



# Selective changes in moral judgment by noninvasive brain stimulation of the medial prefrontal cortex

Paolo Riva<sup>1</sup> · Andrea Manfrinati<sup>1</sup> · Simona Sacchi<sup>1</sup> · Alberto Pisoni<sup>1</sup> · Leonor J. Romero Lauro<sup>1</sup>

Published online: 8 November 2018  
© Psychonomic Society, Inc. 2018

## Abstract

Multiple cortical networks intervene in moral judgment, among which the dorsolateral prefrontal cortex (DLPFC) and the medial prefrontal structures (medial PFC) emerged as two major territories, which have been traditionally attributed, respectively, to cognitive control and affective reactions. However, some recent theoretical and empirical accounts disputed this dualistic approach to moral evaluation. In the present study, to further assess the functional contribution of the medial PFC in moral judgment, we modulated its cortical excitability by means of transcranial direct current stimulation (tDCS) and tracked the change in response to different types of moral dilemmas, including *switch*-like and *footbridge*-like moral dilemmas, with and without personal involvement. One hundred participants (50 males) completed a questionnaire to assess the baseline levels of deontology. Next, participants were randomly assigned to receive anodal, sham, or cathodal tDCS over the medial prefrontal structures and then were asked to address a series of dilemmas. The results showed that participants who received anodal stimulation over the medial PFC provided more utilitarian responses to *switch*-like (but not *footbridge*-like) dilemmas than those who received cathodal tDCS. We also found that neurostimulation modulated the influence that deontology has on moral choices. Specifically, in the anodal tDCS group, participants' decisions were less likely to be influenced by their baseline levels of deontology compared with the sham or cathodal groups. Overall, our results seem to refute a functional role of the medial prefrontal structures purely restricted to affective reactions for moral dilemmas, providing new insights on the functional contribution of the medial PFC in moral judgment.

**Keywords** Moral judgment · Moral dilemmas · Dual-process theories · Medial prefrontal cortex · Transcranial direct current stimulation

Individuals often have to make decisions in which they must choose whether to follow a universal moral imperative or implement a cost-benefit analysis that reflects a utilitarian approach. Put simply, they are faced with a moral dilemma, and moral dilemmas are both interesting and useful, because they evoke competing, incompatible judgments (Hauser, 2006; Mikhail, 2011; Greene, 2014). Cognitive neuroscience has identified several key brain regions involved in moral

decision making. Critically, two main cortical areas have emerged as part of a cortical network that plays a pivotal role on individual responses to moral dilemmas: the dorsolateral prefrontal cortex (DLPFC; see Tranel, Bechara, & Denburg, 2002; Tassy et al., 2012) and the medial prefrontal structures (i.e., medial PFC and ventromedial PFC; see Moll, Zahn, de Oliveira-Souza, Krueger, & Grafman, 2005; Moll & de Oliveira-Souza, 2007; Greene, Nystrom, Engell, Darley, & Cohen, 2004). A dualistic approach to moral judgment suggested that the DLPFC is mainly responsible for cognitive control, whereas the medial prefrontal structures underlie emotional impulses (Greene, Sommerville, Nystrom, Darley, & Cohen, 2001; Greene, 2014). However, past research disputed this role for the DLPFC (Tassy et al., 2012). In the present study, we focused on the functional contribution of medial PFC in the context of moral dilemmas. We modulated the cortical excitability of this cortical region and observed changes in participants' reactions when confronted with different moral dilemmas.

---

This work was supported by a Grant from the Italian Ministry of Education, University, and Research (FIRB: RBFR128CR6). Correspondence concerning this article should be addressed to Paolo Riva, University of Milano-Bicocca, Department of Psychology, Piazza Ateneo Nuovo, 1, 20126 – Milano (Italy).

✉ Paolo Riva  
paolo.riva1@unimib.it

<sup>1</sup> University of Milano-Bicocca, Milan, Italy

## Models of morality

Moral issues are not all alike, and people respond differently to different types of dilemmas. Consider in this regard the longstanding philosophical debate over two well-known moral dilemmas (Foot, 1967, Thomson, 1986). The *trolley* dilemma asks people to imagine a runaway trolley is about to run over five workers on the track. The question is whether you would hit a switch turning the trolley on another track where only one man is standing and will be killed. The *footbridge* dilemma is similar to the *trolley* dilemma, but with one difference: the only way to save the five workmen is to push a man who is standing on the footbridge into the path of the trolley, thus killing the stranger but preventing the trolley from reaching the others.

Decades of research shows that the majority of people faced with the two types of dilemmas choose to adopt a utilitarian approach in the *trolley* dilemma by deciding to switch the track, whereas only a small minority (roughly 10%) decide to push the stranger off the bridge in the *footbridge* dilemma (Hauser, Cushman, Young, Kang-Xing Jin, & Mikhail, 2007). Accounting for these differences, studies found higher levels of emotional activation for *footbridge*-like dilemmas than *switch*-like dilemmas (Borg, Hynes, Van Horn, Grafton, & Sinnott-Armstrong, 2006; Greene et al., 2001). *Footbridge*-like dilemmas require the decision to use a person as a means to an end, whereas in *switch*-like dilemmas, killing one person to save more people is a foreseen but unintended consequence of their action (Manfrinati, Lotto, Sarlo, Palomba, & Rumiati, 2013). The ratio of lives and deaths in switch-like and footbridge-like dilemmas is the same; what changes is whether the harm can be considered as a side effect of the implementation of utilitarian behavior (*switch*-like dilemmas) or whether direct harm is required to implement the utilitarian behavior (*footbridge*-like dilemmas). Thus, the direct infringement of a moral taboo (e.g., personally and intentionally killing a person as a means to save more lives) in *footbridge*-like dilemmas is thought to trigger emotional responses that prevent most people from choosing that course of action. Deontological preferences appear to be guided mostly by affective processes, whereas utilitarian judgments appear to be shaped mostly by deliberative responses.

In past research, moral dilemmas typically considered hypothetical real-life scenarios involving the life or death of *other* people. However, another relevant factor that can influence people's responses to moral dilemmas relates to the degree of *self*-involvement in these scenarios (Lotto, Manfrinati, & Sarlo, 2014). Dilemmas can be construed by manipulating *self*-involvement in such a way that the main character's life is at risk (in both *footbridge*-like and *switch*-like dilemmas). In these scenarios, killing might result in saving one's own life while saving others' lives (*self*-involvement dilemmas) or may involve saving the lives of only other people

(*other*-involvement dilemmas). Unsurprisingly, people are more willing to kill to save themselves and others than to save only others (Lotto et al., 2014). In this case, feelings of self-preservation may make people more prone to choose a given course of action (e.g., killing someone else) regardless of the ultimate tradeoff between costs and benefits. Accordingly, empirical data showed higher levels of self-reported emotional activation for *self*-involvement when compared with *other*-involvement dilemmas (Lotto et al., 2014).

Overall, judgments on *footbridge*-like dilemmas appear to be primarily led by emotional reactions and prescriptive rules (such as the deontological precept of “not to kill” and self-preservation). In contrast, *switch*-like dilemmas, which are characterized by lower levels of emotional activation and a lack of prescriptive rules, may allow for calculated responses that depend on cost-benefit analyses. Which of these moral dilemmas would be more strongly influenced by changes in cortical excitability of the anterior portion of medial prefrontal structures? The answer to this question would enable us to make causal inferences on the functional role of this brain region on moral judgment.

## The various roles of medial PFC in moral judgment

Existing neuropsychological and functional imaging (fMRI) studies identified a potential network of neural regions involved in moral judgment that includes several cortical and subcortical regions, including the prefrontal cortex, specifically the ventromedial, dorsolateral, and medial and lateral orbitofrontal portions, the anterior temporal lobes, the superior temporal sulcus and the anterior and posterior cingulate cortex, amygdala, and precuneus (Forbes & Grafman, 2010; Fumagalli & Priori, 2012). Among these regions, two cortical areas have emerged for their pivotal role in moral judgment. The first region, the dorsolateral prefrontal cortex (DLPFC), has been traditionally linked with cognitive control over moral decision making and utilitarian responses. The second region, the medial PFC, has been associated with emotional reactions to moral dilemmas and deontological responses.

Within cognitive neuroscience, this dichotomy between reason and feeling has been mostly endorsed by one of the best-known theories of moral judgment (i.e., *dual process theory*; Greene, 2014). According to this theory, the medial PFC, and in particular its ventral part (i.e., the vmPFC) primarily acts as an “alarm bell” such that when facing severe moral transgressions (e.g., pushing a man off a footbridge to stop a trolley), the emotional value detected by the vmPFC prevents the decision maker from choosing a particular course of action. By contrast, utilitarian responses when facing severe moral transgression arise from cognitive control mechanisms based in the DLPFC.

