A causal account of the brain network computations underlying strategic social behavior

Christopher A Hill[1] , Shinsuke Suzuki[2,3], Rafael Polania[1], Marius Moisa[1,4], John P O'Doherty[5,6] & Christian C Ruff[1]

During competitive interactions, humans have to estimate the impact of their own actions on their opponent's strategy. Here we provide evidence that neural computations in the right temporoparietal junction (rTPJ) and interconnected structures are causally involved in this process. By combining inhibitory continuous theta-burst transcranial magnetic stimulation with model-based functional MRI, we show that disrupting neural excitability in the rTPJ reduces behavioral and neural indices of mentalizing-related computations, as well as functional connectivity of the rTPJ with ventral and dorsal parts of the medial prefrontal cortex. These results provide a causal demonstration that neural computations instantiated in the rTPJ are neurobiological prerequisites for the ability to integrate opponent beliefs into strategic choice, through system-level interaction within the valuation and mentalizing networks.

Competition is a principle of nature[1]. When multiple agents want access to a finite resource, an agent's outcome can become contingent on the choices of other agents. Success in these situations depends critically on the ability to predict the future actions of competitors. Strategically sophisticated agents therefore employ mentalizing to simulate the opponents' belief about the agent's next course of action[2–4]. This is supported by theoretical accounts[4], behavioral observations[3] and neurobiological data[2].

In particular, corresponding signals have been linked with neural activity in the temporoparietal junction (TPJ), a key area of the so-called mentalizing network[5], both when agents generate false signals to influence the opponents[2] and when they detect that their opponent uses such strategies[6]. Moreover, blood-oxygen-level-dependent (BOLD) signal changes in the right TPJ can predict how much people act upon these beliefs to manipulate opponents[7,8]. This evidence suggests that the rTPJ computes higher-order beliefs about the opponent's intentions, based on correlations between parameters of fitted learning or choice models and measured neural activity. However, this leaves unclear whether neural computations in the rTPJ are indeed functionally relevant for behavior in the manner implied by models incorporating higher-order beliefs. This is because identical predictions for neural signals can be generated by various computational models that differ substantially in algorithmic formulation[9]. Correlations between model-derived parameters and neural activity therefore do not imply a causal mapping between neural activity and computations implemented in the model; they may also reflect compounds of very different computations that together resemble the predicted signal[10]. Hence, establishing that the rTPJ indeed computes

representations used for strategic social choices requires tests of how external modulations of rTPJ excitability affect both behavior and neural activity during competitive social interactions.

Here we provide such a test by combining functional MRI (fMRI) with repetitive transcranial magnetic stimulation (rTMS). For this purpose, we employed the inspection game[2], a two-player competitive game with a 2 × 2 strategy space and asymmetric payoffs (Online Methods). As performing above chance in this game depends on the player's ability to predict the opponent's behavior in any given round, mentalizing-related processes are essential to correctly infer the strategy of the opponent. The use of these mentalizing processes can be studied with a model[2] (Online Methods) that formalizes both how the players react to their opponent's past choices (first-order beliefs) and how they anticipate the influence of their own choice on the opponent's behavior (i.e., mentalizing-related second-order beliefs). The latter computation, termed 'influence signal', is thought to be expressed at the time of observing the choice outcome and has been found to correlate with activity in the inferior portion of TPJ (posterior superior temporal sulcus, pSTS)[2]. The impacts of first-order and second-order beliefs on choices are captured by distinct parameters, allowing us to disentangle these two processes. We employed rTMS to test whether disrupting rTPJ excitability prevents an agent from anticipating the influence of their past actions on the opponent's choice. By employing fMRI together with the stimulation, we explored how disruption of neural processing in the rTPJ affected the expression of influence signals, both for behavior and for the associated neural processes in the stimulated area and interconnected brain regions.

[1]Laboratory for Social and Neural Systems Research (SNS-Lab), Department of Economics, University of Zurich, Zurich, Switzerland. [2]Frontier Research Institute for Interdisciplinary Sciences, Tohoku University, Sendai, Miyagi, Japan. [3]Institute of Development, Aging and Cancer, Tohoku University, Sendai, Miyagi, Japan. [4]Biomedical Engineering, University and ETH of Zurich, Zurich, Switzerland. [5]Division of the Humanities and Social Sciences, California Institute of Technology, Pasadena, California, USA. [6]Computation and Neural Systems, California Institute of Technology, Pasadena, California, USA. Correspondence should be addressed to C.A.H. (christopher.hill@econ.uzh.ch) or C.C.R. (christian.ruff@econ.uzh.ch).

