Cogn Affect Behav Neurosci (2016) 16:626–634
DOI 10.3758/s13415-016-0419-2

Interfering with activity in the dorsomedial prefrontal cortex via TMS affects social impressions updating

Chiara Ferrari[1] · **Tomaso Vecchi**[1,2] · **Alexander Todorov**[3] · **Zaira Cattaneo**[2,4]

Abstract In our everyday social interactions we often need to deal with others' unpredictable behaviors. Integrating unexpected information in a consistent representation of another agent is a cognitively demanding process. Several neuroimaging studies point to the medial prefrontal cortex (mPFC) as a critical structure in mediating social evaluations. Our aim here was to shed light on the possible causal role of the mPFC in the dynamic process of forming and updating social impressions about others. We addressed this issue by suppressing activity in the mPFC by means of 1 Hz offline transcranial magnetic stimulation (TMS) prior to a task requiring participants to evaluate other agents' trustworthiness after reading about their social behavior. In two different experiments, we found that inhibiting activity in the mPFC increased perceived trustworthiness when inconsistent information about one agent's behavior was provided. In turn, when only negative or positive behaviors of a person were described, TMS over the mPFC did not affect judgments. Our results indicate that the mPFC is causally involved in mediating social impressions updating—at least in cases in which judgment is uncertain due to conflicting information to be processed.

✉ Zaira Cattaneo
zaira.cattaneo@unimib.it

[1] Department of Brain and Behavioral Sciences, University of Pavia, Pavia 27100, Italy

[2] Brain Connectivity Center, National Neurological Institute C. Mondino, Pavia 27100, Italy

[3] Department of Psychology, Princeton University, Princeton, NJ 08544, USA

[4] Department of Psychology, University of Milano–Bicocca, Piazza dell'Ateneo Nuovo 1, 20126 Milano, Italy

Keywords Medial prefrontal cortex · Social impressions · Transcranial magnetic stimulation · Social judgments · Trustworthiness

When interacting in social contexts, individuals continuously generate impressions about other agents and expectations about their possible behavior, often on the basis of very limited amount of information (Todorov & Uleman, 2003). When expectations about others are violated, the social impressions of them need to be updated (e.g., Hamilton, Driscoll, & Worth, 1989; Hastie & Kumar, 1979; Reeder & Coovert, 1986). In these situations, individuals usually take longer to integrate information about others that contradicts rather than matches their initial impressions (Reeder & Coovert, 1986). Indeed, integrating new inconsistent information in preexisting schemata is cognitively demanding, likely tapping on executive functions (Macrae, Bodenhausen, Schloerscheidt, & Milne, 1999; Payne, 2005).

Neuroimaging and brain stimulation evidence suggests that forming social impressions about others and/or judging their trustworthiness involves the medial prefrontal cortex (mPFC) (Cattaneo, Mattavelli, Platania, & Papagno, 2011; Cloutier, Gabrieli, O'Young, & Ambady, 2011; Ferrari et al., 2016; Ma et al., 2012; Mattavelli, Cattaneo, & Papagno, 2011; Mende-Siedlecki, Baron, & Todorov, 2013; Mende-Siedlecki, Cai, & Todorov 2012; Schiller, Freeman, Mitchell, Uleman, & Phelps, 2009). Interestingly, this region is also involved in processing socially relevant emotions (such as arrogance or guilt) beyond basic emotions (Jankowski & Takahashi, 2014; see also D'Agata et al., 2011). Accordingly, alterations in the functioning or structural abnormalities of the medial prefrontal cortices have been associated with abnormal biases in social evaluation that characterize psychiatric disorders, such as schizophrenia (Brüne, 2005; Pia & Tamietto, 2006;Yamada et al., 2007) or depression (Foland-Ross et al., 2014; Holmes et al., 2012; Thoma, Norra, Juckel, Suchan, & Bellebaum, 2015). Moreover, the mPFC seems to be particularly sensitive to violations of expectations in updating social impressions: For instance, when presented with consistent (either morally good or morally bad) or inconsistent behaviors of an agent, dorsal sectors of the mPFC (dmPFC) preferentially activate when evaluating behaviors that contradict the initially formed impression about the agent (Ma et al., 2012; Mende-Siedlecki et al., 2012).

In this study we used TMS to shed light on the role of the dmPFC in evaluating other agents' trustworthiness when presented with either consistent or inconsistent information about their social behavior. Since forming and updating impressions about other agents is a sequential process that requires integrating different pieces of information over time, we used an offline TMS paradigm in which activity in the dmPFC was suppressed before asking participants to evaluate a person's trustworthiness on the basis of reading verbal descriptions of the person's behavior. Interfering with dmPFC activity via TMS should affect the formation and updating of social impressions, especially when conflicting information about one's behavior needs to be integrated (Ma et al., 2012; Mende-Siedlecki et al., 2012).