However, recent theory and research disputed the predominant “rational” cognitive control role of the DLPFC (Moll and de Oliveira-Souza, 2007; Talmi and Frith, 2007). More specifically, Tassy et al. (2012) adopted a neuromodulatory technique (i.e., repetitive transcranial magnetic stimulation; rTMS) to interfere with right DLPFC activity while participants were confronted with moral dilemmas. The authors found that disrupting rDLPFC activity increased utilitarian tendencies. This result is in contrast to the dual system hypothesis that would predict that DLPFC should underlie cognitive control over emotional impulses. By contrast, the right DLPFC may be part of a psychological system that participates to the integration of representational emotions during moral evaluation (Moll et al., 2005). This raises the possibility of a different framework of prefrontal cortex in which cognitive control and emotion are not competitive mechanisms but integrated and interactive processes (Pessoa, 2013).

In a similar vein, the predominantly affective role of the medial PFC during decision making in the context of moral dilemmas is currently under debate. On one hand, classical studies supported the idea of medial prefrontal structures, in particular their ventral part, as an emotional area, by showing that patients with vmPFC lesions exhibit decreased emotional responsivity and social emotions (e.g., compassion; Damasio, Tranel, & Damasio, 1990; Damasio, 1994; Koenigs et al., 2007) and a greater frequency of utilitarian judgments in dilemmas typically triggering strong emotions (Ciaramelli et al., 2007; Koenigs et al., 2007; Ciaramelli, Braghittoni, & di Pellegrino, 2012). Moreover, fMRI studies found that the reasoning behind emotionally engaging dilemmas (e.g., *footbridge*-like) is associated with increased vmPFC activation (Greene et al., 2001). Thus, deontological options may be the product of negative emotional responses that at least partially depend on vmPFC activity (Greene et al., 2001; Greene, Nystrom, Engell, Darley, & Cohen, 2004).

On the other hand, the theory of moral judgment proposed by Moll et al. (2005; 2007), in which moral decision making is implemented by a single set of brain areas, represents a valid alternative to Greene’s dual-process theory. A central issue in studies on the relationship between morality and brain damage relates to the precise anatomical distribution of the lesions. In the Koenigs et al. (2007) study, in the vmPFC group prefrontal damage extended bilaterally to the medial frontopolar cortex (FPC) in five of the six patients and to the lateral FPC (including the anterior dLPFC) in four of them. In 2005; 2007, the vmPFC and the FPC, in conjunction with temporal cortex and limbic and paralimbic systems, play a distinguished role in the experience of prosocial sentiments (i.e. guilt, compassion and interpersonal attachment), whereas the ventrolateral PFC (with lateral sectors of DLPFC) is more relevant for the experience of anger or

indignation. Based on this framework, medial prefrontal structures represent different aspects of social knowledge, which are bound to emotional relevance. Consequently, these representations guide the assessment of social-emotional outcomes associated with behavioral choices, such as prospective thinking and representing multiple outcomes of events and actions. Furthermore, this network is not only necessary for the experience of prosocial moral sentiments but also plays a role in moral calculus: “A moral calculus results from the ability to envision a number of action-outcome options in a parallel fashion, and compare their relative weights” (Moll, De Oliveira-Souza, Zahn, Grafman, 2008b, p. 6). Indeed, the authors predict that a lesion of the anterior PFC would lead to selective impairments in moral evaluations that rely on predicting the long-term outcomes of one’s own actions. More specifically, the activation of this network during moral judgment results from representing possible outcomes and how they branch into the future. This would explain anterior PFC activation in reflective moral reasoning (Moll, de Olivera-Souza, & Eslinger, 2003) and in utilitarian moral judgments (Greene et al., 2004).

In line with this perspective, an fMRI study (Prehn, Wartenburger, Mériaux, Scheibe, Goodenough, et al., 2007) showed that activity in the vmPFC is modulated by individual differences in moral judgment competence (i.e., the ability to apply moral orientations and principles in a consistent and differentiated manner in varying social situations; Lind, 2008). When identifying social norm violations, participants with lower moral judgment competence recruited the left vmPFC more often than participants with greater competence. Because increased activation in individuals with lower moral judgment competence may be due to the increased recruitment of mental resources, the authors proposed that the augmented activity in the vmPFC corresponds to an increased involvement of social cognitive and emotional processes, such as mentalizing or estimating the value of possible outcomes of a behavior and the experience of moral emotions during moral judgment (see also Amodio & Frith, 2006).

In conclusion, these findings are in opposition to the dual process theory proposed by Greene and colleagues (Greene et al. 2001), because they advance the hypothesis that moral reasoning and emotion depend on associatively linked representations within *fronto-temporo-limbic* networks (Moll et al., 2005). According to this view, all morally relevant experiences are considered to be essentially cognitive/emotional association complexes. Instead of competing with each other, cognition and emotion are continuously integrated during moral decision making, and the key site in which this integration is made possible is the medial prefrontal cortex (Moll et al., 2005; Moll, de Olivera-Souza, & Zahn, 2008a; Pessoa, 2013).

## Individual differences

Besides some research programs aimed to identify universal principles of moral cognition (Hauser, 2006; Mikhail, 2007), more recent research suggested that moral judgment is a phenomenon subject to major interindividual differences (Bartels, 2008; Feltz & Cokely, 2008; Lind, 2008; Prehn, Wartenburger, Mériaux, Scheibe, Goodenough, et al., 2007). For instance, a set of studies showed a positive relation between utilitarian preferences and working memory capacity (Moore, Clark, & Kane, 2008), tendency to deliberative rather than intuitive thinking (Bartels, 2008), measures of psychopathy, Machiavellianism, and life meaninglessness (Bartels & Pizarro, 2011). Considering more traditional variables, Atran (2002) and Boyer (2003) suggested that across different cultures, a common ground to justify moral decisions entails the religious belief on supernatural agents, and that imagining empathetic support from a supernatural agent may facilitate the adjudication of moral dilemmas and the justification of hard moral decisions. Existing research shows that moral beliefs are also deeply connected with political inclination (Haidt & Graham, 2007; Lakoff, 2002). Consider, for example, the different moral views that are endorsed by liberal and conservative in the American political environment.

In addition to these variables, a key dimension that could predict utilitarian responses on moral dilemmas relates to the individual differences on deontology. People with a strong deontological orientation endorse the existence of moral obligations requiring or prohibiting certain actions regardless of their consequences (Baron & Spranca, 1997; Sacchi, Riva, Brambilla, & Grasso, 2014). Standing with categorical imperatives (Kant, 1785/1959) should imply lower support for utilitarian-consistent solutions. Research showed that the greater an individual's endorsement of deontological principles, the lower the endorsement of utilitarian solutions was (Xu & Ma, 2015). Thus, it is plausible to expect that people who indicate greater agreement with deontological principles will be less likely to engage in a utilitarian calculus and perceive a five-lives-for-one tradeoff permissible.

Another factor known to influence moral judgment is gender. The available research shows that women tend to exhibit stronger deontological inclinations than men (Friesdorf, Conway, & Gawronski, 2015; Gilligan, 1982; Jaffee, & Hyde, 2000). Crucially to the purposes of the present investigation, research found differences in the neural structures involved in moral judgment of females and males (Harenski, Antonenko, Shane, & Kiehl, 2008). Similarly, previous studies showed that anodal transcranial direct current stimulation (tDCS) of the ventral prefrontal cortex (VPC) increased utilitarian responses to moral dilemmas, whereas cathodal tDCS tended to decrease it (Fumagalli et al., 2010). However, this

effect occurred only in females; males were unaffected by the manipulation of cortical excitability.

## The present study

Recent research showed that it is possible to shift individuals' moral judgment by modulating the cortical excitability of brain regions involved in moral behavior (Fumagalli et al., 2010; Tassy et al., 2012). In the present study, we modulated the cortical excitability of medial prefrontal structures through transcranial direct current stimulation (tDCS) and observed participants' reactions to different types of moral dilemmas.<sup>1</sup> Our first goal was to test the hypothesis that neuromodulation differently influenced moral judgment according to the dilemma type (e.g., *switch*-like vs. *footbridge*-like dilemmas). These data can provide additional information on the functional role of medial prefrontal structures in moral judgment. Indeed, different types of dilemmas relate to different underlying constructs (e.g., emotional activation). By the dual process theory in which is postulated a primarily emotional function of the medial PFC (Greene et al., 2001), we would predict a predominant effect of tDCS on *footbridge*-like dilemmas, because the emotional activation is greater. However, following the "moral calculus" hypothesis (Moll et al., 2008a), modulating cortical excitability over medial prefrontal structures should mainly influence dilemmas that require deliberative responses and a cost-benefit analysis of a given action (i.e., *switch*-like dilemmas).

Our second goal was to explore the role of individual differences in deontology on the effect of tDCS on moral judgment. People with a strong deontological orientation should show less support for utilitarian-consistent solutions. However, past studies suggested that the effects of brain stimulation vary across individuals (Krause & Cohen Kadosh, 2014; Peña-Gómez, Vidal-Piñeiro, Clemente, Pascual-Leone, & Bartrés-Faz, 2011). Thus, we explored the potential role of deontology on the effects of brain stimulation on moral judgment.

