after the stimulation, we explored the occurrence of momentary ruminative self-referent thought. The results demonstrated that the influence of anodal tDCS of the left DLPFC (and not of sham stimulation) on momentary ruminative self-referent thinking is mediated by the enhancement of WM operations for angry faces. Moreover, the more individuals ruminate in everyday life (as measured using the Ruminative Response Style), the larger this mediation effect was. These findings suggest that enhancing cognitive self-regulation, by increasing the ability to update and shift away from negative representations in working memory, might help individuals to control unintentional streams of self-referent thoughts that are self-critical and self-evaluative, a thinking style known as rumination.

205 words
1. Introduction

We all spend time thinking about our own feelings, thoughts and behavior. Self-reflection enables us to create and clarify the meaning of past and present experiences (Boyd & Fales, 1983). This issue of self-referent thinking has ever inspired many scholars. Ancient philosophers have pondered on the nature and fundamental conditions of human self-focus as part of a greater understanding of personal identity. The classical aphorism “Nosce te ipsum” (Know thyself), attributed to Socrates, reflects the ultimate goal of using systematic self-questioning to increase one’s self-knowledge and, as a result, to attain wisdom. However, although this process facilitates coping under a variety of circumstances (Austenfeld & Stanton, 2004; Pennebaker & Seagal, 1999; Trapnell & Campbell, 1999), it might turn out to be unconstructive when self-referent thoughts become rather abstract, evaluative and self-critical (e.g., “Why do I always react this way?”, Watkins, 2008). Moreover, self-referent thoughts might evolve in rumination – “a class of conscious thoughts that revolve around a common instrumental theme and that recur in the absence of immediate environmental demands requiring the thoughts” (Martin & Tesser, 1996, pp. 7). Although these repetitive thoughts don’t necessarily have unconstructive consequences (Watkins, 2008), depression vulnerable individuals have the tendency to focus their thoughts on negative information and personal concerns. It is therefore crucial to understand how self-evaluative ruminative thoughts can be regulated in order to prevent them from becoming unintentional and unconstructive, particularly in individuals who demonstrate a tendency to ruminate in everyday life.

Here, a fundamental question is whether the occurrence of ruminative thoughts could be attributed to underlying cognitive phenomena, such as information processing. In studies investigating this question, participants are usually required to make retrospective reports on the content of their usual thoughts, whether or not referring to a moment when they feel
down, sad or depressed (e.g., Joormann, Dkane, & Gotlib, 2006; Joormann & Gotlib, 2008). These studies tap into stable dispositions of rumination, and have shown that these trait scores are associated with cognitive information processes such as working memory for negative information (i.e., angry faces) (De Lissnyder, Koster, Derakshan, & De Raedt, 2010). During every day functioning, however, ruminative thoughts of healthy individuals fluctuate continuously from one moment to another (Moberly & Watkins, 2008). Therefore, it might be important to assess these thoughts as how they occur naturally. In a prior study, we looked into brain activation over a period of 8 minutes, and afterwards asked participants to report their momentary thoughts during this period. Results revealed that the more unwanted self-referent ruminative thoughts were present during the 8 minutes rest period, the less dorsolateral prefrontal cortex (DLPFC, Brodmann 9/46) activation was observed (Kühn, Vanderhasselt, De Raedt, & Gallinat, 2012). Interestingly, the DLPFC seems not only involved in momentary ruminative thinking (e.g., Kühn et al., 2012; Cooney, Joormann, Eugène, Denis, & Gotlib, 2010), but also in the active maintenance of attentional demands and contextual information during cognitive tasks (Ochsner & Gross, 2005; MacDonald, Cohen, Stenger, & Carter, 2000) and cognitive control for mental representations in working memory (Goldman-Rakic, 1994; Barch et al., 1997; Braver et al., 1997; Braver et al., 2001; Curtis & D’Esposito, 2003).

Working memory (WM) refers to a limited-capacity system to simultaneously maintain and select information over a short period of time in the service of current cognitive processes (Baddeley, 1992; Smith & Jonides, 1999; Miyake et al., 2000). The representations maintained in WM can be emotional or not, and the selection of information depends on its importance in the attainment of goals. In fact, it has been hypothesized that the relation between DLPFC and rumination would be mediated by WM operations (Joormann, Yoon, & Zetsche, 2007). In other words, the cognitive capacity to adequately update negative mental
representations in WM, which is associated with DLPFC activation, might influence the occurrence of momentary self-referent thoughts, preventing these thoughts to become uncontrollable and unconstructive (e.g., rumination). This is in line with cognitive theories proposing that a reduced capacity to shift away from internal representations of negative information could be the functional process underlying the inability to control self-referent ruminative thoughts, which can result in a tendency to continuously ruminate over negative thoughts and feelings (e.g., Koster, De Lissnyder, Derakshan, & De Raedt, 2011; Mor & Winquist, 2002; Watkins, 2008). Based on these models, we specifically hypothesized that momentary ruminative self-referent thoughts would be associated with WM processes to update and shift away from angry faces, an ability which is associated with the DLPFC.