Experiment 1

Method

Participants

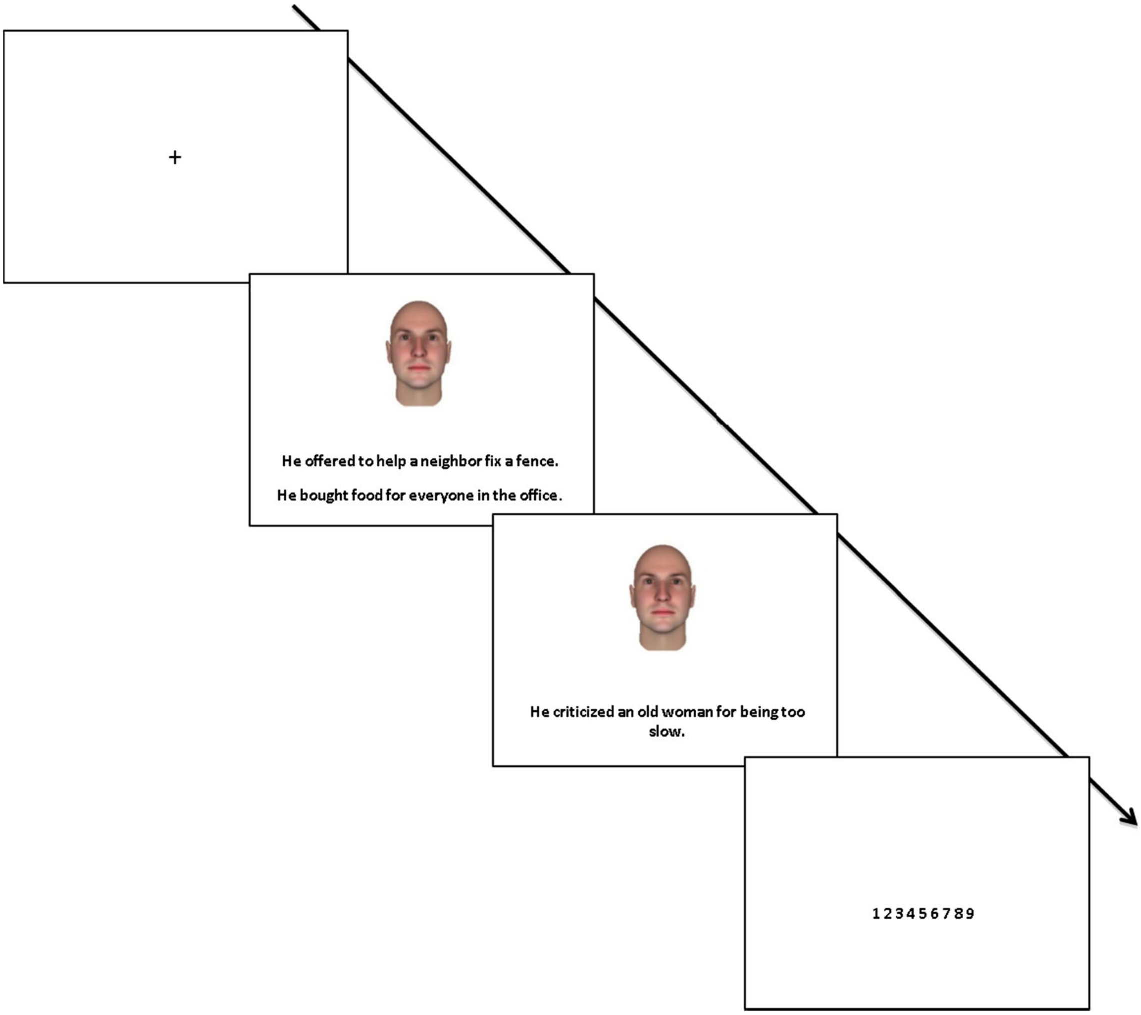
Twenty right-handed (Oldfield, 1971) Italian students (10 males, mean age = 24.1 years, *SD* = 2.0) participated in the experiment. Prior to the TMS experiment, each participant filled in a questionnaire (Rossi, Hallett, Rossini, & Pascual-Leone, 2011) to evaluate compatibility with TMS. None of the participants reported neurological problems or history of seizures. None was taking medications that could interfere with neuronal excitability. Written informed consent was obtained from all participants before the experiment. The protocol was approved by the local ethical committee, and participants were treated in accordance with the Declaration of Helsinki.

Stimuli and procedure

Participants were seated comfortably at a distance of 57 cm from a 17-in. (1,024 × 768) TFT-LCD computer monitor and wore earplugs to minimize TMS click sound interference. Experimental stimuli consisted of male faces (each measuring 7 × 7 of visual angle) and written sentences (white ink, 12-point Courier New font). We selected the face stimuli from a larger database (http://tlab.princeton.edu/databases/trustworthinessfaces2, see Oosterhof & Todorov, 2008) that included seven computer-generated variations along the trustworthiness dimension (i.e., falling 1, 2, 3 standard deviations below or above the original neutral version of each face) for 25 different Caucasian male identities. From this set, we chose 15 different-identity faces with average trustworthiness (i.e., neutral faces, coded as trustworthiness level "0" in the original database). Hence, faces were all similar (i.e., all neutral) along the trustworthiness trait. Sentences were adapted from Mende-Siedlecki et al. (2012). Each sentence described the behavior of a male individual in a particular situation. Half of the sentences described a good/socially valuable behavior (e.g., *He gave out toys to the Children's Hospital at Christmas*) and half described a bad/socially questionable behavior (e.g., *He told a colleague in public that she should lose weight*). Sentences referred to "ordinary" positive or negative behaviors, with no reference to extremely bad acts (such as murders) or to heroic gestures. A total of 45 sentences describing positive behaviors and 45 sentences describing negative behaviors were created.

Figure 1 shows the timeline of an experimental trial. Each trial started with a fixation cross appearing in the middle of the screen for 500 ms. A face was then presented in the middle of the screen together with two sentences appearing below it describing two behaviors that were either both positive or negative. Participants were instructed to (silently) read the sentences and form an impression of the person depicted in the picture, and press the space-bar key (with the left hand) when ready. A third sentence was then presented, with the same face still visible in the middle of the screen, that described a behavior that could be either of the same valence of the previous two (congruent condition) or of opposite valence (incongruent condition). Participants were instructed to update the impression they had just formed, integrating the additional information, and to press the space-bar key (with the left hand) when they were ready. After this, participants rated the person on a Likert scale, ranging from 1 (*not trustworthy at all*) to 9 (*very trustworthy*) by pressing (with the right hand) the corresponding number on the keyboard. The next trial followed their response.

Each TMS block (see below) consisted of 60 trials (half starting with positive and half with negative behaviors described). The behaviors were of the same valence in the first two sentences of both congruent and incongruent trials. The order of trials was randomized within each experimental block and for each participant. Participants performed the task at self-pace but were encouraged to be fast. Participants performed four practice trials before receiving TMS (see below) to familiarize with the task. Moreover, all sentences were presented once in a random order before TMS was given to ensure that participants had knowledge of the type of actions described and of their "morality" range, and could adjust their rating criterion accordingly before the experiment (Palmer, Schloss, & Sammartino, 2013). The experimental task took



He offered to help a neighbor fix a fence.