Our final goal was to determine whether sex-related differences in utilitarian thinking underlie the responses to tDCS. Previous studies on moral judgment showed stronger effects of anodal tDCS for females (Fumagalli et al., 2010). Thus, we expect the strongest tDCS effect to occur on females' responses to our presented set of moral dilemmas.

<sup>1</sup> Previous studies adopted this technique to modulate the activity of the vmPFC (e.g., Chib, Yun, Takahashi, & Shimojo, 2013). However, with tDCS is not possible to selectively target vmPFC without including surrounding areas. Thus, in the current work, we refer to a larger portion of the prefrontal cortex, that is, the medial PFC.

## Methods

### Participants

The study participants consisted of 100 healthy university students (50 males;  $M_{\text{age}} = 24.68$ ,  $SD = 7.44$ ) with a negative history of medical disorders, substance abuse or dependence, use of central nervous system medications, and, in particular, psychiatric and neurological conditions, including brain surgery, tumor, or intracranial metal implantation (Poreisz, Boros, Antal, & Paulus, 2007). Considering the effect sizes obtained in our previous tDCS research (Riva, Gabbiadini, Lauro, Andrighetto, Volpato, & Bushman, 2017; Riva, Lauro, DeWall, & Bushman, 2012; Riva et al., 2014; Riva, Lauro, Vergallito, DeWall, Chester, & Bushman, 2015), *a priori* power analysis suggested a sample ranging from 73 ( $f = 0.33$ ) to 152 ( $f = 0.23$ ). Thus, the sample size of the current study ( $N = 100$ ) fell within this range.

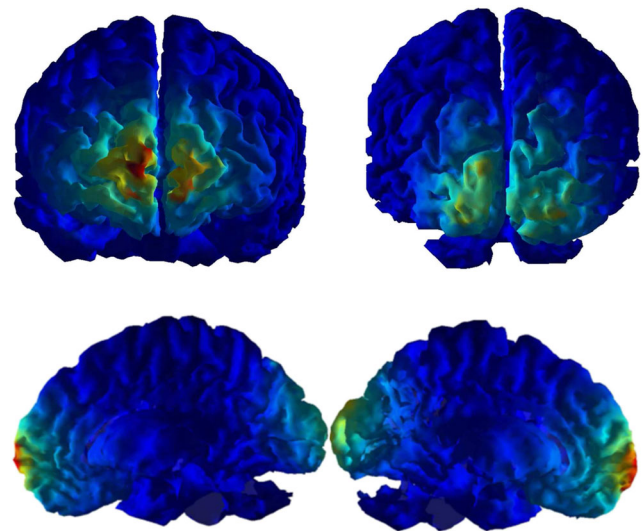
### Procedure

Participants were tested individually. After informed consent was obtained, participants completed eight items created ad hoc for this study (Appendix) that assessed individual differences in deontology. In contrast to utilitarianism, from a deontological perspective some choices cannot be justified by their effects: No matter how morally good are the consequences, some choices are morally forbidden. According to deontology, justifications should match principles that are universal and obeyed by each moral agent. In defining the eight-item scale, we focused on this attribute of universality. Example items from the scale are as follows: “*In general, I tend to make decisions consistently with the moral principles that a person must follow*” and “*In general, I tend to make decision thinking that there are absolute moral principles that apply to all situations* (1 = completely disagree to 7 = completely agree; alpha: 0.77 – see Appendix for the complete items list).

Next, participants were randomly assigned to receive anodal tDCS, cathodal tDCS or sham stimulation over the medial PFC. The tDCS device (DC-STIMULATOR, NeuroConn GmbH, Germany) included a study mode for a double-blind procedure. Namely, a numeric code, corresponding to either anodal, sham, or cathodal tDCS, input by the experimenter, started the stimulation, thus preventing awareness in both participants and experimenters of which stimulation condition was delivered. More specifically, the experimenter randomly extracted one of three codes for each participant. One code triggered anodal tDCS, another one triggered cathodal stimulation, and the last one triggered sham stimulation. This procedure led to a certain degree of variation between the number of participants assigned to the anodal (i.e.,  $N = 32$ ), cathodal (i.e.,  $N = 28$ ), and sham (i.e.,  $N = 40$ ) tDCS conditions.

Stimulation was applied using a constant current stimulator via sponge-soaked electrodes (DC-STIMULATOR, NeuroConn GmbH, Germany). The target electrode was  $9 \text{ cm}^2$  ( $3 \times 3 \text{ cm}$ ) and was placed between the nasion and FPZ (MNI coordinates: 2, 32, -10; Boorman, Rushworth, & Behrens, 2013), according to the international 10-20 system for EEG electrode placement. A  $25 \text{ cm}^2$  ( $5 \times 5 \text{ cm}$ ) reference electrode was placed over OZ. We used two differently sized electrodes to increase the focality of the stimulation. A constant current with an intensity of 0.75 mA was applied for 20 minutes. This provided a greater current density ( $0.08 \text{ mA/cm}^2$ ) for the stimulation electrode relative to the current density of the cephalic reference electrode ( $0.03 \text{ mA/cm}^2$ ; Nitsche et al., 2008). This electrode montage was modeled using Comets (COMputation of Electric field due to Transcranial current Stimulation, Jung et al., 2013). As shown in Fig. 1, considering our montage parameters, the peak of the electrical field occurred underneath the target electrode in an area corresponding to the medial prefrontal structures, including the frontopolar and ventromedial portion of PFC. For sham stimulation, the electrodes were placed in the same position, but the stimulator was turned on for only 30 s (Gandiga et al., 2006).

During the stimulation, participants were presented with 40 dilemmas on a computer screen. These dilemmas constituted a subset of a dataset that has been validated in Italian (Lotto et al., 2014). We selected 20 *switch-like* (i.e., killing one individual is a foreseen but unintended consequence of saving others) and 20 *footbridge-like* (i.e., killing one individual as



**Fig. 1** Computational model of the current flow related to our montage parameters showing the distribution of the electrical field. In the upper row, an anterior (on the left) and posterior (on the right) view of the brain is shown. The strongest electric field occurs around the cortical area underneath the target electrode. In the lower row, a lateral view the right (on the left) and left (on the right) hemisphere is illustrated. The peak of current flow is located in the medial PFC, affecting also the frontopolar region and ventromedial portion of medial PFC

an intended means to save others) dilemmas. Each of these two classes of dilemmas was varied for *self*-involvement. Thus, in 20 dilemmas, killing one individual resulted in saving one's own and other people's lives (*self*-involvement dilemmas), whereas in the other 20 dilemmas, killing one individual resulted in saving only other people (*other*-involvement dilemmas). The presentation order of the 40 dilemmas was randomized across subjects.

Each dilemma was presented in a series of three screens of text. The first screen described a scenario. The second screen described a possible action. The third screen posed a question related to the degree with which participants intended to implement the behavior described in the scenario ("To what extent would you implement this behavior?"; from 0 = *Not at all* to 7 = *Completely*). Higher ratings corresponded to higher levels of utilitarian responses. Participants were allowed to read through the screens and answer the questions at their own pace.

Sociodemographic information (Table 1), including sex, age, nationality, political orientation, and degree of religiosity, were collected, and then participants underwent a debriefing. During the debriefing, participants were asked whether they perceived any physical sensation from the electrodes.

## Analyses

Preliminary analyses were run to test: a) the equality of variances across stimulation groups of dilemma and involvement type; b) differences in age, gender, deontology, religiosity, and political orientation between the stimulation groups; c) potential differences in procedural sensations elicited by the stimulation protocol.

Concerning the primary analytical procedure, these were performed with the statistical program R (R Development Core Team, 2008). The considered dependent variable was the implementation rating subjects expressed for each dilemma. Data were submitted to a series of linear mixed effects models (Baayen et al., 2008), using the "lme4" R package (version 1.1-5, Bates, Maechler and Bolker, 2014). First, we tested whether the inclusion of a fixed effect or interaction contributed to the model goodness-of-fit. This was assessed

by looking at the goodness of fit measures, namely the likelihood ratio tests (LRT), AIC and BIC, including only effects that significantly increased the model's goodness-of-fit (Gelman & Hill, 2006). Critically, LRT significance and the information provided by the AIC and BIC were used to decide whether to include a parameter in the model. When the three indices were not in agreement, a decision was made according to the information provided by two of three parameters. Model selection results are reported in Table 2. As fixed factors, tDCS (3 levels: anodal vs. sham vs. cathodal), sex (2 levels: male vs. female), dilemma type (2 levels: *switch*-like vs. *footbridge*-like), and involvement type (2 levels: *self*-involvement dilemmas vs. *other*-involvement dilemmas), and their interactions were tested, together with participants' political orientation, religiosity, and baseline deontology as continuous independent variables. By subjects and by trial random intercepts were included, and the addition of random slopes for the fix effects included in the final model was tested as previously described. Results from the ANOVA on the final, best fitting model will be reported, with factors significance level based on Satterthwaite's degrees of freedom approximation in the "lmerTest" R package (version 2.0-29, Kuznetsova, Brockhoff, and Christensen, 2015).