To experimentally address this hypothesis, we used Transcranial Direct Current Stimulation (tDCS) which is a non-invasive neuromodulatory technique. This neuromodulatory technique consists of the application of a weak, direct electric current through electrodes positioned over one’s scalp, which are able to reach the neuronal tissue and induce polarization-shifts on the resting membrane potential (Brunoni et al., 2011). Anodal stimulation generally facilitates cortical activity, whereas cathodal tDCS has opposite effects. In many previous studies it could already be demonstrated that tDCS of the left DLPFC enhances cognitive processes, both for non-emotional (e.g., Fregni et al., 2005; Leite, Carvalho, Fregni, & Goncalves, 2011; Mulquiney et al., 2011) as emotional processes (Boggio et al., 2007; Wolkenstein & Plewnia, 2013). In the current experiment, we used anodal left tDCS to causally modulate DLPFC neural activation in order to enhance WM operations, and subsequently explored the occurrence of self-referent ruminative thoughts. Based on the literature on cognitive deficits in rumination, we used the Internal Shift Task (IST, De Lissnyder, Koster, & De Raedt, 2012), an experimental paradigm to evaluate the ability to update and shift between specific emotional and non-emotional representations in
WM (Garavan, 1998; Chambers, Lo, & Allen, 2008). Twenty minutes after the end of the stimulation, we explored the occurrence of momentary ruminative self-referent thoughts during an unguided rest period. We asked individuals to rest without any specific task - which is known to result in a stream of undirected free thoughts - and self-referent ruminative thoughts - were assessed during this period using a short self-report questionnaire (subjects were not aware of the purpose of our study). In this paper, momentary self-referential ruminative thinking refers to a temporary cognitive thought pattern that is highly dependent on situational cues but that is independent of mood. Trait rumination, on the other hand, is defined as “behaviors and thoughts that focus one’s attention on one’s depressive symptoms and on the implications of those symptoms”, and is considered a habitual thinking response to sadness (Nolen-Hoeksema, 1991, p. 569).

Our study hypothesis is based on cognitive theories which propose that a reduced capacity to shift away from internal representations of angry faces (e.g., negative information) could be the functional process underlying the inability to control self-referent ruminative thoughts. Therefore, we hypothesized that the effect of anodal left DLPFC stimulation (versus sham) on momentary ruminative self-referent thoughts would be mediated by the influence of tDCS on WM operations to update and shift away from negative representations. Because the level of momentary self-focus is related to the habitual tendency to ruminate in everyday life (Moberly & Watkins, 2008), we predicted that trait rumination scores would moderate the relationship between neuromodulation, working memory operations for negative material and momentary ruminative self-referent thoughts.
2. Methods and Material

2.1 Participants

Thirty-two right handed volunteers (20F/12M) ranging from 18 to 36 years in age ($M = 22.28$, $SD = 3.74$) participated in this study. All participants were university students or had a university degree and reported no history of depression or anxiety. The study was approved by the institutional ethics committee of the Mackenzie Presbyterian University, Brazil and by the National Ethics Committee (SISNEP, Brazil).

2.2 Transcranial Direct Current Stimulation

Direct electrical current was applied by a saline-soaked pair of surface sponge rubber electrodes ($35 \text{ cm}^2$) and delivered by a battery-driven stimulator. To stimulate the left DLPFC, the anode electrode was positioned centered over F3 according to the 10–20 international system for electroencephalogram electrode placement. The cathode was placed over the contralateral supra orbital area. This electrodes placement and method of DLPFC localization is in accordance with prior tDCS studies over the left DLPFC looking at WM processes (Fregni et al., 2005). A constant current of 2 mA intensity was applied for 20 minutes. We used a sham controlled within subjects design in which all participants are their own control. For sham stimulation, the position of the electrodes was exactly the same as during tDCS stimulation; however, the current was ramped down after 30 seconds. This procedure is commonly used by tDCS researchers and has been found to be an almost optimal and reliable sham condition (Brunoni et al., 2011). Most participants (27/32) could not distinguish tDCS from sham stimulation. When they received the choice to guess whether they had received tDCS or sham stimulation, most of them did not want to guess and these participants were considered as unable to distinguish between tDCS and sham stimulation. To avoid carry-over effects from the previous stimulation, the second session was carried out after an interval of at
least forty-eight hours. The experiment number assigned to the participant defined the order of stimulation (first tDCS or first sham) for that specific participant.

2.3 Materials

2.3.1 Internal Shift Task (IST). The task was programmed using E-prime 2.0 software package and ran on a Windows XP computer with a 75 Hz, 19-inch color monitor. The stimuli were angry and neutral faces of males and females, based on a validation (Goeleven, De Raedt, Leyman, & Verschuere, 2008) of the Karolinska Directed Emotional Faces (Lundqvist, Flykt, Öhman, 1998). For complete information on these stimuli, we refer to De Lissnyder et al. (2010).

In the IST, faces are presented at the centre of the computer screen one by one. All participants were asked to complete two separate task conditions (in counterbalanced order), a non-emotion and an emotional condition. In the ‘non-emotion condition’, participants had to focus on the gender dimension of the face (the faces had to be categorized as male or female). In the ‘emotion condition’, participants had to focus on the facial emotional expression (the faces had to be categorized as neutral or angry). The participant’s task was to keep a silent mental count of the number of faces in each category, presented within a block of trials (participants had to count the number of male and female faces in the non-emotion condition; participants had to count the number of neutral and angry faces in the emotion condition). When a face was presented, participants were asked to press the spacebar as fast as possible to indicate that they had updated the number of faces seen in one of both categories, which means that they had updated the internal counters of their WM. Then, the next face appeared on the screen after a 200ms intertrial interval. Participants had to report the number of faces of both categories, using the number path of the keyboard at the end of each block in a fixed order to encourage a consistent counting strategy (e.g., in the emotion condition they had to
report their counts first for the neutral and then for the angry faces, in the non-emotion condition the order was male-female).