More specifically, we compared choices and neural activity between two groups of participants who received an inhibitory rTMS protocol (cTBS[11]) either over the rTPJ or over a vertex control site (to mimic any unspecific rTMS effects; Online Methods). By comparing both groups, we tested the hypothesis that disrupting the neural coding of the influence signal localized in rTPJ will cause the player to become less strategic, in that he or she will be hampered in the ability to estimate how his or her own actions influence the opponent's future behavior. Moreover, we were able to test for a direct correspondence between the computational model and the neural computations instantiated by the rTPJ, by testing whether cTBS over rTPJ compared to over the control site led to weaker BOLD activity in the rTPJ associated with the computationally derived influence signal.

Because rTMS coupled with fMRI can test for remote influences of the stimulated site on interconnected brain structures[12,13], we also tested how the altered influence-update signal in the rTPJ interacted with processing in connected brain structures. Two areas of particular interest were the dorsomedial prefrontal cortex (dmPFC)[14] and the ventromedial prefrontal cortex (vmPFC)[15], key nodes of the mentalizing[16] and valuation networks[17], respectively. Several studies have suggested that the dmPFC is involved in representing information about another agent's strategy[18,19]. Notably, it has been reported that the influence signal is increasingly represented in the dmPFC as a player's behavior increasingly conforms to the predictions of the influence model[2]. Moreover, functional coupling between the signals in rTPJ and dmPFC peaks at the time of feedback[2]. However, these correlational results alone cannot characterize the network properties of the mentalizing network. For instance, it is unclear whether the computations occurring in dmPFC depend on the outputs of rTPJ or whether the outputs of both regions are integrated in a third region. In the present study, we tease apart these possibilities: for instance, a decrease in connectivity between rTPJ and dmPFC due to cTBS would show that both regions are interdependent in the computations they perform. Alternatively, activity disruptions in rTPJ may result in an enhanced representation of the influence update in dmPFC by triggering compensatory activity[20].

For the vmPFC, communication between the mentalizing network and the valuation network has been proposed to play a role in integrating the intentions of others into choice values[15,21,22]. In the context of competitive auctions, it has been reported that rTPJ–vmPFC coupling predicts individual overbidding[21]. Furthermore, vmPFC–rTPJ coupling has been shown to increase during episodes of market bubbles[15] and during charitable giving[22]. Thus, if the influence-update signal is relevant for goal-directed behavior guided by vmPFC value representations, we expect that disrupting neural processing in rTPJ will affect value computations in the vmPFC by means of altered functional connectivity.

RESULTS
Behavioral results
In total, 120 healthy subjects (60 females, ages 18–25) partook in the experiment. Sixty subjects underwent cTBS to either the right TPJ (30 subjects) or the vertex control site (30 subjects), followed by fMRI scanning (**Fig. 1a**) during which they played 160 trials of the inspection game in the role of employee (**Fig. 1b**). The other 60 subjects (who were neither scanned nor stimulated with cTBS) took the role of the employer in a real competitive game. Subject groups were matched in terms of gender (15 males, 15 females per condition) and did not differ significantly with regards to age, baseline strategic sophistication, working memory, cTBS-related beliefs, cTBS-related discomfort and reaction time (all $P > 0.05$; Online Methods and **Supplementary**

Table 1). Thus, any behavioral and neural differences between the groups resulted from the application of rTPJ-cTBS versus control vertex-cTBS.

Results from a hierarchical Bayesian model estimation procedure indicated that subjects indeed employed mentalizing-related computations to win the game, as the Influence model afforded the best fit to the data (**Fig. 1c**) compared to alternative models simulating random responding (mixed Nash equilibrium), reinforcement learning and fictitious play. Moreover, our results show that the weighting given to second-order beliefs (i.e., the influence parameter κ; Online Methods) reflects a strategically important aspect of behavioral control: κ was positively related to total payoff ($t_{54} = 2.38$, $P = 0.02$), indicating that participants who were more sensitive to the influence of their actions on the opponent won more often (**Fig. 1d**). This is consistent with the view that individuals capable of complex strategic behavior (who therefore have high κ in the context of our experiment) are less numerous in the population[4] than those who resort to strategies guided by the simpler first-order learning (captured by η). The former can exploit the behavior of the latter to their advantage, which concurs with the positive relationship between κ and earnings found here.