## Results

### Preliminary analysis

**Equality of Variances** Levene's test indicated that the variances for dilemma type (*switch*-like vs. *footbridge*-like) and involvement type (*self*-involvement dilemmas vs. *other*-involvement dilemmas) were not different across the anodal, sham, and cathodal conditions,  $F_s(2,97) < 1.09$ ,  $p_s > 0.338$ .

**Age and sex differences** A between-subjects one-way ANOVA revealed that the mean age of those who received anodal stimulation was similar to those who received sham or cathodal stimulation,  $F(2,97) = 1.61$ ,  $p = 0.21$ . Moreover, the number of males and females did not differ across the experimental conditions,  $\chi^2(2) = 0.23$ ,  $p = 0.89$ .

**Table 1** Demographics

	<i>N</i>	Age (SD)	Females	Nationality (# of Italians)	Political orientation (from 1 to 7; higher values = right-wing)	Degree of religiosity (from 1 to 7; higher values = more religious)	Deontology (from 1 to 7; higher values = more deontological)
Anodal tDCS	32	24.78 (6.38)	15 (47%)	31 (97%)	3.19 (1.40)	2.53 (1.70)	4.53 (1.20)
Sham tDCS	40	25.98 (9.83)	21 (53%)	40 (100%)	3.18 (1.41)	2.45 (1.60)	4.44 (.92)
Cathodal tDCS	28	22.71 (3.16)	14 (50%)	25 (89%)	3.18 (1.39)	2.54 (1.58)	4.45 (.85)

**Table 2** Model selection procedure reporting the three goodness of fit indicators, i.e., AIC, BIC, and LRT, on utilitarian responses. LRT significance level was used to decide whether to include the parameter in the model, considering the information provided by the AIC and BIC.

When the indices were not in agreement, decision was made according to the information provided by two of three parameters. The table indicates whether each parameter was included or not and reports the resulting model for each step

Fix effects							
	$\chi^2$	df	p	AIC	BIC	Inclusion	Resulting model
Empty model				8918.6	8943.8		
tDCS	1,13	2	0,570	8921.5	8959.3		
Involvement type	1,97	1	0,160	8918.7	8950.1		
Dilemma type	53,50	1	<,001	8867.1	8898.6	*	Dilemma type
Deontology	30,44	1	<,001	8838.7	8876.5	*	Dilemma type + Deontology
Political orientation	0,57	1	0,450	8840.1	8884.2		
Religiosity	0,26	1	0,610	8840.4	8884.5		
Gender	8,83	1	0,003	8831.9	8875.9	*	Dilemma type + Deontology + Gender
tDCS * Gender	3,05	4	0,55	8836.8	8906.1		
tDCS * Involvement type	10,60	5	0,057	8831.0	8906.6		
tDCS * Dilemma type	30,26	4	<,001	8809.6	8874.8	*	Dilemma type + Deontology + Gender + tDCS * Dilemma type
tDCS * Deontology	36,88	6	<,001	8807.0	8868.8	*	Dilemma type + Deontology + Gender + tDCS * Dilemma type + tDCS * Deontology
Involvement type *	11,25	2	0,004	8799.7	8894.2	*	Dilemma type + Deontology + Gender + tDCS * Dilemma type + tDCS * Deontology + Involvement type * Dilemma type
Dilemma type							
Involvement type * Deontology	42,06	1	<,001	8759.7	8860.4	*	Dilemma type + Deontology + Gender + tDCS * Dilemma type + tDCS * Deontology + Involvement type * Dilemma type + Involvement type * Deontology
Gender * Deontology	0,11	1	0,74	8761.6	8868.6		
Gender * Dilemma type	3,2149		0,073	8757.5	8864.5		
Gender * Involvement type	0,93	1	0,33	8760.8	8867.8		
Dilemma type * Deontology	2,73	1	0,090	8759.0	8866.0		
tDCS * Involvement type * Dilemma type	2,44	4	0,650	8765.2	8891.1		
tDCS * Involvement type * Deontology	0,38	2	0,830	8763.3	8876.6		
tDCS * dilemma type * deontology	7,43	3	0,070	8758.3	8877.8		
Involvement type* Dilemma type * Deontology	3,03	2	0,220	8760.7	8874.0		
tDCS * involvement type * dilemma type * deontology	9,54	8	0,300	8766.1	8917.2		
Random effects slopes							
Subject							
Final fix effects model				8759.7	8860.4		tDCS + Involvement type + Dilemma type + Deontology + Gender + tDCS * Dilemma type + tDCS * Deontology + Involvement type * Dilemma type + Involvement type * Deontology
tDCS	NC						
Dilemma type	NC						
Involvement	88,19	3	<,001	8677.5	8797.1	*	(0 + Involvement type ID)
tDCS * Dilemma type	NC						
Involvement type * Dilemma type	173,4	7	<,001	8518.1	8681.8	*	(0 + Involvement type + Involvement type : Dilemma type ID)
Item							
tDCS	NC						
Deontology	0,001	1	0,99	8520.1	8690.1		
Gender	NC						
tDCS*Deontology	5,76	6	0,45	8524.4	8725.8		

**Deontology, religiosity, and political orientation** A between-subjects one-way ANOVA revealed that deontology scores did not differ among participants receiving anodal, sham,

and cathodal tDCS,  $F(2,97) = 0.09, p = 0.92$ . The same analysis showed that neither religiosity [ $F(2,97) = 0.03, p = 0.97$ ] nor scores related to political orientation [ $F(2,97) = 0.01, p =$

0.99] differ among participants receiving anodal, sham, and cathodal tDCS.

**Physical sensation from electrodes** In line with previous research (Nitsche et al., 2008), we found that only one participant (1/100) reported experiencing a physical sensation from the electrodes.

### Primary analyses

Table 2 shows the results of the model selection procedures. The final model included as fix effects the main effects of involvement type, dilemma type, sex, tDCS and deontology, as well as the dilemma type by tDCS, deontology by tDCS, involvement type by deontology, and involvement type by dilemma type interactions. The random structure included the by subject and by item random intercepts as well as the by subject random slopes for involvement type and involvement type by dilemma type (see Table 3 for the final model's parameters).

The final model showed no main effect of tDCS type,  $F(2,93.1) = 2.78, p = 0.07$ . However, we found the main effect of sex,  $F(1,93) = 5.12, p = 0.026$ . Male participants provided more utilitarian responses ( $M = 3.68, SD = 1.20$ ) than female participants ( $M = 3.05, SD = 0.93$ ). Furthermore, we found the main effect of dilemma type,  $F(1,53.17) = 109.02, p < 0.001$ . In particular, we found higher levels of utilitarian responses for *switch*-like dilemmas ( $M = 4.33, SD = 1.30$ ) than *footbridge*-like dilemmas ( $M = 2.40, SD = 1.11$ ). The analysis also showed the main effect of involvement type,  $F(1,124.8) = 20.61, p < 0.001$ . Participants provided more utilitarian

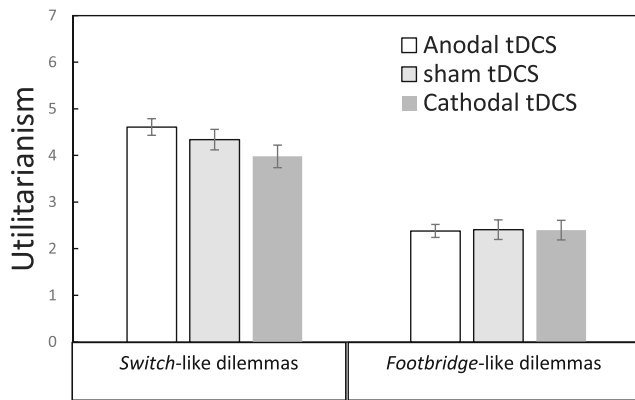
responses for personal involvement scenarios ( $M = 3.61, SD = 1.28$ ) than scenarios without personal involvement ( $M = 3.12, SD = 1.11$ ). Finally, deontology influenced participants' ratings,  $F(1,93) = 46.23; p < 0.001$ , with higher deontology associated with lower levels of utilitarian responses.