Due to the sequence of the faces, there were shift and no shift trials in each block. A trial is regarded as a shift if a target trial has to be updated in WM on a different category as its preceding trial (i.e., in the emotion condition an angry face following a neutral face). A trial is regarded as a no-shift if a target trial has to be updated in WM on the same category as its preceding trial (i.e., in the emotion condition an angry face following an angry face).

The practice phase consisted of 3 blocks of trials and the experimental phase of 12 blocks of items in each condition (non-emotion and emotion condition). Within each block, a random number between 10 to 14 trials (or faces) was presented. An example of the stimulus display in one single trial is presented in Figure 1.

--- Figure 1 about here ---

2.3.2 The Momentary Ruminative Self-focus Inventory. To measure the occurrence of self-referent thoughts following each session (tDCS or sham), we used a questionnaire that measures momentary self-reflective rumination (Mor, Marchetti, & Koster, 2013) (Portuguese version translated by the authors). All questions relate to self-referent, ruminative thoughts as a particular self-focus on feelings, reactions and sensations without immediate environmental demands. The statements are not inherently negative or positive, and are considered as a state measure of ruminative thinking. Given the timeframe of our study, we have used an earlier version of this questionnaire that consisted of 8 questions (instead of 6 questions that are part of the current version) (see Table 1), for example “Right now, I am thinking about how happy or sad I feel” and “Right now, I wonder why I react the way I do”. Participants were requested to indicate whether they were engaging in these thoughts during the relax period (see procedure). They were asked to respond using a seven-point Likert scale ranging from 1 (totally disagree) to 7 (totally agree) in order to measure the intensity of self-referent thinking.
In the current study, the questionnaire had acceptable internal consistency during tDCS (Cronbach’s alpha=.70) and sham (Cronbach’s alpha=.79) condition.

--- Table 1 about here ---

2.3.3 Rumination questionnaire. In order to assess trait tendencies to ruminate, the Rumination Response Scale was administered (RRS, Treynor, Gonzalez & Nolen-Hoeksema, 2003, Portuguese version translated by the authors). The RRS consists of items that describe responses to a depressed mood, related to focussing on the self, on symptoms, and on the origin and consequences of the distress. This self-report questionnaire consists of 22 questions to which participants respond on a 4-point Likert scale how often they engage in these responses (i.e. 1 = almost never, 2 = sometimes, 3 = often, 4 = most of the time-). In the current study, the questionnaire has good internal consistency (Cronbach’s alpha=.81).

2.3.4 Mood Ratings. The Positive and Negative Affect Schedule (PANAS, state version) (Watson, Clark, & Tellegen, 1988) was administered to measure potential mood changes induced by electrical stimulation. The PANAS is a commonly used 20-item self-report questionnaire (10 for positive affect (PA) and 10 for negative affect (NA)) developed to measure positive and negative affect. PA represents emotions such as enthusiasm and alertness, and NA represents emotions such as subjective distress and un-pleasurable engagement.

2.4 Procedure

After completing the informed consent form, participants filled out the RRS during the first session of the experiment. Subsequently, electrodes were soaked in saline solution and placed on the participant’s scalp using the electrode montage described above. Following 5 minutes of stimulation (tDCS or sham), participants started with the IST during the remaining 15 minutes of stimulation. Participants were asked to perform the task as quickly and accurately as possible. Approximately 20 minutes after stimulation (during which the
tDCS device was removed, and the PANAS was administered), participants were asked to relax for a well-defined period of eight minutes. No specific instruction was given, except to sit still, relax and wait for our sign that the eight minutes would be over. This way, a post-stimulation situation was created with no additional environmental demands that could trigger specific thoughts. After these eight minutes, participants filled out the Momentary Ruminative Self-focus Inventory. The PANAS was administered at three time points: baseline (T₀), immediately after stimulation (T₁), and approximately 60 minutes after stimulation (T₂). The order of tDCS/sham (within-subjects design) was counterbalanced across participants (16 participants received first tDCS, 16 participants received first sham). Participants were fully debriefed at the end of the study. This experiment was part of a larger project investigating also EEG resting state.

2.5 Data analytic plan

Statistical analysis was performed using the SPSS software package (version 16.0). The significance level was set at an alpha level of .05. Effect sizes are reported in the form of partial eta-squared ($\eta_p^2$).

To examine possible effects of tDCS on mood states, two Stimulation (tDCS, sham) x Time (T₀, T₁, T₂) repeated measures ANOVAs were performed, one for Positive Affect (PA) and another one for Negative Affect (NA), state versions. The dependent variables were positive affect (sum of the scores of all emotions related to positive mood (n=10) and all the negative affect items (negative affect of all emotions related to negative mood (n=10)). Where necessary, we applied the Greenhouse-Geisser correction to ensure the assumption of sphericity.

Regarding behavioral performance on the IST, we investigated the effects of neuromodulation on reaction times and error rates. For each participant, median reaction times (RT) for each category over all trials were calculated in order to reduce any influence of
outliers in the within-subject data. All trials, with correct and incorrect responses, were included in the RT data analyses\(^1\). To investigate specific (differential) effects of tDCS on RT for emotional and non-emotional information, a mixed model with fixed effects for \textit{Stimulation} (tDCS, sham), \textit{Condition} (emotion, nonemotion), \textit{Stimulus Type} nested within \textit{Condition} (angry and neutral in emotion, and male and female in nonemotion conditions), \textit{Shift Type} (shift, non shift) and its interactions, and a random intercept for each subject was fitted. The dependent variable was the median RT. Given the strong a priori hypotheses outlined in the introduction, post-hoc tests were also run to assess the differential effect of tDCS in the shift condition in the emotion and non-emotion condition separately. These RT latencies in the IST shift condition are used in the first step of the mediation analysis (see later).

