

Downregulation of the Posterior Medial Frontal Cortex Prevents Social Conformity

Vasily Klucharev,^{1,2,3} Moniek A. M. Munneke,^{1,4} Ale Smidts,² and Guillén Fernández^{1,5}

¹Donders Institute for Brain, Cognition, and Behaviour, Radboud University Nijmegen, NL-6500 HB Nijmegen, The Netherlands, ²Rotterdam School of Management, Erasmus University Rotterdam, 3000 DR Rotterdam, The Netherlands, ³Department of Psychology, University of Basel, 4055 Basel, Switzerland, and Departments of ⁴Neurology and ⁵Cognitive Neuroscience, Radboud University Nijmegen Medical Centre, 6500 HB Nijmegen, The Netherlands

We often change our behavior to conform to real or imagined group pressure. Social influence on our behavior has been extensively studied in social psychology, but its neural mechanisms have remained largely unknown. Here we demonstrate that the transient downregulation of the posterior medial frontal cortex by theta-burst transcranial magnetic stimulation reduces conformity, as indicated by reduced conformal adjustments in line with group opinion. Both the extent and probability of conformal behavioral adjustments decreased significantly relative to a sham and a control stimulation over another brain area. The posterior part of the medial frontal cortex has previously been implicated in behavioral and attitudinal adjustments. Here, we provide the first interventional evidence of its critical role in social influence on human behavior.

Introduction

Humans demonstrate various forms of herding—alignments of the thoughts or behaviors of individuals in a group (herd) without centralized coordination (Raafat et al., 2009). For instance, the well documented phenomenon of conformity refers to the act of changing one's behavior to match the responses of others (Cialdini and Goldstein, 2004). A functional magnetic resonance imaging (fMRI) study (Klucharev et al., 2009) revealed that social conformity is reflected in the activity of the ventral striatum and the dorsal aspect of the posterior medial frontal cortex (pmFC), brain areas that are often implicated in reward processing and behavioral adjustments (Berns et al., 2001; Holroyd and Coles, 2002; Knutson and Wimmer, 2007). Other fMRI studies showed that tendencies for adolescents (Berns et al., 2010) and adults (Campbell-Meiklejohn et al., 2010; Falk et al., 2010) to conform to the opinion of their reference group or experts were positively correlated with activation in the pmFC. Furthermore, the pmFC was also found to be central to learning about social information (advice) and for determining the extent to which it guides behavior (Behrens et al., 2008). Interestingly, imaging studies (van Veen et al., 2009; Izuma et al., 2010) demonstrated that the activation of the pmFC reflected cognitive dissonance—an important cognitive mechanism underlying various forms of social influence (Festinger, 1957). Overall, there is cumulative evidence

for a hypothesis suggesting that some important forms of social influence affect our behavior via the pmFC. However, the studies published so far provide correlational evidence only. Consequently, one can always argue that activity of the pmFC is just an epiphenomenon or a consequence of social influence. Thus, we set out to use an interventional approach to probe whether the pmFC is a critical neural substrate of conforming behavioral adjustment.

In an effort to further study the role of the pmFC in social influence, we used repetitive transcranial magnetic stimulation (TMS) to modulate social conformity to group opinion. We hypothesized that pmFC downregulation evoked by TMS reduces the extent of conforming behavioral adjustments, the probability of their occurrence, or both.

Materials and Methods

We applied a short protocol of offline TMS—continuous theta-burst stimulation (Huang et al., 2005)—briefly before a social conformity task.

Stimulation sites. The stimulation site in the right pmFC was selected based on the activation observed in our previous fMRI study of social conformity (Klucharev et al., 2009): rostral cingulate zone, Brodmann areas 24, 32, 6, and 9; center of mass at MNI coordinates (x, y, z) = [8, 16, 52] mm. The control stimulation site was located in the medial parietal cortex (precuneus). At the beginning of the experiment, we localized the central sulcus using single TMS pulses and determined the active motor threshold (see the next paragraph for details). Since both the toe/leg representation and the pmFC (i.e., rostral cingulate zone) are located at the same depth level of the medial cortex, we determined the pmFC location by moving the coil in front of the central sulcus. The individual distance of the pmFC from the central sulcus was determined by the size of the subject's head in the same fashion as described for the international 10–20 system (Klem et al., 1999)—the approach commonly used in TMS studies (e.g., Griskova et al., 2007; Knoch et al., 2009). Overall, for stimulation, the TMS coil was placed anterior (mean distance: 4.5 cm, during pmFC and Sham stimulations) or posterior (mean distance: 4.6 cm,

Received April 14, 2011; revised June 27, 2011; accepted July 5, 2011.

Author contributions: V.K., A.S., and G.F. designed research; V.K. and M.A.M.M. performed research; V.K., A.S., and G.F. analyzed data; V.K., M.A.M.M., A.S., and G.F. wrote the paper.

We gratefully acknowledge the financial support by the Erasmus Research Institute of Management (ERIM).

We thank Jörg Rieskamp for the incisive discussion of the results. We thank the referees for their valuable and instructive comments.

Correspondence should be addressed to Vasily Klucharev, Department of Psychology, University of Basel, Misionsstrasse 62A, 4055 Basel, Switzerland. E-mail: vasily.klucharev@unibas.ch.