He bought food for everyone in the office.

He criticized an old woman for being too slow.

Fig. 1 The timeline of an experimental trial. Participants read about either two positive or two negative social behaviors of an agent whose face was also simultaneously presented. Then, a third behavior was presented that could be either consistent (in valence) or inconsistent with the first two. After reading about the third behavior, participants were asked to evaluate the trustworthiness of the person (i.e., output of the impression formation and updating process) on a 1 to 9 Likert scale. Offline 1 Hz TMS was delivered for 15 minutes before the beginning of of the task to suppress activity in the dmPFC

approximately 7 minutes and was performed immediately after the end of the TMS stimulation (see below).

Transcranial magnetic stimulation (TMS)

Offline neuronavigated 1 Hz TMS was administered over the dmPFC via a Magstim Rapid^2 stimulator (Magstim Co Ltd, Whitland, UK) connected to a 70-mm butterfly coil at a fixed intensity of 50% of the maximum stimulator output for 15 minutes. Similar stimulation parameters (i.e., fixed intensity of stimulation, 15 minutes of stimulation) have been used before to suppress activity of prefrontal regions prior to task requiring social judgments (e.g., reciprocal fairness, moral judgments), with the effects of stimulation continuing after the end of the actual stimulation (e.g., Baumgartner, Knoch, Hotz, Eisenegger, & Fehr, 2011; Eisenegger, Treyer, Fehr, & Knoch, 2008; Knoch et al., 2006; Knoch, Pascual-Leone, Meyer, Treyer, & Fehr, 2006; Tassy et al., 2012). Talairach coordinates (Talairach & Tournoux, 1988) for the dmPFC were $x = 1.5$, $y = 31.5$, and $z = 35.5$; these coordinates were taken from previous neuroimaging work demonstrating an activation of this region during social impression updating (Mende-Siedlecki et al., 2012).

The target location corresponding to the dmPFC was identified on each subject's scalp using the SofTaxic navigator system (E.M.S., Bologna, Italy). The procedure involves the computation of an estimated volume of head MRIs in participants for whom individual MRIs are unavailable (i.e., all participants of our study). The estimated MRIs, referred to the Talairach space, are calculated by means of a warping procedure, operating on a template MRI volume on the basis of a set of around 60 points digitized from the participant's scalp by means of a Polaris Vicra Optical Tracking System (Northern Digital, Inc., Waterloo, ON, Canada). The digitized points are used to compute a subsequent set of reference points that are analogous to a set of points prelocalized on the scalp of the template. The warping procedure is performed using these two corresponding sets of reference points. This procedure has been proven to ensure a global localization accuracy of roughly 5 mm (Carducci & Brusco, 2012), and it has been successfully used in many previous TMS studies (Capotosto, Babiloni, Romani, & Corbetta, 2009; Jacquet & Avenanti, 2015; Renzi et al., 2013; Urgesi, Berlucchi, & Aglioti, 2004). The coil was placed tangentially to the scalp with the handle pointing backward and held parallel to the midsagittal line. Participants underwent both a real TMS and a sham TMS session. In the sham condition, the same stimulation parameters were used but the coil position was tilted 90° (e.g., Zanto, Rubens, Thangavel, & Gazzaley, 2011). Sham and real stimulation were performed in two separated sessions on different days (intermixed by a minimum of 2 days and a maximum of 3 days). The order of the TMS condition (sham vs. real) was counterbalanced across participants.

Results

The dependent variables were mean trustworthiness (1–9 Likert scale) scores and mean response times (RT; ms). Trials in which individual RT (as recorded from onset of the response slide) were more than 3 standard deviations from the participant's mean performance in each block were removed from the analysis (a total of 1.6% trials were excluded).

For each dependent variable, we carried out a repeated-measures ANOVA with TMS (sham vs. real), valence of the first impression (positive vs. negative, as conveyed by the first two sentences), and congruence (i.e., whether the final information conveyed by the third sentence was in line with the previous two) as within-subjects factors, and participants' gender and order of TMS sessions (real first vs. sham first) as between-subjects factors. The ANOVA on trustworthiness scores (see Fig. 2) revealed no significant main effect of either TMS, $F(1, 16) = 1.92$, $p = .18$, participants' gender, $F(1, 16) < 1$, $p =. 48$, or session order, $F(1, 16) < 1$, $p = .47$. The main effects of valence of the first impression, $F(1, 16) = 175.14$, $p < .001$, $\eta_p^2 = .92$, and of congruence, $F(1, 16) = 19.31$, $p < .001$, $\eta_p^2 = .55$, were significant, as well as their interaction, $F(1, 16) = 422.16$, $p < .001$, $\eta_p^2 = .96$. The interaction TMS by congruence was also significant, $F(1, 16) = 11.65$, $p = .004$, $\eta_p^2 = .42$. No other interactions reached significance (*p*s > .09). The main effects of valence and congruence were analyzed in light of their significant interaction. In congruent trials, negative statements (as expected) lowered perceived trustworthiness of the face compared to positive statements, $t(19) = 21.36$, $p < .001$ (Bonferroni-Holm correction applied). In incongruent trials, order of presentation of positive and negative information did not impact on face trustworthiness rating: Face trustworthiness scores were similar when negative information was followed by positive one and when positive information was presented first, $t(19) < 1$, $p = .73$. The interaction TMS by congruence was further analyzed via post hoc comparisons that showed that TMS affected face trustworthiness evaluation in incongruent trials, $t(19) = 2.97$, $p = .032$ (Bonferroni-Holm correction applied), but not in congruent trials, $t(19) < 1$, $p = .61$. As shown in Fig. 2, when incongruent information was provided, participants rated faces as more trustworthy following real rather than sham TMS.