To investigate whether accuracy was different between tDCS and sham stimulation, error rates in all the IST conditions were calculated as the difference between the correct number and the stated number of faces seen in the preceding block (i.e., series of faces), always depending on the condition (emotion or non-emotion). Error rates refer to the total number of errors over all blocks in that condition. We performed a repeated measures ANOVA with \textit{Stimulation} (tDCS, sham) x \textit{Condition} (emotion, non-emotion), followed by paired sample t-tests for both the emotion and non-emotion condition. Error rate data are presented in Table 3.

Finally, in order to test our hypothesis whether WM processes for emotionally negative material (as measured with the IST using angry faces) mediate the relationship between tDCS stimulation and momentary ruminative self-referent thinking (see Figure 2 for an illustration of this model), a within subjects mediation analysis was performed with the analytic approach described by Judd, Kenny, & McClelland (2001). Following the causal step

\(^{1}\)RT analyses with only correct trials (e.g., when people make a mistake reporting the number of faces according the instruction (non-emotion or emotion), this trial is considered erroneous and was left out of the analysis) yielded similar results.
approach of Baron and Kenny (1986), we first assessed the total effect of tDCS on self-referent ruminative thinking using a paired t-test. Even in the absence of a total effect, one can check for the presence of an indirect (i.e., mediated) effect as the latter may have more power to be detected (Rucker, Preacher, Tormala, & Pettyan, 2011). Following Judd et al. (2001), the indirect effect is explored using a 2-step procedure. In a first step, the effect of tDCS on shifting latencies during the IST conditions (the arrow from tDCS/sham to IST in Figure 2) is assessed using a paired t-test, while in a second step the effect of stimulation induced change in shifting latencies during the IST on momentary ruminative self-referent thinking (the arrow from IST to self-referent thinking in Figure 2) is assessed using ordinary least squares estimation. For this second step, it is necessary to calculate: 1) difference scores for the IST shifting latencies and the momentary ruminative self-referent thinking scores, subtracting the scores in the sham condition from the scores in the tDCS condition; and 2) sum scores for the IST shifting latencies in the sham and tDCS condition. After that, the momentary ruminative self-referent thoughts difference score is regressed on the IST shifting latency difference score and the (mean-centered) IST shifting latency sum score. The IST shifting latency difference score is used to assess mediation, while the IST shifting latency sum score is used to exclude “moderated mediation” (i.e. a different effect of IST shifting latencies on momentary ruminative self-referent thinking between tDCS and sham). If both the paired t-test in step 1 and the effect of IST shifting latency difference score in step 2 are significant, one may conclude that stimulation induced shifting latencies during IST, mediate the relationship between tDCS stimulation and momentary ruminative self-referent thinking. To check the specificity of the effects (we had a specific hypothesis for angry → neutral), we performed analyses for all the IST shift variables (emotion and non-emotion condition). Finally, to assess whether the mediated effect is further moderated by trait rumination scores (the arrows from trait rumination in Figure 2), one can look at the effect of (mean-
centered) trait rumination on the stimulation induced change in IST shifting latencies and its interaction effect with IST shifting latencies on self-referent thinking difference scores.

--- Figure 2 about here ---

3. Results

3.1. Effects of tDCS on Mood: analysis of variance

The two Stimulation (tDCS, sham) x Time (T\textsubscript{0}, T\textsubscript{1}, T\textsubscript{2}) repeated measures ANOVAs with positive and negative affect as dependent variable revealed a significant main effect of Time for PA, \(F(2, 30)=17.86, p < .0001, \eta^2_p = .54\], and for NA, \(F(2, 30) = 6.82, p = .004, \eta^2_p = .31\]. Paired t-tests revealed that both in the sham and tDCS condition, participants reported less PA and NA towards the end of the experiment: T\textsubscript{1} to T\textsubscript{0} (\(p < .02\)) and T\textsubscript{2} to T\textsubscript{1} (\(p < .05\)). For both PA and NA, the main effect of Stimulation and interaction with Stimulation did not reach significance, \(Fs < 1.19 \& ps > .32\). This suggests that changes in mood are not different between both stimulation conditions, and therefore are unlikely to confound with the effects on cognitive control. Mean (with SD) scores of PA and NA during tDCS and sham stimulation at three time points - baseline (T\textsubscript{0}), immediately after stimulation (T\textsubscript{1}), and approximately 60 minutes after stimulation (T\textsubscript{2}) - are reported in Table 2.

--- Table 2 about here ---

3.2. Effects of tDCS on IST shifting latencies: analysis of variance

The mixed model approach yielded a main effect for Stimulation \([F(1, 471)=7.55, p = .006]\], Condition, \([F(1, 471)= 7.05, p = .008]\], and Shift Type, \([F(1, 471)=63.92, p < .001]\]. A significant Stimulation x Condition interaction, \([F(1, 471)=56.80, p < .001]\] was observed. Post-hoc tests indicated that participants were faster responding to trials depicting an angry (\(t(31)=4.04, p < .001\)) and a neutral (\(t(31)=3.49, p = .001\)) face in the tDCS compared to the
sham stimulation condition (no differences were found for trials depicting a male and female face between tDCS and sham stimulation, $t$s<$1.59$, $p$s>.12). Moreover, a significant interaction between Stimulation x Shift Type [$F(1, 471)=30.67$, $p$<.001] was found, with post-hoc tests indicating that participants were faster to respond to IST shift trials during the tDCS compared to the sham stimulation ($t(31)=3.61$, $p$=.001) (no differences were observed for the no shift trials, $t(31)=1.03$, $p$=.31). No other main or interaction effects were significant, $F$s<1.85, $p$s>.16. RTs are presented in Figure 3.