DOI:10.1523/JNEUROSCI.1869-11.2011

Copyright © 2011 the authors 0270-6474/11/3111934-07\$15.00/0

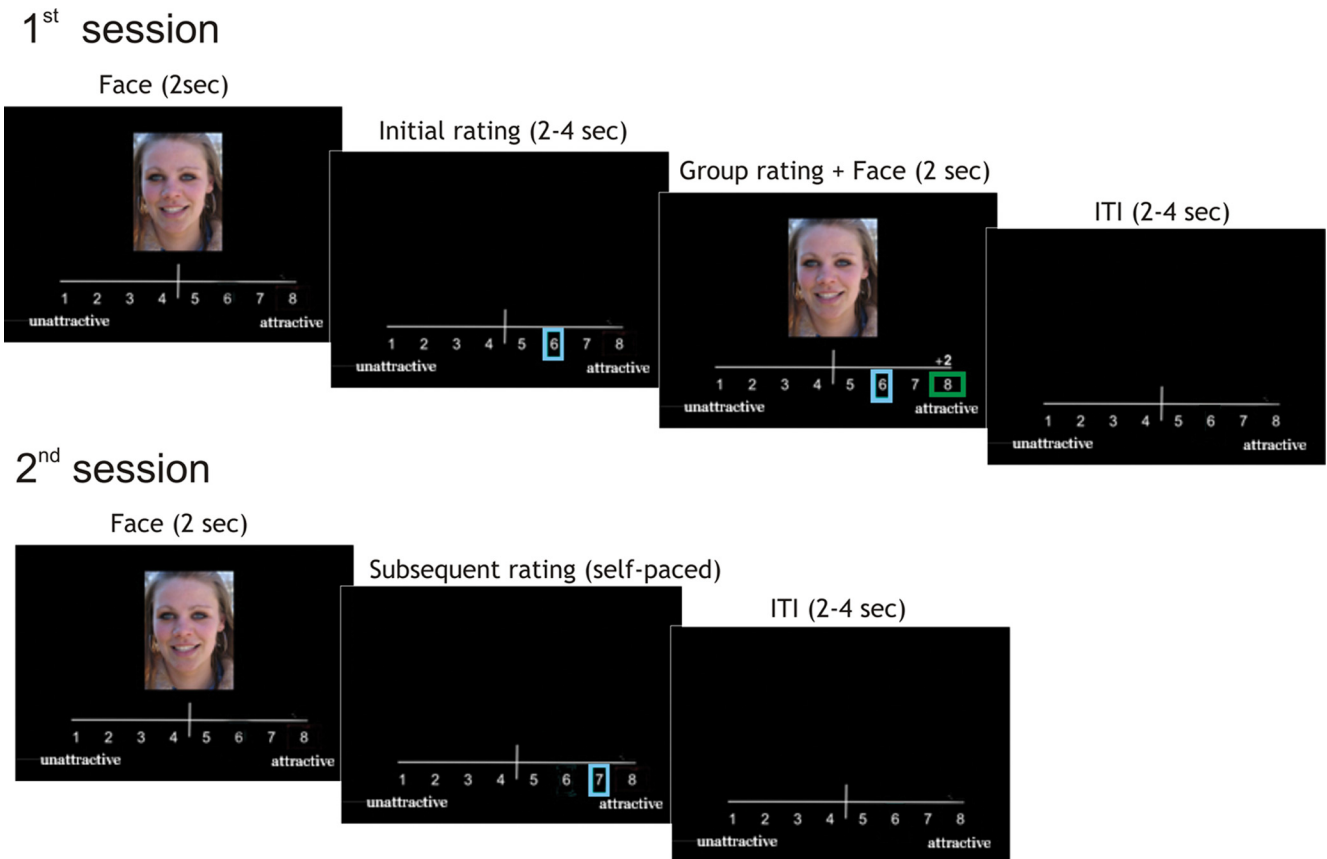


Figure 1. The task (first session) evoking a conflict with group ratings followed immediately after the TMS. The second session assessed the subsequent behavioral change induced by group ratings and its interaction with TMS. The sequence of events within one trial is shown by one example. During the first session, female subjects rated the attractiveness of female faces and were subsequently presented with group ratings that might be similar (no conflict with group ratings), below, or above (as is shown in the figure) subjects' rating (conflict with group ratings). Twenty minutes later, subjects rated the same faces again in a second session to identify the subsequent conformity effects. For each subject, we calculated the mean values of the conformal adjustments (changes) in rating between the first and second session.

during Control stimulation of the parietal cortex) to the central sulcus. Since our conformity paradigm includes a behavioral session unanticipated by subjects (see below), we decided not to use an individually fMRI-guided TMS, because repeating the task for the fMRI localizer and the actual TMS study would reveal the goal of the experiment.

Stimulation parameters. Previous studies successfully demonstrated the effects of standard repetitive TMS of the pMFC on Stroop interference effect (Hayward et al., 2004) and behavioral changes in task switching (Rushworth et al., 2002). Here we used a new TMS protocol, continuous theta-burst stimulation, in which a 40 s train of patterned 50 Hz stimulation is given (600 pulses), resulting in significant reduction of neural activity lasting for nearly 60 min (Huang et al., 2005). A 110 mm double cone coil (Magstim Company) was connected to the magnetic stimulator (Magstim Rapid). This configuration uses two angled windings to improve coupling to the head, increasing its effectiveness at stimulating relatively deep brain areas as targeted here. During stimulation of the midline toe/leg area of primary motor cortex (Cz), active motor thresholds were established as the minimum stimulus intensity that produced a liminal EMG response in 50% of trials during isometric contraction of the tibialis anterior muscle. The stimulation intensity was set at 80% of this active motor threshold (mean values: 25% of the stimulator's maximum output, SD = 1.79, min value: 21%, max value: 27%). The coil was held radially to the subject's head. Subjects received the 40 s theta-burst train over either the pMFC or parietal cortex (control stimulation), or sham stimulation (a very low subthreshold level of TMS—10% of the maximum output). Subjects performed the task immediately after the stimulation train ("offline paradigm") in the same laboratory. As they had received the instructions for the task before the TMS train, they were able to begin the task ~3–5 min after the stimulation train ended, thus when under the influence of the TMS effect.

Participants. Forty-nine young right-handed women (students, aged 19–23 years, mean 21.1 years) participated in the study consisting of two experimental sessions. None of the subjects reported any history of drug abuse, head trauma, neurological, or psychiatric illness. Two participants were rejected from the study because they had expressed concerns or nervousness about the TMS procedure. Three participants were assigned to the Sham group due to a high active motor threshold. The remaining subjects were randomly assigned to one of three groups: a group receiving TMS to the pMFC (pMFC group, 17 subjects); a group receiving TMS to the medial parietal cortex (Control group, 15 subjects); or a Sham group receiving a subthreshold TMS (10% of the maximum stimulator output over pMFC, 17 subjects). The study was approved by the ethical committee and written informed consent was obtained from all subjects.

Stimuli. A set of 222 digital photos of European females (aged 18–35 years, from free internet sources) were used as stimuli. Color portraits of moderately attractive females (mean 4.2, SD = 1.2 of the eight-point scale) were selected from a set of 1000 stimuli, all highly similar in both appearance and photographic style (see Klucharev et al., 2009, for details).