Moreover, to ensure that TMS effects (that are known to fade over time) covered the entire task, we split participants' responses in two halves, and we repeated the same analysis taking into account whether the TMS effects we reported were different when considering the first 30 trials versus the latest 30

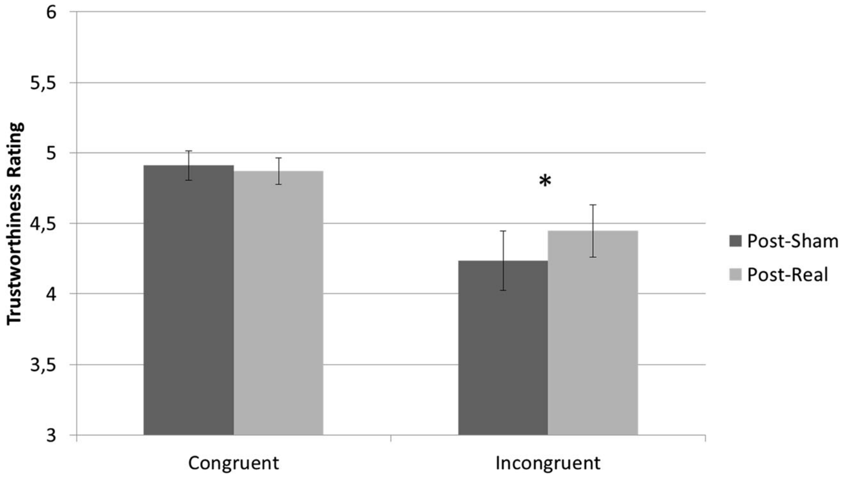


Fig. 2 Mean participants' trustworthiness rating scores as a function of TMS (sham vs. real) and congruence of the behaviors described (congruent = both positive or negative vs. incongruent). Error bars represent ±1 *SEM*. The asterisk indicates a significant difference between sham and real TMS: suppression of the dmPFC by real TMS resulted into significantly more positive evaluations compared to sham TMS when conflicting information was provided (incongruent trials)

trials. The analysis revealed that the timing of the response (first half vs. latest half of the test block) was not significant, $F(1, 16) = 2.85$, $p = .11$, neither it interacted with any other variable ($ps > .09$), confirming that the suppressive effect of the stimulation persisted for the entire task duration (cf. Thut & Pascual-Leone, 2010).

The ANOVA on mean RT revealed no significant main effect of TMS, $F(1, 16) < 1$, $p= .88$, and no significant main effect of valence, $F(1, 16) < 1$, $p = .42$. The main effect of congruence was significant, $F(1, 16) = 15.75$, $p = .001$, $\eta_p^2 = .50$, indicating that participants took longer to judge faces associated with inconsistent information (mean RT = 405 ms) than those associated with consistent information (mean RT= 366 ms). The between-subjects variables gender, $F(1, 16) = 2.59$, $p = .13$, and session order, $F(1, 16) < 1$, $p = .43$, were not significant. None of the interactions were significant (all $ps > .08$). The RT we analyzed were recorded from the *onset* of the response slide in which only the Likert scale was presented; however, for the way the experiment was designed (self-paced reading of information), it is likely that participants came up with their judgment and prepared to respond before moving to the response slide. Even when considering cumulative trial time, TMS effects were not significant (main effect of TMS: $p = .66$, cumulative trial time following real TMS: 6,088 ms, following sham TMS: 6,016 ms; interaction TMS by congruence, $p = .84$).

Experiment 2

The results of Experiment 1 suggest that TMS on the dmPFC affected the integration of inconsistent information, increasing the perceived trustworthiness of the agent. Experiment 2 was carried out to ensure that the effects of dmPFC TMS on social impression updating could be replicated in another group of participants, and also when modulating response uncertainty by varying the amount of consistent and inconsistent information provided.

Method

Participants

Fourteen Italian students (three males, mean age= 24.0 years, *SD* = 1.6) participated in the experiment. None of them had participated in Experiment 1. Inclusion criteria were the same as for Experiment 1.

Stimuli, procedure, and TMS

Stimuli and procedure were similar to those used in Experiment 1, but participants were only presented with descriptions of inconsistent behaviors. Moreover, we varied the amount of the information provided, including a condition in which only two behaviors rather than three were described. Each block contained 60 trials, 15 for each condition (i.e., 2 positive + 1 negative; 2 negative + 1 positive; 1 positive + 1 negative; 1 negative + 1 positive). To familiarize with stimuli and the procedure, participants read all the sentences in random order before TMS was given and performed four practice trials. The task took approximately 7 minutes and started immediately after the end of the stimulation. TMS parameters were identical to Experiment 1.

Results

The dependent variables were computed as in Experiment 1. A total of 1.29% trials were excluded due to RT (as recorded from onset of the response slide) falling more than 3 *SD* from the participant's mean RT in each block.

Mean trustworthiness scores are reported in Fig. 3. A repeated measure-ANOVA with TMS (sham vs. real), task

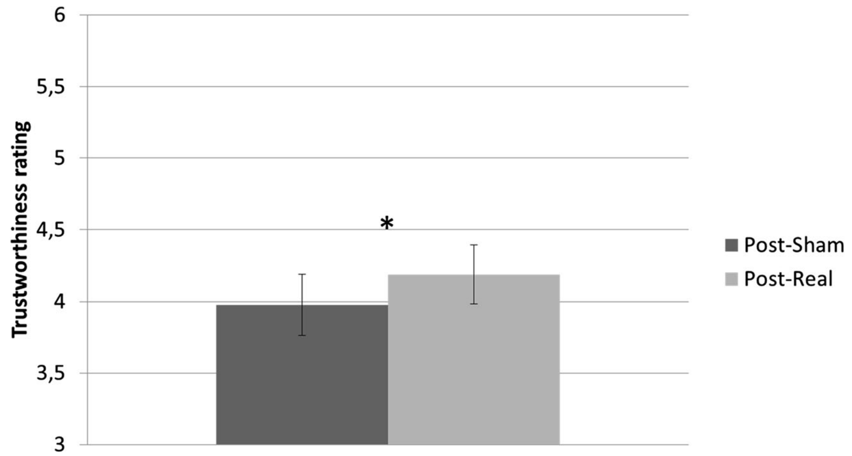


Fig. 3 Mean participants' trustworthiness rating scores after reading description of one agent's inconsistent behaviors in Experiment 2. Error bars represent ±1 *SEM*. The asterisk indicates a significant difference between sham and real TMS: Participants rated social agents as more trustworthy following real compared to sham TMS over the dmPFC, replicating the findings of Experiment 1