--- Figure 3 about here ---

3.3. Effects of tDCS on IST errors: paired t-tests

Overall, the error rates were less than 5% (see Table 3 for the data). A Stimulation (tDCS, sham) x Condition (emotion, non-emotion) repeated measures ANOVA yielded no significant main or interaction effect ($F$s<3.34, $p$s>.08). Moreover, paired t-tests demonstrated no significant differences between error rates during sham and tDCS, both for the emotion ($t(31)=1.13$, $p$=.27) and the non-emotion ($t(31)=1.38$, $p$=.18) condition. Also, we observed no correlations between IST reaction times and error rates during both sham and tDCS, $p$s>.17.

3.4. Effects of tDCS on momentary ruminative self-referent thinking: the total effect

We first assessed the total effect of tDCS on momentary ruminative self-referent thinking scores. The amount of self-referent thinking was only numerically smaller in the tDCS versus the sham condition (tDCS: $M=26.4$, $SD=6.5$ versus sham: $M=27.1$, $SD=8.7$, $p$=.64, see Table 4). Despite the absence of a total effect, it is still worth exploring a mediation effect. Indeed, following Rucker et al. (2011), one may have much more power to detect an indirect effect than a total effect, especially in small samples with small total effects. In other words, a mediator variable could lead to a stronger indirect effect than total effect from the independent to dependent variable. As hypothesized and described in the data
analytic plan, we therefore investigate the influence of updating negative and neutral representations in WM as a possible mediator in this association.

3.5. Effects of tDCS on momentary ruminative self-referent thinking: the mediated effect by WM

In a first step, we checked whether tDCS (compared to sham) increased the ability to specifically shift and update internal representations (negative to neutral, male to female, and vice versa) in WM, as measured by the IST. Although no significant three-way interaction was found in the repeated measured ANOVA (see section 3.2) (even though we observed two two-way interactions), we proceed with testing these particular post-hoc tests given our a priori hypotheses in the emotion condition. To look for the specificity of the effects, we also analyzed the effects of tDCS (versus sham) on updating performance for neutral (neutral → angry in the emotion condition) and non-emotion conditions. We observed that both shifting away from angry to neutral, \( t(31) = 6.80, p < .001 \), and shifting away from neutral to angry, \( t(31) = 5.81, p < .001 \), was faster in the tDCS as compared to the sham condition (see Figure 3). The shifts in the non-emotion condition (male → female; female → male) were not significantly different in the tDCS versus sham condition, \( t < .33 \) (see Figure 3). Because of this non-significant association between tDCS/sham and WM for the non-emotion condition, further mediation analysis for these non-emotion conditions were not performed.

In a second step, we assessed the effect of stimulation induced IST shifting latencies (during sham and during tDCS) on the corresponding momentary ruminative self-referent thinking scores. We first discuss the stimulation induced change of a shift from angry to neutral (angry → neutral). The decrease in IST shift latencies for angry faces during tDCS was significantly associated with a stimulation induced decrease in momentary ruminative self-referent thinking (\( b = 0.009, \beta = 0.363, t = 2.133, p = .041 \)). Given that shifting from angry to neutral is faster during tDCS as compared to sham, and given the effect of tDCS induced
change of IST shifting latencies on momentary ruminative self-referent thinking, we can conclude that the relationship between tDCS and momentary ruminative self-referent thinking is mediated by shifting ability from angry to neutral in working memory. More specifically, the better participants shifted away from angry faces (angry → neutral) during neuromodulation of the DLPFC, the less momentary ruminative self-referent thoughts were reported after the neuromodulation. Importantly, we found no evidence of moderated mediation (p=.836), i.e. no differential effect of IST shifting latencies on momentary ruminative self-referent thinking between tDCS and sham.

Next, we discuss the shift from neutral to angry. We observed that the decrease in IST shifting latencies was no longer significantly associated with a decrease in momentary ruminative self-referent thinking ($b=0.005$, $\beta=0.197$, $t=1.102$, $p=.279$). We also found no evidence of moderated mediation ($p=.785$). Altogether, further analysis using this mediator will not be performed$^2$.

### 3.6. Mediation effect moderated by trait rumination

Finally, we explored whether the mediation effect of shifting away from angry to neutral representations in WM on momentary ruminative self-referent thoughts, is moderated by trait rumination (see Figure 2). First, a linear regression model for the difference in IST shifting latencies during both stimulation conditions (subtracting the scores in the sham condition from the scores in the tDCS condition) on (mean-centered) trait rumination, revealed that the difference in shifting away from angry to neutral is larger with higher trait rumination ($b=-14.7, \beta=-0.50, t=-3.142, p=.004$). Second, we regressed the momentary ruminative self-referent thinking difference score on (1) the IST shifting latencies difference scores, (2) the (mean-centered) trait rumination and (3) the interaction between the stimulation induced IST shifting latencies difference scores and trait rumination (table 5). We

---

$^2$ We also analyzed the data with the non shift trials in the emotion condition (angry -> angry and neutral -> neutral), but these mediation analyses did not reach significance.
observed that the magnitude of the effect of IST shifting latencies on reducing momentary ruminative self-referent thinking does depend on trait rumination ($p=.002$): the effect of tDCS through IST shifting latencies on reduced momentary ruminative self-referent thinking, increases with higher trait rumination (see Figure 4). To illustrate this point, we classified our sample into low and high ruminators based on a median split in Figure 4. While in the majority of the high ruminators, a decrease in RT from tDCS (represented by bullet points) to sham (represented by triangles) is associated with a decrease in momentary ruminative self-referent thinking (i.e. the dashed lines are mostly decreasing from the observed value under the sham condition to the observed value under the tDCS condition). No such association is observed in the low trait ruminators (i.e. the solid lines are mostly flat). From the above described analyses, we can thus conclude that the mediation effect is moderated by trait rumination, more specifically its effect is most pronounced with high levels of trait rumination.