Experiment. Subjects were informed that they were participating in a project "Seeing Beauty," which was studying human perceptions of attractiveness. During the first session subjects were exposed to a series of 222 photographs of female faces [stimulus duration = 2 s, intertrial interval (ITI) = 2–4 s] (Fig. 1). They were instructed to rate the face on an eight-point scale ranging from very unattractive (1) to very attractive (8). Subjects indicated their rating by pressing the appropriate button, using eight buttons on a PC keyboard. The subject's rating (initial rating, blue rectangle frame) was visualized on the screen immediately after the face stimulus. Three to five seconds later, at the end of each trial, the subject was informed (by a green rectangle frame) of the average rating of

the same face given by 200 students from the same Dutch university (group rating). The difference between the subject's rating and the group rating was also indicated by a score presented above the scale (0, ± 2 , or ± 3 points). Importantly, the frame and the number indicating the conflict with group opinion were visible during both "conflict" and "non-conflict" trials. Actual group ratings were programmed using the following criteria: in 33% of trials, group ratings agreed with subjects' ratings ("no conflict"), whereas in 67% of trials, group ratings were pseudorandomly above or below subject's rating by 2 points ("small conflict") or 3 points ("large conflict"). This was done by using an adaptive algorithm that kept the overall ratio of "more negative" or "more positive" group ratings approximately equal during the experiment. We informed subjects that group ratings that matched with their own rating within a ± 1 point range produced a group rating frame that overlapped visually with the frame of their own rating. Subjects were not informed initially about the real purpose of the experiment or the manipulation of the group ratings. All photographs were randomized across subjects and conditions.

In the unanticipated second session, subjects were instructed to rate again (self-paced) the attractiveness of the same faces presented in a new randomized order (subsequent rating, Fig. 1). See our previous fMRI study (Klucharev et al., 2009) for the behavioral validation of the task. All subjects were debriefed after the experiment. None of the subjects reported any suspicions about the cover story. None reported alternative behavioral strategies, e.g., while giving their ratings, they never used a strategy to simply predict group ratings. Importantly, the observed conformity effects were calculated relative to judgments given during the second session in absence of group ratings. The large number of stimuli and the long break between sessions ensured that ratings during the second session were not affected by explicit memory but indicated subjects' own opinion. Importantly, blinded TMS stimulation is almost impossible due to the fact that the subjects feel the stimulation and the experimenter can see the TMS effect when observing scalp muscle contraction. Therefore, written instructions and computerized behavioral tasks ensured researchers' biases had minimal influence on the subjects' conformity.

Overall, the setup of the current study imitates social psychological studies investigating persuasion, where subjects are informed about a dominant behavior in a group (Cialdini, 2007). Social psychology suggests two types of social norms: injunctive norms have a moral tone and characterize what people should do, whereas descriptive norms represent typical behavior or what most people actually do, regardless of its appropriateness (Cialdini and Goldstein, 2004). In the current study, we investigated descriptive social norms, which send the message, "If a lot of people are doing this, it's probably a wise thing to do." It is also important to note that in our study, subjects were not involved in a standard reinforcement task; i.e., they could not learn correct answers, or correct evaluation criteria, because there was no correct answer: the normative feedback was pseudorandom. Critically, attractiveness is a socially important facial feature (Langlois et al., 2000); judgments of facial attractiveness are fast, effortless, and consistent across subjects (Willis and Todorov, 2006). Therefore, a mismatch between individual judgments of facial attractiveness and group opinion should create a strong normative conflict.

Data analysis. To detect conformal behavioral adjustments, we analyzed changes of ratings between the two sessions: the mean differences between the second and the first ratings were calculated separately for conflict and no-conflict trials. Additionally, we calculated conformal adjustments relative to a baseline of no-conflict trials to reduce variability related to the baseline condition: the mean change in no-conflict trials was subtracted from the mean change in conflict trial ratings. The effect of TMS on conformal adjustments was analyzed using a two-way ANOVA with group rating (three levels: group ratings are more positive, equal, or more negative)—as a within-subject factor and TMS (three levels: pMFC TMS, Control, and Sham TMS)—as a between-subject factor. The Tukey HSD test was used whenever appropriate to test specific contrasts. The data were analyzed using SPSS for Windows, version 7.

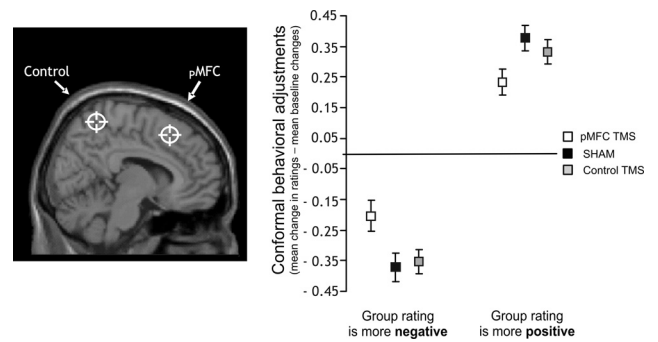


Figure 2. Effects of offline theta-burst TMS of the pMFC, Sham, and Control (medial parietal cortex) site on mean conforming adjustments of judgments. Conformal behavioral effect (right panel)—The mean change in rating after group ratings that had been more negative (or positive) than subject's own initial rating (conflict trials) subtracted from the change in rating after group ratings that had been equal to subject's own initial rating (no-conflict trials). The sites of the TMS stimulations (left panel) are overlaid on a sagittal view of a standard T1 MRI. The crosshairs indicate the intended target regions for the Control stimulation and for the pMFC stimulation placed at the coordinate of the maximal conjunction effect in the previous fMRI study (Klucharev et al., 2009). This conjunction analysis combined the effects of conflicts with group ratings and prediction of subsequent conformal behavioral adjustments. SHAM, Sham stimulation, i.e., subthreshold TMS of the pMFC (10% of the maximum stimulator output). Error bars indicate SEM.