The tDCS type by dilemma type was significant,  $F(1,96.88) = 3.69, p = 0.028$ . As shown in Fig. 2, participants receiving anodal stimulation provided higher levels of utilitarian responses to *switch*-like dilemmas ( $M = 4.61, SD = 1.23$ ) compared with *footbridge*-like dilemmas ( $M = 2.38, SD = 1.08$ ) than those receiving cathodal stimulation (*switch*-like:  $M = 3.98, SD = 1.35$ ; *footbridge*-like:  $M = 2.40, SD = 1.24$ ;  $b = 0.29; t(96.9) = 2.71, p = 0.008$ ). In contrast, there were no differences between the sham (*switch*-like:  $M = 4.34, SD = 1.49$ ; *footbridge*-like:  $M = 2.41, SD = 1.44$ ) and the anodal stimulation groups,  $b = 0.11; t(96.9) = 1.19, p = 0.24$ , and between the cathodal and sham stimulation groups,  $b = 0.17; t(96.85) = 1.7, p = 0.09$ . The two-way interaction between dilemma type and involvement type was not significant,  $F(1,37.38) = 2.87, p = 0.09$ . However, the involvement type by deontology interaction was significant,  $F(1,98.1) = 13.65; p < 0.001$ . Specifically, Fig. 3 shows that higher levels of utilitarian responses were provided by those with low levels of baseline deontology when confronted with *self*-involvement dilemmas (vs. *other*-involvement ones). Finally, the tDCS by deontology interaction resulted significant,  $F(2,93) = 3.45; p = 0.036$ . The influence of deontology on moral choices was lower for the anodal condition compared with both the sham ( $b = -0.2; t(93) = -2.18; p = 0.032$ ) and the cathodal condition ( $b = -0.23; t(93) = -2.17; p = 0.033$ ), whereas no difference was present between the sham and the

**Table 3** Parameters of the final, best fitting model on utilitarian responses to moral dilemmas

	b	df	t	p
(Intercept)	0.95	116.97	3.06	0.003
tDCS: anodal vs. sham	0.82	96.18	1.89	0.062
tDCS: anodal vs. cathodal	0.79	95.84	1.57	0.120
tDCS: cathodal vs. sham	-0.03	95.39	-0.06	0.951
Dilemma type: <i>footbridge</i> vs. <i>switch</i> -like	-1.15	69.77	-8.85	<0.001
Involvement type: yes vs. no	0.71	132.05	3.54	0.001
Deontology	-0.08	100.72	-1.26	0.212
Sex: Female vs. male	-0.19	93	-2.26	0.026
Dilemma type: <i>footbridge</i> vs. <i>switch</i> -like * tDCS: anodal vs. sham	0.12	96.91	1.19	0.238
Dilemma type: <i>footbridge</i> vs. <i>switch</i> -like * tDCS: anodal vs. cathodal	0.29	96.9	2.71	0.008
Dilemma type: <i>footbridge</i> vs. <i>switch</i> -like * tDCS: cathodal vs. sham	0.17	96.85	1.70	0.092
tDCS: anodal vs. sham * Deontology	-0.20	93	-2.18	0.032
tDCS: anodal vs. cathodal * Deontology	-0.24	93	-2.17	0.033
tDCS: cathodal vs. sham * Deontology	-0.03	93	-0.28	0.782
Involvement type: yes vs. no * Dilemma type: <i>footbridge</i> vs. <i>switch</i> -like	0.26	37.38	1.69	0.099
Involvement type: yes vs. no * Deontology	-0.14	98.08	-3.69	<0.001





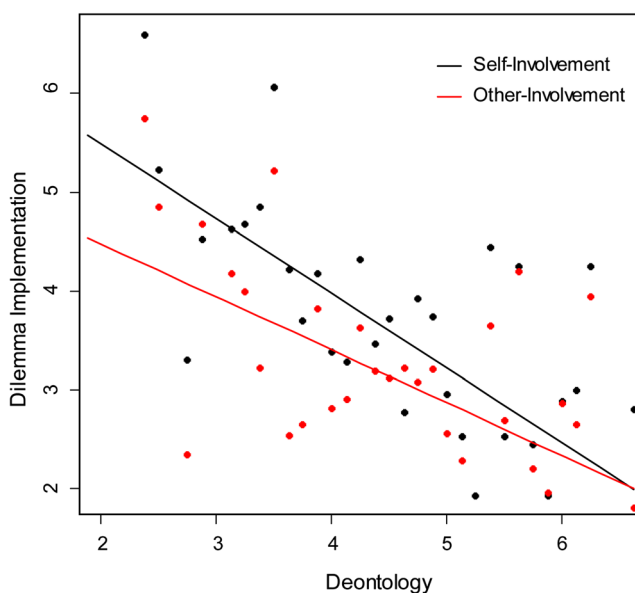
**Fig. 2** Utilitarian responses to switch-like and footbridge-like dilemmas for participants given anodal, sham, or cathodal stimulation. Capped vertical bars denote 1 *SE*

cathodal stimulation groups ( $b = -0.03$ ;  $t(93) = -0.28$ ;  $p = 0.78$ ; Fig. 4).

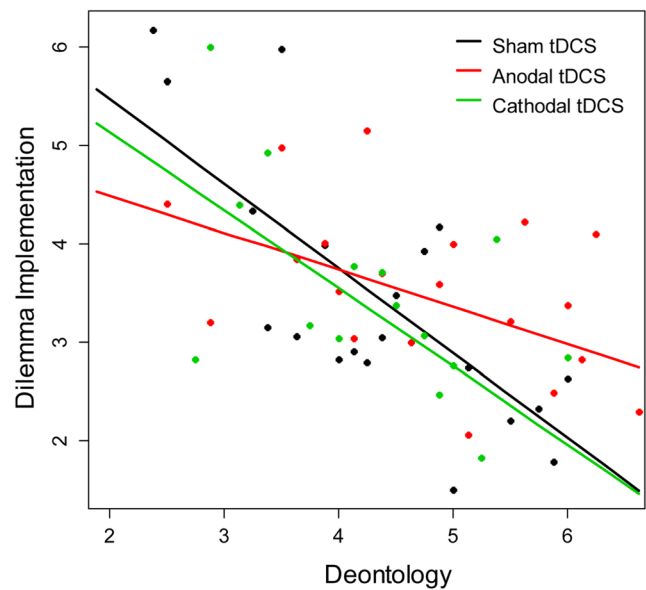
## Discussion

To investigate the functional contribution of the medial prefrontal structures in moral judgment, we manipulated the cortical excitability of this brain region using tDCS and observed the changes in responses to different categories of moral dilemmas. Elucidation of the effects of tDCS over the medial PFC can further our understanding of the role of this cortical region in moral judgment.

Our results showed that tDCS over medial PFC differently modulates the response to moral dilemmas in a fashion



**Fig. 3** Two-way interaction between deontology and dilemma type on utilitarian responses. For people with a weak deontological morality, self-preservation amplified the endorsement of a utilitarian resolutions in moral dilemmas



**Fig. 4** Two-way interaction between tDCS and deontology on utilitarian responses. Anodal tDCS (vs. sham and cathodal tDCS) significantly reduced the influence of people's baseline deontology on their moral choices

dependent upon the polarity of the stimulation, the types of dilemma, and individual differences in deontology. Previous research investigated the effect of tDCS over the VPC on moral judgments. Fumagalli et al. (2010) applied tDCS over the ventral prefrontal cortex (VPC) and found that cathodal tDCS reduced reaction times for utilitarian responses but did not affect the proportion of utilitarian responses. It decreased, albeit not significantly, utilitarian responses only in females. However, the results of this study are controversial. In particular, the authors collapsed different dilemma types (e.g., *switch-like*, *footbridge-like*, and non-moral) in the data analysis, thus excluding the possibility to determine whether tDCS reduced or increased utilitarian responses in different scenarios. Furthermore, this study adopted the original set of dilemmas used by Greene et al. (2001) while disregarding several criticisms that have been raised on this material. Specifically, several decisional scenarios were nondilemmas, because there were no conflicts between two actions or two obligations (McGuire, Langdon, Coltheart, & Mackenzie, 2009). To overcome such methodological limitations and to examine the role played by the medial PFC while individuals are facing various types of moral dilemmas, we adopted a more stable and standardized set of moral dilemmas with varying factors (Lotto et al., 2014).

In line with previous studies, we found that respondents provided higher levels of utilitarian responses when confronted with *switch-like* dilemmas, when *self-involvement* occurred, and when they were males. These findings replicate numerous studies showing that people react differently in *switch-like* versus *footbridge-like* scenarios (because of a combination of factors, such as emotional activation and

rules; Nichols & Mallon, 2006). The finding that people provided higher levels of utilitarian responses in *self*-involvement scenarios compared with *other*-involvement scenarios makes intuitive sense (e.g., self-preservation) and is in accordance with past research on this dimension (Lotto et al., 2014). We also found that for people with a weak deontological morality, self-preservation amplified the endorsement of a utilitarian resolutions in moral dilemmas. Conversely, for people with high levels of deontology, namely, people who generally tend to reject harm regardless of the outcomes of an action, the distinction between killing to save one's own life while saving others' lives (*self*-involvement dilemmas) and saving the lives of only other people (*other*-involvement dilemmas) did not matter. Indeed, highly deontological people tend to support the notion of "sanctity of life" that claims that human life is inherently valuable and precious, demanding respect both for *others* and for *oneself* (Singer, 1993). Finally, our finding that male participants provided more utilitarian responses than females also supports published studies on sex differences in moral reasoning (Friesdorf et al., 2015).