--- Table 3, 4, 5 and Figure 4 about here ---

4. Discussion

We investigated the influence of DLPFC neuromodulation on working memory (WM) operations and momentary ruminative self-referent thinking. Anodal tDCS of the left DLPFC (compared to sham) did not directly influence the occurrence of momentary ruminative self-referent thoughts. However, we observed that the influence of anodal tDCS (and not of sham stimulation) on momentary ruminative self-referent thinking was mediated by the enhancement of WM operations for angry faces. In other words, the better participants could update and shift away from internal representations of angry faces (angry $\rightarrow$ neutral) during neuromodulation of the DLPFC, the less momentary ruminative self-referent thoughts were reported during an 8-minutes rest period after neuromodulation. Moreover, the strength of this
indirect effect was moderated by the tendency to ruminate in everyday life (as measured using the Ruminative Response Style). In other words, the more individuals have the habitual tendency to ruminate, the stronger the mediation effect (see Figure 4 for an illustration of this model). Interestingly, not only the indirect effect was moderated by trait rumination, but we also found evidence for a marginal interaction effect of stimulation (tDCS, sham) and trait rumination on momentary ruminative self-referent thinking ($p=.10$), i.e. a moderated total effect. One possible explanation for the absence of the total effect (i.e., the effect of tDCS on momentary ruminative self-referent thinking) in our study with healthy participants may therefore be that this total effect is a mixture of no effect in low habitual ruminators and an effect in high habitual ruminators. In other words, in low habitual ruminators, it may be hard to see any decrease in momentary ruminative self-referent thinking. Future neuromodulation studies regarding momentary ruminative self-referent thinking should therefore focus on high ruminators and/or a more clinical population with larger variability in habitual rumination scores.

Our results are completely in line with the impaired disengagement hypothesis (Koster et al., 2011), which proposes that a reduced capacity to disengage from internal representations of angry faces could be the functional process underlying the prolonged processing of self-referential material, which can result in a tendency to continuously ruminate over negative thoughts and feelings. Important for the present study, mood was not differentially influenced by the sham and tDCS condition, which means that our effects are not influenced by mood changes. Moreover, tDCS (always compared to sham) did not influence WM operations for non-emotional material (gender), which signifies that there is no interplay between the DLPFC and these WM operations related to ruminative self-referent thoughts. Also, participants were overall very accurate during the IST (less than 5 % errors), and error rates were similar between sham and tDCS.
First, how could we relate the specific interplay between DLPFC neuromodulation, improved mental ability to shift away from negative representations in WM and decreased momentary ruminative self-referent thoughts to everyday life functioning? WM, although classically understood as an executive system that holds and manages external (visual and auditory) information (Baddeley, 2012), also regulates shifting away from internally generated thoughts and feelings. This includes abilities to (i) shield the flow of thoughts from distractive, emotional stimuli and (ii) update goals, by shifting away from failed goals in order to pursue new ones (e.g., Hofmann, Schmeichel, & Baddeley, 2012). These WM operations are part of an everyday self-regulation process to update and shift away from negative representations that seem, based on the present findings, to influence the level of momentary (state dependent) self-focused ruminative thoughts later on. Indeed, well functioning individuals regulate self-critical negative thoughts (e.g., internal representations in WM) which are not in line with existing positive self-views, in order to focus on more positive thoughts, and this cognitive regulation is related to enhanced activation in the DLPFC and a reduced momentary ruminative self-focus.

Second, how could we interpret this interplay between DLPFC neuromodulation, WM for negative representations and momentary ruminative self-referent thoughts as a function of trait rumination? It is known that updating in the context of negative content in WM generates the highest cognitive conflict in well-functioning individuals who report a tendency to ruminate in everyday life (e.g., De Lissnyder et al., 2012; Vanderhasselt et al., 2011; 2013). Therefore, in high ruminators there is a larger possible range for enhancing the cognitive regulation of self-referent ruminative thoughts, and this enhanced mental self-regulation can be achieved by modulating WM operations to shift away from negative representations. Because trait rumination is a thought processing style with unintentional streams of ruminative self-referent thoughts that are rather abstract and self-evaluative (Papageorgiou &
Wells, 2000), the enhancement of specific WM operations might facilitate the regulation of a momentary self-focus to eventually prevent an excessive self-focus on negative content (Banich et al., 2009). Indeed, individuals that have difficulties to disengage from negative cognitive representations, seem to become trapped in vicious cycles of negative self-referent thoughts, and in turn rumination might become an habitual mode of thinking (Koster et al., 2011). To summarize, DLPFC neuromodulation enhanced the cognitive ability to shift away from negative material in WM, and in turn reduced the occurrence of momentary ruminative self-referent thoughts, possibly by making habitual ruminators more able to regulate the incidence of evaluative, ruminative self-referent thoughts.