Results

Effect of TMS on the size of conformal adjustments

Overall, subjects rated faces as moderately attractive (pMFC stimulation = 4.3, Sham stimulation = 4.2, Control stimulation = 4.2). For each subject, we calculated the mean values of the conformal adjustments (changes) in rating between the first and second session. In line with our expectations, participants across all groups changed their ratings of attractiveness, aligning themselves with group ratings: participants decreased their attractiveness ratings when group ratings had been more negative than their own initial rating, whereas more positive group ratings were associated with more positive reevaluation of faces. ANOVA analysis revealed a main effect of the factor group rating (three levels: group ratings that are more negative, equal, or more positive than subjects' own ratings) on the mean values of the conformal adjustments in attractiveness ratings across all subjects: $F_{(2,45)} = 171.79$, $p = 0.0001$. As predicted, we observed an effect of TMS on conformity supported by an interaction between the factors group rating (three levels) \times TMS (three levels) for mean values of the conformal adjustments: $F_{(4,92)} = 6.35$, $p = 0.004$ due to lower conformity in the pMFC condition.

Since our analysis specifically focuses on the effect that TMS has on conflicts trials, we also analyzed conformity effects relative to a baseline of no-conflict trials. Thereby, we reduced variability related to differences in the baseline condition. Figure 2 shows the effect of TMS on mean conformity adjustments to more negative or more positive group ratings that were calculated relative to the baseline adjustments following equal group ratings (see Table 1 for the underlying data). The baseline-corrected effect was supported by a significant interaction between the factor of group ratings (two levels: mean values of the conformal adjustments due to group ratings that are more negative or more positive than subjects' own ratings, both subtracted by mean values of the baseline adjustments in attractiveness ratings following equal group ratings) and the factor of TMS (three levels): $F_{(2,46)} = 5.49$, $p = 0.007$. *Post hoc* comparisons using the Tukey HSD test indicated that the conformity adjustments after pMFC stimulation decreased relative to the Sham ($p = 0.019$) and the Control stimula-

Table 1. Original and corrected^a conformity effects and SDs

	Group ratings		
	More negative	Equal	More positive
Means (SD)			
pMFC TMS	−0.18 (0.20)	0.02 (0.25)	0.26 (0.21)
Sham	−0.55 (0.23)	−0.18 (0.27)	0.20 (0.22)
Control TMS	−0.40 (0.25)	−0.05 (0.28)	0.28 (0.24)
Corrected means^a (SD)			
pMFC TMS	−0.20 (0.18)		0.24 (0.16)
Sham	−0.37 (0.15)		0.38 (0.14)
Control TMS	−0.35 (0.13)		0.33 (0.13)

^aCorrected conformity effects—the mean change in rating after group ratings that had been more negative (or positive) than subject’s own initial rating (conflict trials) subtracted from the change in rating after group ratings that had been equal to subject’s own initial rating (no-conflict trials).

tions $p = 0.042$). In turn, the Control stimulation did not significantly differ from the Sham stimulation $p = 0.25$ as expected. After the pMFC stimulation, the mean size of the conformal adjustments decreased on average by 0.14 points on the eight-point scale: from 0.33–0.38 point adjustments after the Sham and the Control stimulations to 0.20–0.24 point adjustments after the pMFC stimulations (Fig. 2, Table 1). In addition, to exclude possible unspecific effects of TMS, we analyzed the effect of TMS only on those trials in which conformity changes were observed. The mean size of the conformal adjustments (calculated relative to the baseline adjustments) significantly decreased after TMS of the pMFC as indicated by a significant interaction: group rating (two levels) \times TMS (three levels), $F_{(2,46)} = 6.03$, $p = 0.005$. The Tukey HSD test further confirmed that the conformity adjustments after pMFC stimulation significantly decreased relative to the Sham ($p = 0.009$) and the Control ($p = 0.011$) stimulations in trials in which conformity changes were observed. Taken as a whole, after the pMFC stimulation, conformal behavioral adjustments decreased relative to the Sham and the Control stimulations.

Effect of TMS on the probability of conformal adjustments

On average, across all conditions, conformity changes occurred with probability 0.41. The probability of conformal changes significantly decreased from 0.43 (Sham stimulation) and 0.42 (Control stimulation) to 0.38 after TMS of the pMFC as confirmed by an ANOVA: TMS (three levels), $F_{(2,46)} = 3.55$, $p = 0.037$. Figure 3 clearly shows that pMFC stimulation reduced the probability of conformal behavioral adjustments relative to the Sham and the Control stimulations. *Post hoc* comparisons using the Tukey HSD test indicated that the probability of conformal changes decreased after pMFC stimulation relative to the Sham ($p = 0.025$) and the Control stimulations ($p = 0.033$).

Interaction of the effects of TMS with the levels of conflicts

Interestingly, the effect of TMS on the mean size of the conformal adjustments calculated relative to the baseline was more robust for small conflicts with group opinion than for large ones (see Table 2 for details). This observation was supported by a significant three-way interaction: group rating (two levels) \times TMS (three levels) \times level of conflict (small vs large): $F_{(4,45)} = 3.40$, $p = 0.042$. Figure 4 shows that TMS had a significant effect for both small and large conflicts with group opinion, but the effect of TMS for large conflicts was reduced.

Effect of TMS on the individual tendencies to conform to group opinion

To establish a closer relationship between group ratings and individual behavior, we performed individual correlation and re-

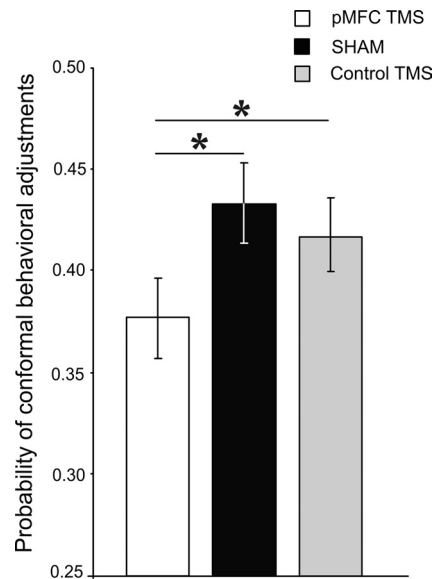


Figure 3. Mean effects of TMS of the pMFC, Sham, and Control (medial parietal cortex) site on the probability of conformal adjustments. Error bars indicate the SEM. *Significant differences at the level of $p < 0.05$ (Tukey HSD test; see Results for details).