rTPJ-cTBS reduces behavioral indices of influence computations
We first tested whether rTPJ-cTBS inhibits mentalizing-related computations, by means of a model-free analysis of the behavioral data. If the neural activity inhibited by cTBS indeed reflects computation of the opponent's beliefs about the player's behavior, then subjects in the rTPJ-cTBS condition should be more likely to repeat their previous choice compared to subjects in the vertex-cTBS condition. This is because an unimpaired representation of the opponent's first-order beliefs promotes switching between choice options, as switching can exploit the naive first-order learning strategy of countering recently observed choices. In other words, the process of anticipating and correcting for the opponent's first-order learning should lead participants to avoid repetitions of the same choice in order to prevent being exploited. Our results indicate that subjects in the rTPJ-cTBS condition indeed switched their choices less frequently ($t_{56} = -2.04$, $P = 0.046$; **Fig. 2a**).

However, note that switching per se did not predict task performance ($t_{55} = 0.91$, $P = 0.36$) and is only a rough proxy for the effect of second-order beliefs on behavior. In fact, the increased tendency to stay with one's previous choice may also reflect a lower responsiveness to the opponent's actions rather than a deficit in the ability to incorporate the influence of one's own actions on the opponent. We therefore employed a logit regression model that allowed us to predict the subject's current choice based on both the subject's own and the opponent's previous choice. If rTPJ-cTBS selectively disrupts an agent's ability to process beliefs about how their actions influence their opponent, then we ought to observe that participants in this group still react normally to the opponent's choices but become more predictable on the basis of their previous choice.