Even though most studies consider ruminative thinking as a trait characteristic, self-referent thoughts fluctuate continuously (especially in healthy individuals) and might provide valuable information to understand the development of a stable trait. In this study, a self-reported (trait) tendency to ruminate in response to negative mood was associated with a momentary (state) self-focus that was not explicitly associated with an emotional content, as assessed by the occurrence of momentary ruminative self-referent thoughts ($r=0.40$). There are many ways to assess momentary self-rumination (see Smith & Alloy, 2009). In studies that have investigated momentary ruminative thoughts, participants are asked randomly during daily life to report the content of their thoughts (Moberly & Watkins, 2008; Killingsworth & Gilbert, 2010), or ruminative thoughts are induced by asking participants to focus their attention on a specific thought (presented by a statement) for some time (e.g., Cooney et al., 2010; Whitmer & Gotlib, 2012). In the current study, we asked individuals to rest without any specific task, which is known to result in a stream of undirected free thoughts (James, 1890; Filler & Giambrata, 1973; Giambra, 1989). Because we did not interfere during this rest period and asked our questions immediately afterwards, we were able to assess naturally occurring self-focused thoughts without linking them to a precise emotional content or response to
negative mood. During this period of idleness, we found that internally generated momentary ruminative self-referent thoughts were causally influenced by the cognitive capacity to regulate negative mental representations in WM, associated with DLPFC neural activation. In fact, it has already been shown that increased WM performance relates to decreased mind wandering during daily life (e.g., Kane et al., 2007; McVay & Kane, 2010), and even predicts the regulation of thought and behavior (Kane & Engle, 2003). Indeed, an enhancement of the executive-control system to adequately combat interfering thoughts that are generated and maintained automatically, is related to decreased ruminative self-referent thoughts (McVay & Kane, 2010). Our findings go beyond these results by showing that cognitive and neural phenomena can be modulated to increase the ability to regulate momentary ruminative self-referent thoughts during a period of idleness, a process closely linked to the ruminative thinking style. This interplay between biological and cognitive factors is in line with a theoretical framework of De Raedt & Koster (2010), which states that cognitive control processes play a central and causal role in the relation between prefrontal neural activation and rumination. Moreover, the current results go beyond correlational findings by using an experimental method that involves neuromodulation of the DLPFC to temporarily enhance its activity, thus allowing causal inferences. This is an important next step for building and refining our understanding of the neural bases of rumination.

Some limitations of the present study should be emphasized. We measured momentary ruminative self-referent thoughts after the end of the unguided rest period, using simple valence a-specific questions that did not require complicated meta-cognitive judgements. Because in this study trait rumination was associated with self-referent ruminative thinking, we might infer that the content of the self-referent thoughts is negative, focused on problems and unattainable goals (Martin & Tesser, 1996; Pyszczynski & Greenberg, 1987). However, we did not assess the content of these self-referent thoughts. It might also be interesting to
relate the content of thoughts during an unguided rest period to operations of updating specific emotional information in working memory. Moreover, because we were not aware of any existing questionnaires to measure momentary ruminative self-referent thoughts, we used a short (so far unpublished) inventory that has only been used in a limited number of studies (Momentary Ruminative Self-focus Inventory; Mor et al., 2013). It is important to underscore that in the current study, the questionnaire has acceptable internal consistency during both the tDCS (Cronbach’s alpha=.70) and the sham (Cronbach’s alpha=.79) condition. In addition, the fact that we obtained these clear results in line with our hypothesis is indicative of sufficient psychometric properties of this questionnaire. A second limitation concerns the absent basic effects of tDCS on shifting capacities for emotion and non-emotion material. Although the general omnibus ANOVA indicated that tDCS (versus sham) enhanced performance during the IST shift trials over all conditions (interaction between Stimulation and Stimulus Type), and enhanced performance in the emotion condition over all trial types (interaction between Stimulation and Switch Type), the crucial three-way interaction was not significant. This might be attributed to a lack of power to detect such interaction in our small sample. Nevertheless, we observed a beneficial effect of tDCS on shift trials specifically in the emotion condition using (hypothesis driven) post hoc tests. Prior studies have shown that tDCS enhances WM for non-emotional information (e.g., Fregni et al., 2005) and cognitive processes for emotional information (e.g., Boggio et al., 2007; Vanderhasselt et al., 2013). Studies on the effects of tDCS on cognitive functioning have, however, never used the IST to measure WM processes. Therefore, more research is needed to further investigate the precise effects of tDCS on WM for emotional and non-emotional information, to replicate the effects of tDCS measured using the IST. Moreover, in order to draw specific conclusions regarding the causal contribution of the DLPFC in updating emotional representations in WM, potential
contributions of and comparisons with other brain regions (e.g., right DLPFC) should be tested as well.

In conclusion, the effect of anodal left DLPFC neuromodulation on the occurrence of momentary ruminative self-referent thoughts was mediated by updating and shifting away from negative representations in WM. This interplay is most pronounced in healthy volunteers who report a higher tendency to ruminate in everyday life, which signifies that working memory processes might be fundamental processes underlying a ruminative self-focus.
Acknowledgements

MAV (FWO08/PDO/168) is a postdoctoral research fellow of the Research Foundation Flanders (FWO). Preparation of this paper was also supported by Grant BOF10/GOA/014 for a Concerted Research Action of Ghent University (awarded to RDR). PSB is supported by a CNPq researcher grant (305718/2009-6). TL would like to thank the Research Foundation Flanders (FWO) for financial support (Grant G.0111.12).
References


Therapy, 37(3), 269-280.