Table 2. Original and corrected^a conformity effects and SDs for different levels of conflict

	Group ratings				
	More negative		Equal	More positive	
Conflict	−3	−2	0	+2	+3
Means (SD)					
pMFC TMS	−0.34 (0.18)	−0.09 (0.22)	0.02 (0.25)	0.18 (0.23)	0.40 (0.19)
Sham	−0.64 (0.21)	−0.47 (0.25)	−0.18 (0.27)	0.10 (0.24)	0.25 (0.21)
Control TMS	−0.49 (0.23)	−0.34 (0.26)	−0.05 (0.28)	0.22 (0.25)	0.40 (0.22)
Corrected means^a (SD)					
pMFC TMS	−0.36 (0.16)	−0.12 (0.21)		0.16 (0.16)	0.38 (0.10)
Sham	−0.46 (0.16)	−0.29 (0.12)		0.28 (0.21)	0.43 (0.09)
Control TMS	−0.44 (0.12)	−0.29 (0.15)		0.26 (0.15)	0.45 (0.12)

^aCorrected conformity effects—see Table 1 for details.

gression analysis between the magnitude of the conflict (i.e., the difference value between subjects’ own and group ratings during the first session: −3, −2, 0, +2, or +3) and the subsequent change in the perceived facial attractiveness. Previously, the individual correlation coefficients were used as conformity scores, i.e., a measure of the individual tendency to conform (Klucharev et al., 2009). We found a significant correlation in all groups (pMFC stimulation: mean $r = 0.18$, $SD = 0.07$, $p = 0.007$; Sham stimulation: mean $r = 0.25$, $p = 0.0002$, $SD = 0.06$; Control stimulation; mean $r = 0.23$, $p = 0.007$, $SD = 0.07$). Overall, we found a significant effect of TMS (three levels) on these conformity correlations: $F_{(2,45)} = 3.67$, $p = 0.033$. Thus, the individual tendency to conform was lower after TMS of the pMFC. These results were further supported by a linear regression analysis. The unstandardized regression coefficients were smaller after the pMFC stimulation than after the Sham or the Control stimulation: $F_{(2,46)} = 4.89$, $p = 0.012$ [mean unstandardized regression coefficient (b) and percentage of significant effects $p < 0.05$): pMFC stimulation: $b = 0.096$, 64.7%; Sham stimulation: $b = 0.155$, 87.7%; Control stimulation; $b = 0.140$, 94.1%]. Overall, the pMFC stimulation decreased the individual tendency to conform to group opinion.

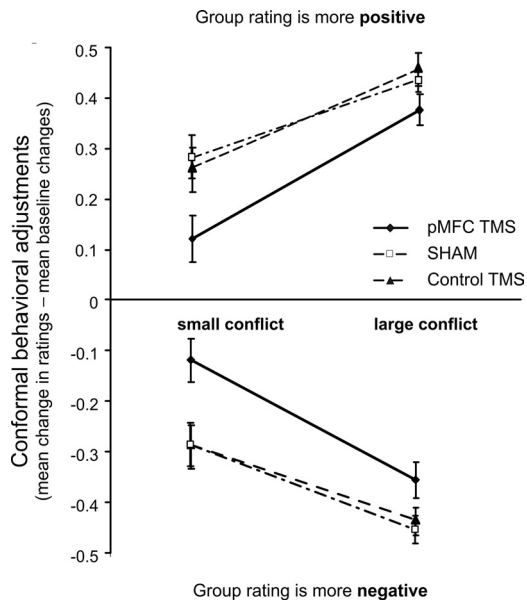


Figure 4. Mean effects of TMS of the pMFC, Sham, and Control site (medial parietal cortex) on mean conforming adjustments of ratings after small (± 2) or large (± 3) conflicts with group opinion [separately for positive conflicts (top panel) and negative conflicts (bottom panel)]. TMS had a significant effect both for small and for large conflicts with group opinion, but the effect of TMS was more significant for small conflicts (pMFC vs Sham stimulation: $t_{(1,34)} = 2.61$, $p = 0.014$; pMFC vs Control stimulation: $t_{(1,32)} = 2.75$, $p = 0.008$) than for large conflicts (pMFC vs Sham stimulation: $t_{(1,34)} = 2.12$, $p = 0.042$; pMFC vs Control stimulation: $t_{(1,32)} = 2.19$, $p = 0.036$).

The observed effect of TMS on conformity was not a “performance-driven” effect

It is also possible to suggest that conformity effects could be driven by unequal distributions of ratings between conflict and no-conflict trials in the initial session (for similar logic, see Chen and Risen, 2010). To compare distributions of ratings in the initial session, we conducted a three-way ANOVA with ratings (eight levels) and conflict (conflict vs no-conflict trials) as within-subjects factors and TMS (three levels) as a between-subjects factor. We found no significant differences of ratings’ distribution (percentages of ratings) between conflict and no-conflict trials in all three TMS conditions: neither for the interaction ratings \times conflict, $F_{(7,322)} = 0.958$, $p = 0.462$, nor for the interaction ratings \times conflict \times TMS, $F_{(14,322)} = 1.383$, $p = 0.159$. Thus, our effects cannot be readily explained by an artifact of unequal distributions of ratings between conflict and no-conflict trials in the initial session.

Additional nonspecific effects of TMS on response range

In addition, we examined the effects of pMFC stimulation on the overall range of ratings. We found a slight but significant range effect of the stimulation of the pMFC: subjects demonstrated a tendency to give fewer extreme ratings (very unattractive and very attractive). Importantly, the range effect when included in the analysis as a covariate (i.e., the ratio of intermediate ratings 4 and 5 to extreme ratings 1 and 8) does not eliminate the behavioral effect of TMS on the mean absolute values of the conformal adjustments: group ratings \times TMS interaction, $F_{(2,45)} = 5.37$, $p = 0.008$. Overall, our behavioral finding of TMS effects cannot be explained by a simple decrease in response variance, but seems to entail a true adjustment to group feedback.