Moving beyond these behavioral effects, in contrast with previous research (Fumagalli et al. 2010), we found no interactions between tDCS and gender on moral choices. Thus, in our sample, tDCS effects did not differ between males and females. However, we found that anodal stimulation increased the endorsement of utilitarian solutions compared with cathodal stimulation. However, this effect was selective for only certain dilemma types, i.e., *switch*-like dilemmas but not *footbridge*-like dilemmas. Thus, compared with cathodal stimulation, anodal tDCS over medial prefrontal structures made people more likely to switch the track to save more lives at the expense of one, *but* it did not make people more likely to push a stranger down a bridge for the same effect. Moreover, while we had no a priori predictions regarding the direction of the relationship between deontology and tDCS, we found that anodal tDCS (vs. sham and cathodal tDCS) significantly reduced the influence of people's baseline deontology on their moral choices. The negative relationship between an individual's endorsement of deontological principles and willingness to endorse of utilitarian solutions has been shown before (Xu & Ma, 2015). However, our work suggests that increasing the cortical excitability over the medial prefrontal structures weakens the connection between ones' baseline deontological orientation and their support for utilitarian-consistent solutions. Thus, anodal stimulation seemed to reduce the influence of deontological principles in favor of a greater calculation on the beneficial costs of a given moral action.

To interpret these effects, we considered that *switch*-like dilemmas are characterized by a lower influence of the intuitive visceral emotional reaction linked with moral violation (Lotto et al., 2014). Indeed, to compare the relative weights

of several options (e.g., killing one to save five), one should be relatively free of the visceral influence of the emotional activation elicited by moral taboos. The dilemmas considered in this work that best allow deliberative responses that depend on calculation are *switch*-like dilemmas. Thus, these dilemmas, compared with *footbridge*-like dilemmas, might facilitate the implementation of deliberative responses. Our results support the view that the medial PFC network is involved in moral calculus (Moll et al., 2005; 2008a). In contrast, our results do not fit with the dual-system hypothesis (Greene, 2014), which predicts that the medial brain structures code mostly for the emotional value of a moral dilemma. The dual-process theory predicts that modulating cortical excitability of the automatic/intuitive emotional system would result in stronger effects on *footbridge*-type dilemmas. At odds with this latter hypothesis, our findings suggest that neuromodulation over the medial PFC affects dilemmas that allow for a cognitive evaluation of outcomes rather than those that are mainly shaped by affect-laden reactions. However, more recently, Shenva and Greene (2014) provided a slightly different explanation of the role of the vmPFC in moral judgment. They found that vmPFC is preferentially engaged when emotional responses and explicit rule-based reasoning must be integrated to form an overall moral acceptability judgment. These findings support the hypothesis that vmPFC serve as a locus of integration for the two principal modes of evaluation in moral judgment: deontological judgments supported by emotional evaluation and utilitarian judgments concerning the consequences of an action or behavior (Shenva & Greene, 2014). Furthermore, the activity in this region (and in the mOFC) reflects a subjective evaluation process that is sensitive to expected value regarding gains and losses, including gains and losses of life (Shenva & Greene, 2010). In this updated view, the vmPFC seems to have lost its primary reactive role of "alarm bell" that generates affective responses to behavioral options (Greene et al., 2001, 2004; Greene, 2008; Greene, 2014)—a role that now seems prerogative of the amygdala—to assume, instead, the role of a "hub" in which disparate value signals are integrated to obtain a more abstract and comprehensive value representation (Shenva & Greene, 2010). As we found that anodal (vs. cathodal) stimulation increased the endorsement of utilitarian responses only for *switch*-like dilemmas, which are usually considered as more "cognitive" and based on cost/benefit trade-off (Green et al., 2001; Green, 2014; Hauser et al., 2007), we may assume that the anodal stimulation can enhance vmPFC sensitiveness to the expected value representation (Shenva & Greene, 2010). Conversely, we may explain the lack of increased utilitarian responses in *footbridge*-like dilemmas considering that these dilemmas are less based on cost/benefit calculation being more emotionally driven. However, it is important to point out that in Shenva & Greene (2010) the BOLD signal for expected moral value of decision options, affects the vmPFC but also the mOFC that is

a much wider network than vmPFC alone, and in some ways more similar to the vmPFC/FPC network proposed by Moll et al. (2005; 2008a).

Notably, as it is well known, tDCS holds a low spatial resolution, hence preventing us from the possibility to selectively stimulate the mPFC. Accordingly, our results should be interpreted as pointing to the involvement of a broader PFC area, namely medial PFC, which includes the ventromedial and the frontopolar portion, as suggested by the computational model of the current flow shown in Fig. 1. Recent studies, mainly using neuroimaging techniques and clinical observation of brain-lesioned patients (see Ramnani & Owen, 2004, for a review) have probed the involvement of the frontopolar portion of PFC, roughly corresponding to the most anterior part of Brodmann area 10, in several “high-level” cognitive tasks. For instance, Christoff and Gabrieli (2000) claimed that this region is specialized for the explicit processing of “internal” information, including one’s thoughts and feelings. According to Koechlin et al. (1999, 2000), the frontopolar region underpins what they called “cognitive branching,” namely the “ability to hold in mind goals while exploring and processing secondary goals” (1999, p. 148). Using Raven’s Progressive Matrices, Kroger et al. (2002) probed the engagement of the frontopolar region in the simultaneous consideration and integration of multiple relations. With an integrative approach concerning these alternative hypotheses, recently Ramnani and Owen (2004) proposed that the key role of the frontopolar region regards dealing with problems implying the coordination and integrations of multiple cognitive operations. All of these putative roles attributed to the frontopolar region of PFC are likely involved in the high cognitive processes underlying the problem solving, reasoning, planning, and decision making on abstract information implied in moral dilemma. Similarly, an activation of the most anterior portion of PFC was reported when a decision had to be made by resolving a choice between two independent probability judgments (Rogers et al., 1999).

Another possible factor that might account for our findings refers to the degree of consensus that people have on different types of dilemmas. For instance, responders are usually more confident in refusing the action for the *footbridge*-like dilemmas than accepting the action for the *switch*-like dilemmas (Hauser et al., 2007). Different degrees of consensus might derive from the existence of moral rules prescribing (e.g., self-preservation) or prohibiting (e.g., one should not kill) the action. Thus, emotion-based explanations should be integrated with the interpretation of moral rules when interpreting people’s reaction to moral dilemmas (Nichols & Mallon, 2006).

There are some main limitations of this study. The first limitation is the already-mentioned low spatial focality of the tDCS, which means that a broad brain region was affected by our tDCS manipulation. However, tDCS effects largely come

from the cortical area beneath the electrode (Zaghi, Acar, Hultgren, Boggio, & Fregni, 2010), and as indicated by the finite elements methods (FEM) based modeling of the current flow related to our montage parameters (Fig. 1), we can assume that changes in the cortical excitability over the medial PFC occurred because the strongest electric field was estimated in that region. Crucially for our study, although the area affected by tDCS encompassed the frontopolar and ventromedial portions of PFC, it clearly set apart from the dorsolateral region. The functional specificity of tDCS at a neurophysiological level has been recently proven (Pisoni et al., 2017), showing that only task-related areas are likely to be modulated by the effects of electrical stimulation. Given the previous literature on medial PFC and moral judgments, it is possible that the behavioral effects reported in the present study are due to modulations of this area. Nevertheless, a functional influence of fronto-polar cortex stimulation cannot be excluded.

Second, our randomization procedure led to some variations in the number of participants assigned to each of the three tDCS conditions. However, the analytical approach we used is typically considered robust against unbalanced designs (Cnaan et al., 1997; Baayen et al., 2008). Random effects are individual-specific effects modelled as coming from a common distribution (usually a normal distribution). Hence, unlike linear models that do not include an individual-level estimation, intercepts are allowed to take different values for each subject. Mixed-effect models work well for balanced as well as unbalanced data sets because estimates for individual effects (intercepts and slopes) are weighted by sample size. These models, thus, handle unbalanced data by simultaneously estimating both the effects and the variance components in a more efficient way compared with fixed effects-only models, as multiple linear regression (Searle, 1988; Burton et al., 1998; Peretz et al., 2002).

Third, although we found that in the sham group participants’ decision were more likely to be influenced by their deontology compared with the anodal tDCS group, in other cases, neither anodal nor cathodal tDCS was statistically different from sham. In this regard, we note that the purpose of our study was not to provide clinical indications about the possibility of designing interventions to modulate moral judgment, for which, larger effects and a difference between cathodal (or anodal) and sham would have been the key finding. Rather, our study speaks about the possible involvement of the medial PFC in different types of moral dilemmas. Still, obtaining a different polarity dependent modulation effect suggests that the medial PFC is involved in decision making in the context of moral dilemmas for *switch*-like dilemmas.