condition (first impression based on one vs. two consistent behaviors, to be updated with a following inconsistent behavior), valence of the first impression (positive vs. negative) as within-subjects variables, and order of TMS sessions as between-subjects variable, was carried out for each dependent variable (participants' gender was not further considered in light of lack of gender differences reported in the previous experiment). The main effect of TMS was significant, $F(1, 12) = 7.16$, $p = .020$, $\eta_p^2 = .37$, indicating that participants perceived faces as more trustworthy following real rather than sham TMS (see Fig. 3). The main effect of task condition, $F(1, 12) < 1$, $p = .81$, and the main effect of valence, $F(1, 12) = 1.15$, $p = .24$, were not significant. The interaction task condition by valence was significant, $F(1, 12) = 12.08$, $p = .005$, $\eta_p^2 = .52$, due to evaluations being more polarized (either toward the positive or the negative side) when the first impression was based on two consistent behaviors compared to trials in which a single positive or negative behavior (first impression) had to be updated with an opposite-valenced one. The main effect of session order was not significant, $F(1, 12) < 1$, $p = .92$. No other interactions reached significance ($ps > .24$).

As in case of Experiment 1, we also repeated the same analysis splitting the trials in the first half versus latest half, to investigate whether the effects of TMS faded away over time or persisted till the end of the test phase. The analysis revealed that the timing of the response (first half vs. latest half of the test block) was not significant, $F(1, 12) < 1$, $p = .61$, neither it interacted with any other variable ($ps > .17$), confirming that the suppressive effect of the stimulation persisted for the entire task duration.

The ANOVA on mean RT (from onset of the response slide) revealed no significant main effects of task condition ($p = .17$), valence ($p = .99$), TMS ($p = .69$, mean RT following sham TMS = 448 ms; following real TMS = 429 ms), and session order ($p = .69$). None of the interactions reached significance ($ps > .10$). The cumulative trial time (from onset of the first slide) following real TMS was 5,587 ms and following sham TMS was 5,658 ms (TMS did not significantly affect cumulative trial RT, $p = .77$).

Discussion

Our findings suggest a causal role of the dorsomedial prefrontal cortex (dmPFC) in the dynamic process of updating social impressions about others, adding to prior neuroimaging evidence showing an involvement of this region in social impression formation and updating (Bhanji & Beer, 2013; Cloutier et al., 2011; Ma et al., 2012; Mende-Siedlecki et al., 2012, 2013; Schiller et al., 2009). In particular, in two experiments we presented participants with descriptions of other agents' behaviors that could be either positively or negatively valenced. We found that inhibiting activity in the dmPFC via TMS compared to control sham stimulation resulted in more positive evaluations of other individuals when inconsistent information was provided. No effect of TMS was found when information about the individual's behavior was entirely consistent (only positive or only negative behaviors; see Experiment 1).

Evaluations were similar whether the first impression was based on one or two behaviors (Experiment 2), indicating that the "updating" worked similarly, regardless of the strength of the initial impression. Overall, the final evaluations on inconsistent trials converged around the midpoint of the scale in both experiments. Together with longer response times for these trials (Experiment 1), this suggests that participants were more uncertain about their trustworthiness decisions when dealing with incongruent information and preferred to give "neutral" judgments. In turn, when the provided descriptions were only negative or only positive, participants' responses were quite polarized. In this case, TMS had no effect in line with prior evidence showing that TMS is more effective in modulating responses in uncertain conditions (Robertson, Theoret, & Pascual-Leone, 2003). Moreover, a prior fMRI study adopting a task similar to the one employed here found enhanced activity in the dmPFC only when inconsistent information had to be integrated (Mende-Siedlecki et al., 2012), suggesting that the dmPFC may be more critical in updating social impressions when conflicting information has to be processed. Also, in that work the dmPFC preferentially responded to inconsistent information regardless the "direction" of the impression updating (from positive to negative or vice versa; Mende-Siedlecki et al., 2012). This is also consistent with our finding that TMS similarly affected evaluation when inconsistent behaviors were described, regardless whether the first impression formed was positive or negative.

When response uncertainty was higher (inconsistent trials), suppressing activation in the dmPFC biased evaluation toward a more positive output. Although TMS in our study did not selectively affect the "weight" of positive or negative behaviors in determining the final evaluation, the positive bias induced by stimulation seems to be in line with prior findings showing that 20 minutes of 1 Hz suppressive TMS over the medial PFC resulted into a bias toward positive emotional stimuli (Schutter & Van Honk, 2006). Also, it has been suggested that the mPFC may be particularly sensitive to violations of morality and social rules (e.g., Fiddick, Spampinato, & Grafman, 2005; Takahashi et al., 2008). Moreover, it is worth mentioning that psychiatric disorders such as schizophrenia, in which mPFC dysfunctions have been observed (Yamada et al., 2007), are often associated with abnormal social evaluations. In particular, schizophrenic patients have been reported to trust unfamiliar faces more than healthy controls (e.g., Baas, Van't Wout, Aleman, & Kahn, 2008; McIntosh & Park, 2014), but also to show abnormal anchoring to prior information, especially when negative-valenced (e.g., Hooker et al., 2011).