Our results also indicate that the stimulation of the pMFC had a slight effect on the baseline condition—no conflict with group

opinion. Normally (i.e., after the Sham stimulation), during the second session subjects gave slightly more negative ratings after no conflict with the group, but the pMFC stimulation significantly reduced this order effect. Baseline adjustments were different in three TMS conditions: mean adjustments after pMFC stimulation = 0.02; Control stimulation = -0.05 , Sham stimulation = -0.18 (see Table 1). Overall, the pMFC stimulation slightly reduced baseline adjustments: Sham stimulation versus pMFC stimulation: $t_{(1,32)} = 2.19$, $p = 0.035$; Sham stimulation versus Control stimulation: $t_{(1,30)} = 1.4$, $p = 0.17$; Control stimulation versus pMFC stimulation: $t_{(1,30)} = 1.7$, $p = 0.07$. This baseline effect may indicate TMS modulation of unspecific order effects: pMFC downregulation could suppress the unspecific trend for negative adjustments of ratings in the second session. Regardless of this, this baseline effect does not contradict the reported effects of TMS on conformal behavioral adjustments, since those effects were reliably stronger in conflict trials.

Discussion

Together, our results show that pMFC downregulation reduced conformity by attenuating the effect that a conflict with a normative group opinion has on behavior. Previous fMRI experiments showed that some forms of social influence are associated with neuronal response in the pMFC (Behrens et al., 2008; Klucharev et al., 2009; Berns et al., 2010; Campbell-Meiklejohn et al., 2010; Falk et al., 2010) and ventral striatum (Klucharev et al., 2009; Campbell-Meiklejohn et al., 2010), two brain areas involved in the monitoring of behavioral outcomes. To our knowledge, the current study directly probed for the first time how critical the pMFC is and demonstrated that its transient downregulation reduces social influence.

The pMFC was previously implicated into conflict detection (Botvinick et al., 2001), cognitive dissonance (van Veen et al., 2009; Izuma et al., 2010), volatility monitoring (Behrens et al., 2007), “goal-based action selection” (Matsumoto and Tanaka, 2004a,b), error likelihood prediction (Brown and Braver, 2005), and error processing (Holroyd and Coles, 2002; Ridderinkhof et al., 2004; Matsumoto et al., 2007). One possible explanation of our results is that TMS of the pMFC could be reducing cognitive/emotional dissonance evoked by conflicts with social norms. Alternatively, TMS of the pMFC could be interrupting a general performance-monitoring mechanism that normally signals the need for performance adjustment (Ridderinkhof et al., 2004).

The role of the pMFC in performance monitoring has been supported by numerous animal and neuroimaging studies (for extensive reviews, see Holroyd and Coles, 2002; Matsumoto and Tanaka, 2004b; Ridderinkhof et al., 2004). In humans, the pMFC is particularly responsive to performance errors in various tasks (Ullsperger and von Cramon, 2001, 2003; Gehring and Willoughby, 2002; Kerns et al., 2004). Neurons in the pMFC are probably involved in generating a “history of action” (Kennerley et al., 2006; Jocham et al., 2009b); distinct neuronal populations of the pMFC encode positive and negative prediction errors of action values (Matsumoto et al., 2007). Activity of the pMFC represents the direction of error in action-value prediction and varies as a function of the environmental context (Jocham et al., 2009a) to allow optimal adaptation to reversing reward contingencies. Furthermore, the pMFC activity might serve as a trigger to shift from an exploratory to an exploitative behavioral mode (Procyk et al., 2000; Quilodran et al., 2008). Thus, the pMFC does not only evaluate feedback, but it also optimizes action adaptation and valuation. In line with the proposed role in performance monitoring, the magnitude of the pMFC’s activity has been

shown to predict the strength of subsequent behavioral adjustments during simple choice decisions (O'Doherty et al., 2003a; Kerns et al., 2004; Cohen and Ranganath, 2007). In agreement with its performance monitoring role, the pmFC has prominent connections with the ventral striatum (Groenewegen et al., 1982; Parkinson et al., 2000; Hauber and Sommer, 2009). This is an area involved in reward prediction (Berns et al., 2001; McClure et al., 2003; O'Doherty et al., 2003b), social learning (Rilling et al., 2002), and social influence (Klucharev et al., 2009; Campbell-Meiklejohn et al., 2010). Furthermore, disconnection of the pmFC from the ventral striatum impairs decision making and stimulus–reward learning in animals (Parkinson et al., 2000; Hauber and Sommer, 2009). Thus, a downregulation of the pmFC could perturb the entire network underlying performance monitoring and social learning. Overall, TMS of the pmFC in the current study may have interrupted the neural performance monitoring mechanism and consequently have led to reduced conforming behavioral adjustments to normative group opinion.

The reinforcement learning theory of performance monitoring suggests that the pmFC activity is modulated by a dopaminergic signal, which indicates whether the outcome of an action is worse or better than expected, regardless of the primary cause of the deviation from the prediction (Holroyd and Coles, 2002; Matsumoto et al., 2007). Theoretically, thus, TMS of the pmFC can influence reinforcement learning mechanisms, and, consequently, affect conforming behavioral adjustments. Interestingly, general accounts of conformity from psychology and behavioral economics also comply with a reinforcement learning mechanism; i.e., both suggest that social norms selectively reinforce certain behaviors. Whereas psychological theories emphasize the rewarding value of social approval or affiliation with others (Cialdini and Goldstein, 2004), behavioral economics focuses more on the effects of punishment for violating the norm (Fehr and Fischbacher, 2004). Our results concur with this general mechanism. Importantly, in contrast to a standard reinforcement learning paradigm, each stimulus was presented only twice in our study, and no systematic feedback was provided. Our data therefore cannot be modeled with a standard reinforcement learning model. Nevertheless, we suggest that the underlying mechanism of social conformity might still be based on a similar mechanism. That is, individuals could compare their own opinion or expectation with the social feedback and this difference could be computed as a prediction error (i.e., the belief of the group subtracted from one's own belief). Anyway, conformity with social norms does require a neural signal indicating deviations from norms (Montague and Lohrenz, 2007). Interestingly, the experiments that were directly designed to model prediction error (Behrens et al., 2008; Burke et al., 2010) indeed demonstrated a prediction-error-like signal generated by some forms of social influence.