Finally, this study was implemented to test two alternative accounts of the functional role of medial PFC on moral judgment. However, as we noted, different dimensions might account for the interactions we found, including moral rules (Nichols & Mallon, 2006) and different degrees of

dilemmaticity for specific types of dilemmas (Brink, 1994). For instance, a higher degree of dilemmaticity might call for a more deliberate elaboration of action outcomes. Intentionality is another dimension that might account for the effect we found. According to Doctrine of Double Effect (Aquinas; trans. 1947), harm is more acceptable when it is a foreseen but unintended consequence than when it is an intended means to an end. Thus, in the *switch*-like dilemmas the harm (i.e., killing one person) can be considered as a side effect of the implementation of utilitarian behavior (i.e., saving five persons), whereas the *footbridge*-like dilemmas require the decision to directly kill a person as a means to save more people (Lotto et al., 2014; Manfrinati, Lotto, Sarlo, Palomba, & Rumiati, 2013). Therefore, future studies are needed to determine the exact mechanisms (e.g., emotions, rules, or an integration of both) involved in the interactive effects identified in this work.

## Conclusions

Our study adds further evidence for the role of medial prefrontal structures in moral decision making. However, our study extends prior knowledge on the functional role of this brain area by showing that altering the cortical excitability over the medial PFC affects the individual's tendency for moral judgment for certain types of moral dilemmas, namely, *switch*-like dilemmas. Furthermore, the present study explored the modulatory role of medial PFC cortical excitability on the link between individual differences in deontology and moral choices. Specifically, we found that anodal stimulation (as compared to the sham or the cathodal group) reduced the influence of people's baseline levels of deontology on their moral choices. These findings are consistent with a moral calculus account of the medial PFC (Moll et al., 2005; 2008a). They also nicely complement previous neuromodulatory data showing that the DLPFC is not solely involved in cognitive control over emotional impulses in the context of moral dilemmas (Tassy et al., 2012). Finally, our results highlight new avenues of research into the key role of individual differences on brain stimulation-induced changes in moral judgment.

## Appendix - Deontology scale

Instructions: Please express your degree of agreement with each of the following statements by choosing the option that best represents your opinion (1 = complete disagreement; 7 = complete agreement).

1. In general, I tend to make decisions consistent with the moral principles a person must follow.
2. In general, I tend to make decisions based on a moral duty.
3. In general, I tend to make decisions excluding morally unacceptable alternatives.
4. In general, I tend to make decisions recalling that certain behaviors are just right or wrong, no matter what the consequences.
5. In general, I tend to make decisions thinking that there are absolute moral principles that apply to all situations.
6. In general, I tend to make decisions thinking that a person can never be used as a means of reaching an end.
7. In general, I tend to make decisions thinking that if one of my behavior damages another person, it should never be done even if this would bring benefits to a greater number of people.
8. In general, I tend to make decisions considering it vital to never sacrifice the dignity of others even when this would bring benefits to a greater number of people.

## References

- Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: The medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, *7*, 268–277.
- Aquinas, T. (1947). *Summa theologiae*. New York: Benzinger Brothers (Originally published in 1265–1272).
- Atran, S. (2002). In Gods We Trust: The Evolutionary Landscape of Religion. New York: Oxford University Press.
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412.
- Baron, J., & Spranca, M. (1997). Protected values. *Organizational Behavior and Human Decision Processes*, *70*, 1–16.
- Bartels, D. M. (2008). Principled moral sentiment and the flexibility of moral judgment and decision making. *Cognition*, *108*, 381–417.
- Bartels, D. M., & Pizarro, D. A. (2011). The mismeasure of morals: Antisocial personality traits predict utilitarian responses to moral dilemmas. *Cognition*, *121*, 154–161.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4*. R package version, 1, 1–23.
- Boorman, E. D., Rushworth, M. F., & Behrens, T. E. (2013). Ventromedial prefrontal and anterior cingulate cortex adopt choice and default reference frames during sequential multi-alternative choice. *The Journal of Neuroscience*, *33*, 2242–2253.
- Borg, J. S., Hynes, C., Van Horn, J., Grafton, S., & Sinnott-Armstrong, W. (2006). Consequences, action, and intention as factors in moral judgments: An fMRI investigation. *Journal of Cognitive Neuroscience*, *18*, 803–817.
- Boyer, P. (2003). Religious thought and behaviour as by-products of brain function. *Trends in Cognitive Sciences*, *7*, 119–124.
- Brink, D. O. (1994). Moral conflict and its structure. *The Philosophical Review*, *103*, 215–247.
- Burton, P., Gurrin, L., & Sly, P. (1998). Tutorial in biostatistics. Extending the simple linear regression model to account for correlated responses: an introduction to generalized estimating equations and multi-level mixed modeling. *Statistics in Medicine*, *17*, 1261–1291.

- Chib, V. S., Yun, K., Takahashi, H., & Shimojo, S. (2013). Noninvasive remote activation of the ventral midbrain by transcranial direct current stimulation of prefrontal cortex. *Translational Psychiatry*, 3, e268.
- Christoff, K., & Gabrieli, J. D. E. (2000). The frontopolar cortex and human cognition: evidence for a rostrocaudal hierarchical organisation within the human prefrontal cortex. *Psychobiology*, 28, 168–186.
- Ciamarelli, E., Braghittoni, D., & di Pellegrino, G. (2012). It is the outcome that counts! Damage to the ventromedial prefrontal cortex disrupts the integration of outcome and belief information for moral judgment. *Journal of the International Neuropsychological Society*, 18, 962–971.
- Ciamarelli, E., Muccioli, M., Ladavas, E., & di Pellegrino, G. (2007). Selective deficit in personal moral judgment following damage to ventromedial prefrontal cortex. *Social Cognitive and Affective Neuroscience*, 2, 84–92.
- Cnaan, A., Laird, N. M., & Slasor, P. (1997). Tutorial in biostatistics: Using the general linear mixed model to analyse unbalanced repeated measures and longitudinal data. *Statistics in Medicine*, 16, 2349–2380.
- Damasio, A. (1994). *Descartes' error: Emotion, reason and the human brain*. New York: Avon Books.
- Damasio, A. R., Tranel, D., & Damasio, H. (1990). Individuals with sociopathic behavior caused by frontal damage fail to respond autonomically to social stimuli. *Behavioural Brain Research*, 41, 81–94.
- Feltz, A., & Cokely, E. T. (2008). The fragmented folk: More evidence of stable individual differences in moral judgments and folk intuitions. In *Proceedings of the 30th annual conference of the Cognitive Science Society*. 1771–1776.
- Foot, P. (1967). The Problem of Abortion and the Doctrine of Double Effect. *Oxford Review*, 5, 5–15.
- Forbes, C. E., & Grafman, J. (2010). The Role of the Human Prefrontal Cortex in Social Cognition and Moral Judgment. *Annual Review of Neuroscience*, 33, 299–324.
- Friedsdorf, R., Conway, P., & Gawronski, B. (2015). Gender differences in responses to moral dilemmas: A process dissociation analysis. *Personality and Social Psychology Bulletin*, 41, 696–713.
- Fumagalli, M., & Priori, A. (2012). Functional and clinical neuroanatomy of morality. *Brain*, 135, 2006–2021.
- Fumagalli, M., Vergari, M., Pasqualetti, P., Marceglia, S., Mameli, F., Ferrucci, R., ..., Barbieri, S. (2010). Brain switches utilitarian behavior: Does gender make the difference?. *PLoS One*, 5, e8865.
- Gandiga, P. C., Hummel, F. C., & Cohen, L. G. (2006). Transcranial DC stimulation (tDCS): A tool for double-blind sham-controlled clinical studies in brain stimulation. *Clinical Neurophysiology*, 117, 845–850.
- Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge university press.
- Gilligan, C. (1982). *In a different voice*. Harvard University Press.
- Greene, J. (2014). *Moral tribes: Emotion, reason and the gap between us and them*. Atlantic Books Ltd.
- Greene, J.D., Nystrom, L.E., Engell, A.D., Darley, J.M., & Cohen, J.D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron*, 44, 389–400.
- Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, 293, 2105–2108.
- Greene, J. D. (2008). The secret joke of Kant's soul. In W. Sinnott-Armstrong (Ed.), *Moral Psychology* (Vol. 3). Cambridge: MIT Press.
- Haidt, J., & Graham, J. (2007). When morality opposes justice: Conservatives have moral intuitions that liberals may not recognize. *Social Justice Research*, 20, 98–116.
- Harenski, C. L., Antonenko, O., Shane, M. S., & Kiehl, K. A. (2008). Gender differences in neural mechanisms underlying moral sensitivity. *Social Cognitive and Affective Neuroscience*, 3, 313–321.
- Hauser, M. (2006). *Moral Minds: How Nature Designed Our Universal Sense of Right and Wrong*. New York: HarperCollins.
- Hauser, M., Cushman, F., Young, L., Kang-Xing Jin, R., & Mikhail, J. (2007). A dissociation between moral judgments and justifications. *Mind & Language*, 22, 1–21.
- Jaffee, S., & Hyde, J. S. (2000). Gender differences in moral orientation: A meta-analysis. *Psychological bulletin*, 126, 703–726.
- Jung, Y. J., Kim, J. H., & Im, C. H. (2013). COMETS: a MATLAB toolbox for simulating local electric fields generated by transcranial direct current stimulation (tDCS). *Biomedical Engineering Letters*, 3, 39–46.
- Kant I. (1785/1959). In: *Foundation of the metaphysics of morals*. Beck LW, translator. Indianapolis: Bobbs-Merrill.
- Koechlin, E., Basso, G., Pietrini, P., Panzer, S. & Grafman, J. (1999). The role of the anterior prefrontal cortex in human cognition. *Nature*, 399, 148–151.
- Koechlin, E., Corrado, G., Pietrini, P. & Grafman, J. (2000). Dissociating the role of the medial and lateral anterior prefrontal cortex in human planning. *Proceedings of the National Academy of Sciences*, 97, 7651–7656.
- Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M., & Damasio, A. (2007). Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature*, 446, 908–911.
- Krause, B., & Cohen Kadosh, R. (2014). Not all brains are created equal: The relevance of individual differences in responsiveness to transcranial electrical stimulation. *Frontiers in Systems Neuroscience*, 8, 25.
- Kroger, J. K. et al., (2002). Recruitment of anterior dorsolateral prefrontal cortex in human reasoning: a parametric study of relational complexity. *Cerebral Cortex*, 12, 477–485.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2015). *Package 'lmerTest'*. R package version, 2(0).
- Lakoff, G. (2002). *Moral politics: How liberals and conservatives think*. Chicago: University of Chicago Press.
- Lind, G. (2008). The meaning and measurement of moral judgment revisited: A dual aspect model. In: Fasko D, Willis W, (eds). *Contemporary Philosophical Perspectives on Moral Development and Education*. Creskill: Hampton press.
- Lotto, L., Manfrinati, A., & Sarlo, M. (2014). A new set of moral dilemmas: Norms for moral acceptability, decision times, and emotional salience. *Journal of Behavioral Decision Making*, 27, 57–65.
- Manfrinati, A., Lotto, L., Sarlo, M., Palomba, D., & Rumiati, R. (2013). Moral dilemmas and moral principles: When emotion and cognition unite. *Cognition & Emotion*, 27, 1276–1291.
- McGuire, J., Langdon, R., Coltheart, M., & Mackenzie, C. (2009). A reanalysis of the personal/impersonal distinction in moral psychology research. *Journal of Experimental Social Psychology*, 45, 577–580.
- Mikhail, J. (2007). Universal moral grammar: Theory, evidence and the future. *Trends in Cognitive Sciences*, 11, 143–152.
- Mikhail, J. (2011). *Elements of moral cognition: Rawls' linguistic analogy and the cognitive science of moral and legal judgment*. Cambridge University Press.
- Moll, J., de Oliveira-Souza, R., & Zahn, R. (2008a). The neural basis of moral cognition. *Annals of the New York Academy of Sciences*, 1124, 161–180.
- Moll, J., de Oliveira-Souza, R., Zahn, R., Grafman, J. (2008b). The cognitive neuroscience of moral emotions. In: Sinnott-Armstrong W, ed. *Moral psychology: The Neuroscience of Morality: Emotion, Brain Disorders, and Development* (pp. 1–17). Vol 3. Cambridge, MA: The MIT Press.
- Moll, J., & de Oliveira-Souza, R. (2007). Moral judgments, emotions and the utilitarian brain. *Trends in cognitive sciences*, 11, 319–321.