Table 1. Questions from the Momentary Ruminative Self-focus Inventory (Mor et al., 2013)

<table>
<thead>
<tr>
<th>Item nr</th>
<th>Sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Right now, I am conscious of my inner feelings.</td>
</tr>
<tr>
<td>2.</td>
<td>Right now, I am reflective about my life.</td>
</tr>
<tr>
<td>3.</td>
<td>Right now, I am aware of my innermost thoughts.</td>
</tr>
<tr>
<td>4.</td>
<td>Right now, I am thinking about how happy or sad I feel.</td>
</tr>
<tr>
<td>5.</td>
<td>Right now, I am thinking about the physical sensations I feel in my body.</td>
</tr>
<tr>
<td>6.</td>
<td>Right now, I wonder why I react the way I do.</td>
</tr>
<tr>
<td>7.</td>
<td>Right now, I am thinking about how tired or alert I feel.</td>
</tr>
<tr>
<td>8.</td>
<td>Right now, I am thinking about the possible meaning of the way I feel.</td>
</tr>
</tbody>
</table>

Table 2. Mean (with SD) scores of positive and negative mood during tDCS and sham stimulation at three time points: baseline (T0), immediately after stimulation (T1), and approximately 60 minutes after stimulation (T2).

<table>
<thead>
<tr>
<th></th>
<th>tDCS</th>
<th>sham</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T1     T2     T3</td>
<td>T1     T2     T3</td>
</tr>
<tr>
<td>Positive Affect</td>
<td>29.78 (6.46) 28.40 (7.07) 25.19 (7.52)</td>
<td>30.84 (4.17) 27.81 (4.80) 25.05 (6.73)</td>
</tr>
<tr>
<td>Negative Affect</td>
<td>14.56 (4.68) 13.28 (3.68) 12.69 (3.79)</td>
<td>14.00 (3.88) 13.25 (3.06) 12.86 (3.24)</td>
</tr>
</tbody>
</table>

Table 3. Mean (with SD) of the total error rates for all IST conditions during tDCS and sham stimulation.

<table>
<thead>
<tr>
<th>Error Rates</th>
<th>tDCS</th>
<th>Sham</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emotion</td>
<td>2.03 (2.19)</td>
<td>2.65 (3.32)</td>
</tr>
<tr>
<td>non-emotion</td>
<td>2.91 (300)</td>
<td>3.78 (4.28)</td>
</tr>
</tbody>
</table>
Table 4. We present an overview of all the variables and statistical findings of our mediation/moderation model. The first two columns present (1) median RT latencies for the IST trials angry -> neutral; (2) the sum score of the Momentary Ruminative Self-focus Inventory (Mor et al., 2013); and (3) the sum score of the Rumination Response Scale (RRS, Treynor et al., 2003) in the two different stimulation conditions (first two columns). In the third column (grey shaded), we present the t-value of the difference between both stimulation conditions. In the last three columns, we present the correlations coefficients for all the variables of our mediation/moderation model. Correlation above and below the diagonal are correlations in tDCS and sham, respectively.

<table>
<thead>
<tr>
<th></th>
<th>tDCS Mean (SD)</th>
<th>sham Mean (SD)</th>
<th>t-value</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RT angry -&gt; neutral (1)</td>
<td>1035 (285)</td>
<td>1426 (444)</td>
<td>6.80**</td>
<td>-</td>
<td>.36*</td>
<td>.27</td>
</tr>
<tr>
<td>Self-referent Thinking</td>
<td>26.4 (6.5)</td>
<td>27.1 (8.7)</td>
<td>.47</td>
<td>.33</td>
<td>-</td>
<td>.15</td>
</tr>
<tr>
<td>Rumination: Total Score</td>
<td>39.9 (11.2)</td>
<td>39.9 (11.2)</td>
<td>-</td>
<td>.55**</td>
<td>.40*</td>
<td>-</td>
</tr>
</tbody>
</table>

Note. *p < .05. **p < .01.

Table 5. Summary table for the linear regression model with difference in momentary ruminative self-referent thinking (tDCS-sham) as outcome.

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>s.e.(B)</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>2.038</td>
<td>2.094</td>
<td>0.973</td>
<td>.339</td>
</tr>
<tr>
<td>Difference in IST</td>
<td>0.002</td>
<td>0.005</td>
<td>0.541</td>
<td>.593</td>
</tr>
<tr>
<td>State self-referent thinking</td>
<td>0.6526</td>
<td>0.2540</td>
<td>2.570</td>
<td>.016</td>
</tr>
<tr>
<td>State self-referent thinking</td>
<td>*</td>
<td>0.0012</td>
<td>3.525</td>
<td>.002</td>
</tr>
</tbody>
</table>

Note. *p < .05.
Figure 1. An example of the stimulus display of a trial during the emotion condition (left) and the non-emotion condition (right) of the Internal Shift task (IST).
Figure 2. An outline of the mediation/moderation model. The effect of tDCS on momentary ruminative self-referent thoughts is mediated by shifting away from angry to neutral representations in WM (as measured using the IST: angry -> neutral). Furthermore, trait rumination (measured using the Ruminative Response Scale) moderates this indirect effect.
Figure 3. A) Median reaction times (with SD) for all the shifting IST condition, both during tDCS and sham stimulation. We observed significant differences for reaction times angry -> neutral and neutral - > angry categories between tDCS and sham stimulation. B) Median reaction times (with SD) for all the non shifting IST conditions, both during tDCS and sham stimulation.

Note: ** < .001.
Figure 4. Scatterplot demonstrating the relation between RT angry to neutral (x-axis) momentary ruminative self-referent thinking (y-axis) with tDCS and sham stimulation (triangle and circle points, respectively). Lines connect observation points under tDCS and sham from the same participant and show whether participants scored low or high on trait rumination (solid and dashed lines, respectively). Note that there are 2 points without connecting line because of missing information on the trait rumination for that particular subject.