Our results also indicate that the effect of the pmFC TMS on the conformal adjustments was more robust for small conflicts with group opinion than for large ones. It has been argued that conformal behavioral changes evoked by conflicts with normative group opinion take effect when the conflict signal reaches a hypothetical threshold that triggers conformity (Klucharev et al., 2009). Thus, only an activity that crosses such a threshold evokes a change of behavior. This would imply that when conflict-related activity is low and further suppressed by TMS, it cannot reach the hypothetical threshold triggering conformity. Large conflicts presumably lead to a larger conflict-related activity at the pmFC, which may overcome to some extent the inhibitory effect of TMS. Therefore, one may speculate that large conflicts

can overcome the medial frontal cortex downregulation, whereas small conflicts are less likely to do so.

Importantly, the pmFC stimulated in the current study is located posterior to the region of the medial prefrontal cortex that is often implicated in specialized social cognitive abilities like theory of mind (Frith and Singer, 2008). Thus, our results suggest that the underlying neural mechanism of the conformal adjustments is unspecific and similar to the fundamental neural mechanism of behavioral adjustments—a fundamental mechanism that is similar for social and for nonsocial situations (Behrens et al., 2008). Therefore, the current study does not need a control experiment, i.e., a “computer feedback” control condition. Computers are commonly used as a control condition in social neuroscience studies to demonstrate the social nature of observed effects. However, a “normative” experiment of this type could only show an attenuated, but fundamentally similar, behavioral effect—a finding that has no additional value for the current study. Despite revealing the unspecific neural mechanism underlying conformity, our task has a social nature, as demonstrated by previous behavioral and fMRI studies (Klucharev et al., 2009). The social relatedness between the subjects and the “group” was correlated with the degree of conformity in our task (Klucharev et al., 2009). Interestingly, a recent fMRI study (Zaki et al., 2011) demonstrated the specific conformity effect of our task on the stimuli's reward value: activity of the orbitofrontal cortex and nucleus accumbens was higher for stimuli that were rated more positively by peers than by participants themselves, as opposed to stimuli that peers rated less positively than participants. Therefore, conformity effects in our task are based on a true reevaluation of stimuli evoked by social norms.

Further experiments are needed to clarify the exact role of the pmFC in social influence. Current results should be interpreted with respect to the female population only, and additional studies will help to generalize the observed mechanisms to the male population and other social situations (including injunctive or moral norms) leading to conformity. Future experiments will also help to investigate factors modulating social conformity, e.g., uncertainty of subjects' opinion, attention, and even the dynamics of the local neuronal metabolism.

Benjamin Franklin observed one day that “Singularity in the right hath ruined many; happy those who are convinced of the general opinion.” Indeed, social psychology, behavioral economics, and neuroimaging studies all indicate that conformity with the general opinion is a relatively automatic process that relates to the power of social norms. Previous fMRI studies (e.g., Behrens et al., 2008; Klucharev et al., 2009; Campbell-Meiklejohn et al., 2010) demonstrate that a “singularity in the right” triggers activity of the pmFC that signals a need for behavioral adjustments. To sum up, the pmFC appears to be an important target of social influence that occurs when specific behavioral changes are induced by other people. In fact, the present study shows that TMS of the pmFC reduces social influence on our judgments. By and large, our results further support the hypothesis that social influence is underlined by a fundamental performance-monitoring neural mechanism that makes a truly independent opinion so difficult to achieve.

References

- Behrens TEJ, Woolrich MW, Walton ME, Rushworth MFS (2007) Learning the value of information in an uncertain world. *Nat Neurosci* 10:1214–1221.
- Behrens TEJ, Hunt LT, Woolrich MW, Rushworth MFS (2008) Associative learning of social value. *Nature* 456:245–249.