In a previous work (Ferrari et al., 2016), we found that online TMS over the dmPFC delayed fast dichotomous (yes/no) responses when participants had to decide whether a face–adjective pair (for instance, a face accompanied by the adjective "selfish") matched the impression they had formed about that agent by reading a description of his behavior. In that work, it was the target stimulus to be "congruent" or "incongruent" with the impression formed, and response accuracy could be measured because the trait–adjective was clearly either in line or in contrast with the behavior described (hence, not surprisingly, accuracy was very high, above 90%). In that study, TMS mainly affected RT, as it is typically the case when accuracies are near ceiling and fast responses are required (Devlin & Watkins, 2008). Moreover, participants in the baseline control condition were faster in responding when the face–adjective pair matched the impression formed, showing that some priming mechanisms were at play. TMS selectively delayed decisions in primed (congruent) trials, according to state-dependent views on the effects of brain stimulation (Cattaneo, Rota, Vecchi, & Silvanto, 2008; Silvanto & Cattaneo, 2014). In the study presented here, the target stimulus was a neutral face, which cannot therefore be defined as congruent or incongruent with the agent's behavior (there were no "correct" responses). In turn, what we measured was whether face trustworthiness decisions (on a 1–9 Likert scale) could be biased by previous knowledge about the agent's behavior. We found that TMS over the dmPFC significantly affected evaluation of face trustworthiness, in particular, when the available information about the agent behavior was inconsistent, including both negative and positive actions. In these instances, participants were more uncertain about their final judgment and their responses were thus more permeable to the effects of stimulation (when all the behaviors described were either positive or negative, responses were more "polarized" and hence less vulnerable to TMS interference; see Robertson et al., 2003). In light of the measurement Likert scale we employed, it is also not surprisingly that RT were not a sensitive measure in our task, since participants had to express their judgment by pressing one out of nine keys, and they probably came up with a final evaluation upon reading (at self-pace, without time pressure) the latest description of the agent's behavior, before moving to the response slide.

Finally, it is important to consider that TMS can modulate activity not only in the neurons under the coil but also in interconnected regions (e.g., Avenanti, Annella, Candidi, Urgesi, & Aglioti, 2013; Siebner, Hartwigsen, Kassuba, & Rothwell, 2009). The amygdala may be particular important here, in light of converging patients and fMRI data suggesting that it is involved in face trustworthiness evaluation (e.g., Adolphs, Tranel, & Damasio, 1998; Baron, Gobbini, Engell, & Todorov, 2011; Todorov & Olson, 2008). Indeed, social impressions are likely to be based on a first perceptual stage in which facial features are analyzed and on a further processing stage in which face appearance is integrated with information stored in memory about that agent's behavior (Rudoy & Paller, 2009). While the amygdala and/or other cortical and subcortical structures could be more relevant in the analysis of face appearance (Tamietto et al., 2005), the dmPFC is likely to intervene at a later stage, combining face appearance with available information about the agent's behavior (Baron et al., 2011; see also Costa et al., 2013). Accordingly, it has been suggested that the dmPFC works as a convergence area for face and behavioral information, interacting with the amygdala's signals (Baron et al., 2011; Kim et al., 2004). In light of this, we cannot exclude that TMS had indirectly affected the amygdala as well as other cortical or subcortical regions, such as the anterior cingulate cortex (important in conflict monitoring; see Botvinick, Braver, Barch, Carter, & Cohen, 2001), and the orbitofrontal cortex (important in processing of positive/rewarding stimuli, e.g., Blair et al., 2013; O'Doherty, Kringelbach, Rolls, Hornak, & Andrews, 2001; Rolls, 2000).

In sum, our study demonstrates that the dmPFC is causally involved in the dynamic process of updating social impressions: when its activity is suppressed, participants tend to be more positive in judging other individuals. Our findings may shed light on the possible role of the dmPFC in mediating abnormal social evaluation in certain psychiatric disorders (such as schizophrenia), thus providing evidence potentially important for the design of clinical treatments employing brain stimulation (see Freitas, Fregni, & Pascual-Leone, 2009).

Acknowledgments This work was supported by a Fund for Investments on Basic Research (FIRB), Italian Ministry of Education, University and Research (RBFR12F0BD) to Z.C.

References

Adolphs, R., Tranel, D., & Damasio, A. R. (1998). The human amygdala in social judgment. *Nature, 393,* 470–474.

Avenanti, A., Annella, L., Candidi, M., Urgesi, C., & Aglioti, S. M. (2013). Compensatory plasticity in the action observation network: Virtual lesions of STS enhance anticipatory simulation of seen actions. *Cerebral Cortex, 23,* 570–580.

Baas, D., Van't Wout, M., Aleman, A., & Kahn, R. S. (2008). Social judgement in clinically stable patients with schizophrenia and healthy relatives: behavioural evidence of social brain dysfunction. *Psychological Medicine, 38,* 747–754.

Baron, S. G., Gobbini, M. I., Engell, A. D., & Todorov, A. (2011). Amygdala and dorsomedial prefrontal cortex responses to appearance-based and behavior-based person impressions. *Social Cognitive and Affective Neuroscience, 6,* 572–81.

Baumgartner, T., Knoch, D., Hotz, P., Eisenegger, C., & Fehr, E. (2011). Dorsolateral and ventromedial prefrontal cortex orchestrate normative choice. *Nature Neuroscience, 14,* 1468–1474.

Bhanji, J. P., & Beer, J. S. (2013). Dissociable neural modulation underlying lasting first impressions, changing your mind for the better, and changing it for the worse. *Journal of Neuroscience, 33,* 9337–9344.

Blair, K. S., Otero, M., Teng, C., Jacobs, M., Odenheimer, S., Pine, D. S., & Blair, R. J. R. (2013). Dissociable roles of ventromedial prefrontal cortex (vmPFC) and rostral anterior cingulate cortex (rACC) in value representation and optimistic bias. *NeuroImage, 78,* 103–110.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Reviews, 108,* 624–652.