- Moll, J., de Oliveira-Souza, R., & Eslinger, P. J. (2003). Morals and the human brain: a working model. *Neuroreport*, *14*, 299–305.
- Moll, J., Zahn, R., de Oliveira-Souza, R., Krueger, F., & Grafman, J. (2005). The neural basis of human moral cognition. *Nature Reviews Neuroscience*, *6*, 799–809.
- Moore, A. B., Clark, B. A., & Kane, M. J. (2008). Who shalt not kill? Individual differences in working memory capacity, executive control, and moral judgment. *Psychological Science*, *19*, 549–557.
- Nichols, S., & Mallon, R. (2006). Moral dilemmas and moral rules. *Cognition*, *100*, 530–542.
- Nitsche, M. A., Cohen, L. G., Wassermann, E. M., Priori, A., Lang, N., Antal, A., ... & Pascual-Leone, A. (2008). Transcranial direct current stimulation: State of the art 2008. *Brain stimulation*, *1*, 206–223.
- Peña-Gómez, C., Vidal-Piñeiro, D., Clemente, I. C., Pascual-Leone, Á., & Barrés-Faz, D. (2011). Down-regulation of negative emotional processing by transcranial direct current stimulation: Effects of personality characteristics. *PLoS one*, *6*, e22812.
- Peretz, C., Goren, A., Smid, T., & Kromhout, H. (2002). Application of mixed-effects models for exposure assessment. *Annals of Occupational Hygiene*, *46*, 69–77.
- Pessoa, L. (2013). *The cognitive-emotional brain: From interactions to integration*. Cambridge: MIT press.
- Pisoni, A., Mattavelli, G., Papagno, C., Rosanova, M., Casali, A. G., & Romero Lauro, L. J. (2017). Cognitive Enhancement Induced by Anodal tDCS Drives Circuit-Specific Cortical Plasticity. *Cerebral Cortex*, 1–9.
- Poreisz, C., Boros, K., Antal, A., & Paulus, W. (2007). Safety aspects of transcranial direct current stimulation concerning healthy subjects and patients. *Brain Research Bulletin*, *72*, 208–214.
- Prehn, K., Wartenburger, I., Mériaux, K., Scheibe, C., Goodenough, O. R., et al. (2007). Individual differences in moral judgment competence influence neural correlates of socio-normative judgments. *Social Cognitive and Affective Neuroscience*, *3*, 33–46.
- R Development Core Team. (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org>. Accessed 30 Aug 2018.
- Ramnani, N., & Owen, A. M. (2004). Anterior prefrontal cortex: insights into function from anatomy and neuroimaging. *Nature Reviews Neuroscience*, *5*, 184.
- Riva, P., Gabbiadini, A., Lauro, L. J. R., Andrighetto, L., Volpato, C., & Bushman, B. J. (2017). Neuromodulation can reduce aggressive behavior elicited by violent video games. *Cognitive, Affective, & Behavioral Neuroscience*, *17*, 452–459.
- Riva, P., Lauro, L. J. R., DeWall, C. N., & Bushman, B. J. (2012). Buffer the pain away: Stimulating the rVLPFC reduces pain following social exclusion. *Psychological Science*, *23*, 1473–1475.
- Riva, P., Lauro, L. J. R., DeWall, C. N., Chester, D. S., & Bushman, B. J. (2014). Reducing aggressive responses to social exclusion using transcranial direct current stimulation (tDCS). *Social Cognitive and Affective Neuroscience*, *10*, 352–356.
- Riva, P., Lauro, L. J. R., Vergallito, A., DeWall, C. N., Chester, D. S., & Bushman, B. J. (2015). Electrified emotions: Modulatory effects of transcranial direct stimulation on the negative emotional reaction to social exclusion. *Social Neuroscience*, *10*, 46–54.
- Rogers, R. D. et al. (1999). Choosing between small, likely rewards and large, unlikely rewards activates inferior and orbital prefrontal cortex. *Journal of Neuroscience*, *19*, 9029–9038.
- Sacchi, S., Riva, P., Brambilla, M., & Grasso, M. (2014). Moral reasoning and climate change mitigation: The deontological reaction toward the market-based approach. *Journal of Environmental Psychology*, *38*, 252–261.
- Searle, S. R. (1988). Mixed models and unbalanced data: where from, where at and where to? *Communications in Statistics-Theory and Methods*, *17*, 935–968.
- Shenhav, A., & Greene, J. D. (2010). Moral judgments recruit domain-general valuation mechanisms to integrate representations of probability and magnitude. *Neuron*, *67*, 667–677.
- Singer, P. (1993). *Practical Ethics*. Cambridge: Cambridge University Press.
- Talmi, D., & Frith, C. (2007). Neurobiology: feeling right about doing right. *Nature*, *446*, 865–866.
- Tassy, S., Oullier, O., Duclos, Y., Coulon, O., Mancini, J., Deruelle, C., ..., Wicker, B. (2012). Disrupting the right prefrontal cortex alters moral judgement. *Social Cognitive and Affective Neuroscience*, *7*, 282–288.
- Thomson, J. J. (1986). *Rights, restitution, and risk: Essays in moral theory*. Cambridge: Harvard University Press.
- Tranel, D., Bechara, A., & Denburg, N. L. (2002). Asymmetric functional roles of right and left ventromedial prefrontal cortices in social conduct, decision-making, and emotional processing. *Cortex*, *38*, 589–612.
- Xu, Z. X., & Ma, H. K. (2015). Does honesty result from moral will or moral grace? Why moral identity matters. *Journal of Business Ethics*, *127*, 371–384.
- Zaghi, S., Acar, M., Hultgren, B., Boggio, P. S., & Fregni, F. (2010). Noninvasive brain stimulation with low-intensity electrical currents: Putative mechanisms of action for direct and alternating current stimulation. *The Neuroscientist*, *16*, 285–307.