- Berns GS, McClure SM, Pagnoni G, Montague PR (2001) Predictability modulates human brain response to reward. *J Neurosci* 21:2793–2798.
- Berns GS, Capra CM, Moore S, Noussair C (2010) Neural mechanisms of the influence of popularity on adolescent ratings of music. *Neuroimage* 49:2687–2696.
- Botvinick MM, Braver TS, Barch DM, Carter CS, Cohen JD (2001) Conflict monitoring and cognitive control. *Psychol Rev* 108:624–652.
- Brown JW, Braver TS (2005) Learned predictions of error likelihood in the anterior cingulate cortex. *Science* 307:1118–1121.
- Burke CJ, Tobler PN, Baddeley M, Schultz W (2010) Neural mechanisms of observational learning. *Proc Natl Acad Sci U S A* 107:14431–14436.
- Campbell-Meiklejohn DK, Bach DR, Roepstorff A, Dolan RJ, Frith CD (2010) How the opinion of others affects our valuation of objects. *Curr Biol* 20:1165–1170.
- Chen MK, Risen JL (2010) How choice affects and reflects preferences: re-visiting the free-choice paradigm. *J Pers Soc Psychol* 99:573–594.
- Cialdini RB (2007) Descriptive social norms as underappreciated sources of social control. *Psychometrika* 72:263–268.
- Cialdini RB, Goldstein NJ (2004) Social influence: compliance and conformity. *Annu Rev Psychol* 55:591–621.
- Cohen MX, Ranganath C (2007) Reinforcement learning signals predict future decisions. *J Neurosci* 27:371–378.
- Falk EB, Berkman ET, Mann T, Harrison B, Lieberman MD (2010) Predicting persuasion-induced behavior change from the brain. *J Neurosci* 30:8421–8424.
- Fehr E, Fischbacher U (2004) Social norms and human cooperation. *Trends Cogn Sci* 8:185–190.
- Festinger L (1957) *A theory of cognitive dissonance*. Evanston, IL: Row, Peterson.
- Frith CD, Singer T (2008) The role of social cognition in decision making. *Philos Trans R Soc Lond B Biol Sci* 363:3875–3886.
- Gehring WJ, Willoughby AR (2002) The medial frontal cortex and the rapid processing of monetary gains and losses. *Science* 295:2279–2282.
- Griskova I, Ruksenas O, Dapsys K, Herpertz S, Höppner J (2007) The effects of 10 Hz repetitive transcranial magnetic stimulation on resting EEG power spectrum in healthy subjects. *Neurosci Lett* 419:162–167.
- Groenewegen HJ, Room P, Witter MP, Lohman AH (1982) Cortical afferents of the nucleus accumbens in the cat, studied with anterograde and retrograde transport techniques. *Neuroscience* 7:977–996.
- Hauber W, Sommer S (2009) Prefrontostriatal circuitry regulates effort-related decision making. *Cereb Cortex* 19:2240–2247.
- Hayward G, Goodwin GM, Harmer CJ (2004) The role of the anterior cingulate cortex in the counting Stroop task. *Exp Brain Res* 154:355–358.
- Holroyd CB, Coles MG (2002) The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol Rev* 109:679–709.
- Huang YZ, Edwards MJ, Rounis E, Bhatia KP, Rothwell JC (2005) Theta burst stimulation of the human motor cortex. *Neuron* 45:201–206.
- Izuma K, Matsumoto M, Murayama K, Samejima K, Sadato N, Matsumoto K (2010) Neural correlates of cognitive dissonance and choice-induced preference change. *Proc Natl Acad Sci U S A* 107:22014–22019.
- Jocham G, Neumann J, Klein TA, Danielmeier C, Ullsperger M (2009a) Adaptive coding of action values in the human rostral cingulate zone. *J Neurosci* 29:7489–7496.
- Jocham G, Klein TA, Neumann J, von Cramon DY, Reuter M, Ullsperger M (2009b) Dopamine DRD2 polymorphism alters reversal learning and associated neural activity. *J Neurosci* 29:3695–3704.
- Kennerly SW, Walton ME, Behrens TE, Buckley MJ, Rushworth MF (2006) Optimal decision making and the anterior cingulate cortex. *Nat Neurosci* 9:940–947.
- Kerns JG, Cohen JD, MacDonald AW 3rd, Cho RY, Stenger VA, Carter CS (2004) Anterior cingulate conflict monitoring and adjustments in control. *Science* 303:1023–1026.
- Klem GH, Lüders HO, Jasper HH, Elger C (1999) The ten-twenty electrode system of the International Federation. The International Federation of Clinical Neurophysiology. *Electroencephalogr Clin Neurophysiol Suppl* 52:3–6.
- Klucharev V, Hytönen K, Rijpkema M, Smidts A, Fernández G (2009) Reinforcement learning signal predicts social conformity. *Neuron* 61:140–151.
- Knoch D, Schneider F, Schunk D, Hohmann M, Fehr E (2009) Disrupting the prefrontal cortex diminishes the human ability to build a good reputation. *Proc Natl Acad Sci U S A* 106:20895–20899.
- Knutson B, Wimmer GE (2007) Splitting the difference: how does the brain code reward episodes? *Ann N Y Acad Sci* 1104:54–69.
- Langlois JH, Kalakani L, Rubenstein AJ, Larson A, Hallam M, Smoot M (2000) Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychol Bull* 126:390–423.
- Matsumoto K, Tanaka K (2004a) Neuroscience. Conflict and cognitive control. *Science* 303:969–970.
- Matsumoto K, Tanaka K (2004b) The role of the medial prefrontal cortex in achieving goals. *Curr Opin Neurobiol* 14:178–185.
- Matsumoto M, Matsumoto K, Abe H, Tanaka K (2007) Medial prefrontal cell activity signaling prediction errors of action values. *Nat Neurosci* 10:647–656.
- McClure SM, Berns GS, Montague PR (2003) Temporal prediction errors in a passive learning task activate human striatum. *Neuron* 38:339–346.
- Montague PR, Lohrenz T (2007) To detect and correct: norm violations and their enforcement. *Neuron* 56:14–18.
- O'Doherty J, Critchley H, Deichmann R, Dolan RJ (2003a) Dissociating valence of outcome from behavioral control in human orbital and ventral prefrontal cortices. *J Neurosci* 23:7931–7939.
- O'Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ (2003b) Temporal difference models and reward-related learning in the human brain. *Neuron* 38:329–337.
- Parkinson JA, Willoughby PJ, Robbins TW, Everitt BJ (2000) Disconnection of the anterior cingulate cortex and nucleus accumbens core impairs Pavlovian approach behavior: further evidence for limbic cortical-ventral striatopallidal systems. *Behav Neurosci* 114:42–63.
- Procyk E, Tanaka YL, Joseph JP (2000) Anterior cingulate activity during routine and non-routine sequential behaviors in macaques. *Nat Neurosci* 3:502–508.
- Quilodran R, Rothé M, Procyk E (2008) Behavioral shifts and action valuation in the anterior cingulate cortex. *Neuron* 57:314–325.
- Raafat RM, Chater N, Frith C (2009) Herding in humans. *Trends Cogn Sci* 13:420–428.
- Ridderinkhof KR, Ullsperger M, Crone EA, Nieuwenhuis S (2004) The role of the medial frontal cortex in cognitive control. *Science* 306:443–447.
- Rilling J, Gutman D, Zeh T, Pagnoni G, Berns G, Kilts C (2002) A neural basis for social cooperation. *Neuron* 35:395–405.
- Rushworth MF, Hadland KA, Paus T, Sipila PK (2002) Role of the human medial frontal cortex in task switching: a combined fMRI and TMS study. *J Neurophysiol* 87:2577–2592.
- Ullsperger M, von Cramon DY (2001) Subprocesses of performance monitoring: a dissociation of error processing and response competition revealed by event-related fMRI and ERPs. *Neuroimage* 14:1387–1401.
- Ullsperger M, von Cramon DY (2003) Error monitoring using external feedback: specific roles of the habenular complex, the reward system, and the cingulate motor area revealed by functional magnetic resonance imaging. *J Neurosci* 23:4308–4314.
- van Veen V, Krug MK, Schooler JW, Carter CS (2009) Neural activity predicts attitude change in cognitive dissonance. *Nat Neurosci* 12:1469–1474.
- Willis J, Todorov A (2006) First impressions: making up your mind after a 100-ms exposure to a face. *Psychol Sci* 17:592–598.
- Zaki J, Schirmer J, Mitchell JP (2011) Social influence modulates the neural computation of value. *Psychol Sci* 22:894–900.