Brüne, M. (2005). Emotion recognition, 'theory of mind', and social behavior in schizophrenia. *Psychiatry Research, 133,* 135–147.

Capotosto, P., Babiloni, C., Romani, G. L., & Corbetta, M. (2009). Frontoparietal cortex controls spatial attention through modulation of anticipatory alpha rhythms. *Journal of Neuroscience, 29,* 5863–5872.

Carducci, F., & Brusco, R. (2012). Accuracy of an individualized MR-based head model for navigated brain stimulation. *Psychiatry Research, 203,* 105–108.

Cattaneo, Z., Mattavelli, G., Platania, E., & Papagno, C. (2011). The role of the prefrontal cortex in controlling gender-stereotypical associations: A TMS investigation. *NeuroImage, 56,* 1839–1846.

Cattaneo, Z., Rota, F., Vecchi, T., & Silvanto, J. (2008). Using state-dependency of transcranial magnetic stimulation (TMS) to investigate letter selectivity in the left posterior parietal cortex: A comparison of TMS-priming and TMS-adaptation paradigms. *European Journal of Neuroscience, 28,* 1924–1929.

Cloutier, J., Gabrieli, J. D. E., O'Young, D., & Ambady, N. (2011). An fMRI study of violations of social expectations: When people are not who we expect them to be. *NeuroImage, 57,* 583–588.

Costa, T., Cauda, F., Crini, M., Tatu, M. K., Celeghin, A., de Gelder, B., & Tamietto, M. (2013). Temporal and spatial neural dynamics in the perception of basic emotions from complex scenes. *Social Cognitive and Affective Neuroscience, 9,* 1690–1703.

D'Agata, F., Caroppo, P., Baudino, B., Caglio, M., Croce, M., Bergui, M., Tamietto, M., Mortara, P., & Orsi, L. (2011). The recognition of facial emotions in spinocerebellar ataxia patients. *The Cerebellum, 10,* 600–610.

Devlin, J. T., & Watkins, K. E. (2008). Investigating language organization with TMS. *The Oxford handbook of transcranial stimulation, 479,* 499.

Eisenegger, C., Treyer, V., Fehr, E., & Knoch, D. (2008). Time-course of "off-line" prefrontal rTMS effects—A PET study. *NeuroImage, 42,* 379–384.

Ferrari, C., Lega, C., Vernice, M., Tamietto, M., Mende-Siedlecki, P., Vecchi, T., Todorov, A., & Cattaneo, Z. (2016). The dorsomedial prefrontal cortex plays a causal role in integrating social impressions from faces and verbal descriptions. *Cerebral Cortex, 26,* 156–165.

Fiddick, L., Spampinato, M. V., & Grafman, J. (2005). Social contracts and precautions activate different neurological systems: An fMRI investigation of deontic reasoning. *NeuroImage, 28,* 778–786.

Foland-Ross, L. C., Hamilton, J. P., Sacchet, M. D., Furman, D. J., Sherdell, L., & Gotlib, I. H. (2014). Activation of the medial prefrontal and posterior cingulate cortex during encoding of negative material predicts symptom worsening in major depression. *NeuroReport, 25,* 324–329.

Freitas, C., Fregni, F., & Pascual-Leone, A. (2009). Meta-analysis of the effects of repetitive transcranial magnetic stimulation (rTMS) on negative and positive symptoms in schizophrenia. *Schizophrenia Research, 108,* 11–24.

Hamilton, D. L., Driscoll, D. M., & Worth, L. T. (1989). Cognitive organization of impressions: Effects of incongruency in complex representations. *Journal of Personality and Social Psychology, 57,* 925–939.

Hastie, R., & Kumar, P. A. (1979). Person memory: Personality traits as organizing principles in memory for behaviors. *Journal of Personality and Social Psychology, 37,* 25–38.

Holmes, A. J., Lee, P. H., Hollinshead, M. O., Bakst, L., Roffman, J. L., Smoller, J. W., & Buckner, R. L. (2012). Individual differences in amygdala-medial prefrontal anatomy link negative affect, impaired social functioning, and polygenic depression risk. *Journal of Neuroscience, 32,* 18087–18100.

Hooker, C. I., Tully, L. M., Verosky, S. C., Fisher, M., Holland, C., & Vinogradov, S. (2011). Can I trust you? Negative affective priming influences social judgments in schizophrenia. *Journal of Abnormal Psychology, 120,* 98–107.

Jacquet, P. O., & Avenanti, A. (2015). Perturbing the action observation network during perception and categorization of actions' goals and grips: State-dependency and virtual lesion TMS effects. *Cerebral Cortex, 25,* 598–608.

Jankowski, K. F., & Takahashi, H. (2014). Cognitive neuroscience of social emotions and implications for psychopathology: Examining embarrassment, guilt, envy, and schadenfreude. *Psychiatry and Clinical Neurosciences, 68,* 319–336.

Kim, H. Y., Somerville, L. H., Johnstone, T., Polis, S., Alexander, A., Shin, L. M., & Whalen, P. J. (2004). Contextual modulation of amygdala responsivity to surprised faces. *Journal of Cognitive Neuroscience, 16,* 1730–1745.

Knoch, D., Gianotti, L. R., Pascual-Leone, A., Treyer, V., Regard, M., Hohmann, M., & Brugger, P. (2006). Disruption of right prefrontal cortex by low-frequency repetitive transcranial magnetic stimulation induces risk-taking behavior. *Journal of Neuroscience, 26,* 6469–6472.

Knoch, D., Pascual-Leone, A., Meyer, K., Treyer, V., & Fehr, E. (2006). Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science, 314*(5800), 829–832.

Ma, N., Vandekerckhove, M., Baetens, K., Van Overwalle, F., Seurinck, R., & Fias, W. (2012). Inconsistencies in spontaneous and intentional trait inferences. *Social Cognitive and Affective Neuroscience, 7,* 937–950.

Macrae, C. N., Bodenhausen, G. V., Schloerscheidt, A. M., & Milne, A. B. (1999). Tales of the unexpected: Executive function and person perception. *Journal of Personality and Social Psychology, 76,* 200–213.

Mattavelli, G., Cattaneo, Z., & Papagno, C. (2011). Transcranial magnetic stimulation of medial prefrontal cortex modulates face expressions processing in a priming task. *Neuropsychologia, 49,* 992–998.

McIntosh, L. G., & Park, S. (2014). Social trait judgment and affect recognition from static faces and video vignettes in schizophrenia. *Schizophrenia Research, 158,* 170–175.

Mende-Siedlecki, P., Baron, S. G., & Todorov, A. (2013). Diagnostic value underlies asymmetric updating of impressions in the morality and ability domains. *Journal of Neuroscience, 33,* 19406–19415.

Mende-Siedlecki, P., Cai, Y., & Todorov, A. (2012). The neural dynamics of updating person impressions. *Social Cognitive and Affective Neuroscience, 8,* 623–631.

O'Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J., & Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neuroscience, 4,* 95–102.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9,* 97–113.

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Science, 105,* 11087–11092.

Palmer, S. E., Schloss, K. B., & Sammartino, J. (2013). Visual aesthetics and human preference. *Annual Review of Psychology, 64,* 77–107.

Payne, B. K. (2005). Conceptualizing control in social cognition: How executive functioning modulates the expression of automatic stereotyping. *Journal of Personality and Social Psychology, 89,* 488–503.

Pia, L., & Tamietto, M. (2006). Unawareness in schizophrenia: Neuropsychological and neuroanatomical findings. *Psychiatry and Clinical Neuroscience, 60,* 531–537.

Reeder, G. D., & Coovert, M. D. (1986). Revising an impression of morality. *Social Cognition, 4,* 1–17.

Renzi, S., Schiavi, S., Carbon, C. C., Vecchi, T., Silvanto, J., & Cattaneo, Z. (2013). Processing of featural and configural aspects of faces is lateralized in dorsolateral prefrontal cortex: A TMS study. *NeuroImage, 74,* 45–51.

Robertson, E., Theoret, H., & Pascual-Leone, A. (2003). Studies in cognition: The problems solved and created by transcranial magnetic stimulation. *Journal of Cognitive Neuroscience, 15,* 948–960.

Rolls, E. T. (2000). The orbitofrontal cortex and reward. *Cerebral Cortex, 10,* 284–294.

Rossi, S., Hallett, M., Rossini, P. M., & Pascual-Leone, A. (2011). Screening questionnaire before TMS: An update. *Clinical Neurophysiology, 122,* 1686.

Rudoy, J. D., & Paller, K. A. (2009). Who can you trust? Behavioral and neural differences between perceptual and memory-based influences. *Frontiers in Human Neuroscience, 3,* 16.

Schiller, D., Freeman, J. B., Mitchell, J. P., Uleman, J. S., & Phelps, E. A. (2009). A neural mechanism of first impressions. *Nature Neuroscience, 12,* 508–514.

Schutter, D. J., & Van Honk, J. (2006). Increased positive emotional memory after repetitive transcranial magnetic stimulation over the orbitofrontal cortex. *Journal of Psychiatry Neuroscience, 31,* 101–104.

Siebner, H. R., Hartwigsen, G., Kassuba, T., & Rothwell, J. C. (2009). How does transcranial magnetic stimulation modify neuronal activity in the brain? Implications for studies of cognition. *Cortex, 45,* 1035–1042.

Silvanto, J., & Cattaneo, Z. (2014). State-dependency protocols. In A. Rotenberg, J. C. Horvath, & A. Pascual-Leone (Eds.), *NeuroMethods: Transcranial magnetic stimulation*. New York, NY: Springer.

Takahashi, H., Kato, M., Matsuura, M., Koeda, M., Yahata, N., Suhara, T., & Okubo, Y. (2008). Neural correlates of human virtue judgment. *Cerebral Cortex, 18*, 1886–1891.

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain*. New York, NY: Thieme Medical.

Tamietto, M., Latini Corazzini, L., Pia, L., Zettin, M., Gionco, M., & Geminiani, G. (2005). Effects of emotional face cueing on line bisection in neglect: A single case study. *Neurocase, 11,* 399–404.

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*(3), 282–288.

Thoma, P., Norra, C., Juckel, G., Suchan, B., & Bellebaum, C. (2015). Performance monitoring and empathy during active and observational learning in patients with major depression. *Biological Psychology, 109,* 222–231.

Thut, G., & Pascual-Leone, A. (2010). A review of combined TMS-EEG studies to characterize lasting effects of repetitive TMS and assess their usefulness in cognitive and clinical neuroscience. *Brain Topography, 22,* 219–232.

Todorov, A., & Olson, I. R. (2008). Robust learning of affective trait associations with faces when the hippocampus is damaged, but not when the amygdala and temporal pole are damaged. *Social Cognitive and Affective Neuroscience, 3,* 195–203.

Todorov, A., & Uleman, J. S. (2003). The efficiency of binding spontaneous trait inferences to actors' faces. *Journal of Experimental Social Psychology, 39,* 549–562.

Urgesi, C., Berlucchi, G., & Aglioti, S. M. (2004). Magnetic stimulation of extrastriate body area impairs visual processing of nonfacial body parts. *Current Biology, 14,* 2130–2134.

Yamada, M., Hirao, K., Namiki, C., Hanakawa, T., Fukuyama, H., Hayashi, T., & Murai, T. (2007). Social cognition and frontal lobe pathology in schizophrenia: A voxel-based morphometric study. *NeuroImage, 35,* 292–298.

Zanto, T. P., Rubens, M. T., Thangavel, A., & Gazzaley, A. (2011). Causal role of the prefrontal cortex in top-down modulation of visual processing and working memory. *Nature Neuroscience, 14,* 656–661.