We found that participants in the rTPJ-cTBS group differed significantly ($t_{56} = -2.03$, $P = 0.047$) from participants in the vertex-cTBS group in how much their actions could be predicted based on their past choice (**Fig. 2b**). The rTPJ-cTBS participants were indeed predictable on the basis of their past actions, in the sense that they tended to repeat their previous choice (t test vs. log-odds ratio = 0, $t_{28} = -5.44$, $P = 8.25 \times 10^{-6}$; **Fig. 2b**). In contrast, participants in the vertex-cTBS condition were not predictable on the basis of their previous choice ($t_{28} = -0.95$, $P = 0.34$; **Fig. 2b**). The difference in predictability between both cTBS conditions held even if we controlled in our model for an estimate of how much the participants themselves
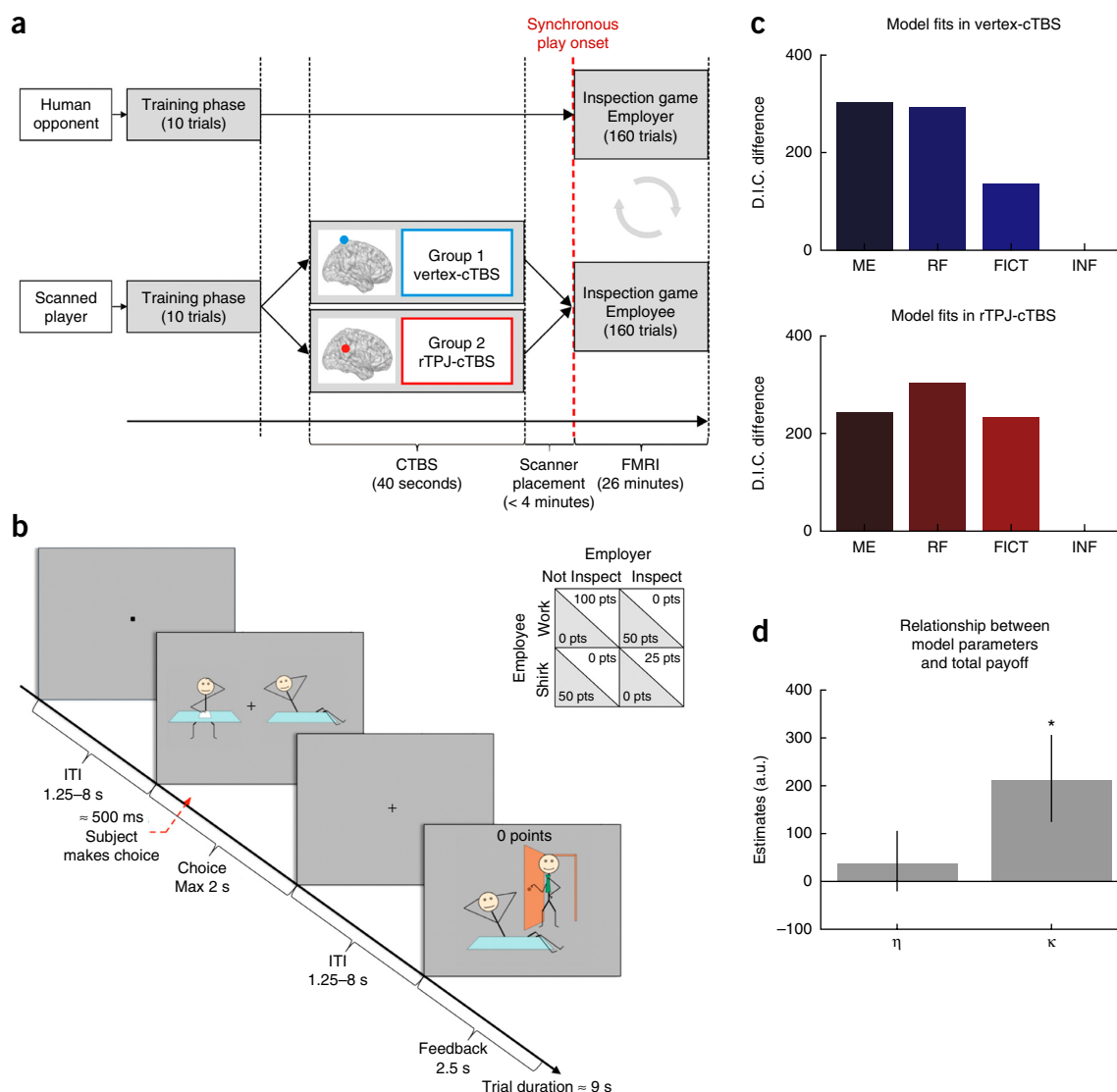
Figure 1 Experimental design and model evaluation. (**a**) Experimental design and timeline. Subjects first underwent neuronavigated TMS motor-threshold calibration and performed a short training phase. They then received cTBS over a control vertex or rTPJ before performing the task in the scanner as the Employee, facing a human opponent playing as Employer. Vertical dotted lines separate the different phases of the experiment and anchor these phases to the timeline below. (**b**) fMRI model and payoff matrix. Choice and Feedback periods were temporally separated. The participant had 2 s to make a choice. Top right: the payoff matrix of the inspection game. The Employee receives 50 points if he or she successfully predicts the action of the Employer. The Employer gets 25 points if he or she Inspects and finds that the Employed Shirked, or 100 points if he or she chooses Not Inspect and the Employee Works. (**c**) Evidence that participants used the influence model (deviance information criterion (DIC) is a generalization of the Bayesian information criterion (BIC) for hierarchical models[41]; the smaller the DIC, the better the model fits to the data). Shown are differences in DIC values for models (fitted to both groups) corresponding to purely random responding (mixed equilibrium; ME), reinforcement learning (RF) and fictitious play (FICT), compared to the influence model (INF). The latter model affords the best fit to the data in both experimental groups. (**d**) Coefficients for the regression of total Employee payoff on the first-order learning rate η ($t_{54} = 0.65$, $P = 0.51$) and second order-beliefs measured by the parameter κ ($t_{54} = 2.3$, *$P = 0.025$). Note that our model controls for the effect of cTBS condition ($t_{54} = 1.15$, $P = 0.25$). This result shows that taking into account the effect of one's actions on the opponent's strategy is associated with a better behavioral outcome. Error bars denote standard errors of the regression coefficients.

reacted to their opponent's past action ($t_{55} = -2.26$, $P = 0.027$; **Fig. 2b**). This suggests that rTPJ-cTBS did not affect overall cognitive ability but rather selectively disrupted the participant's ability to account for how their own actions influenced the opponent's future behavior. Notably, not only the mean but also the variance of these estimates was significantly reduced (Levene's test for equality of variance, $F_{1,56} = 7.6$, $P = 0.007$) in the rTPJ-cTBS condition. This suggests that rTPJ-cTBS may have predominantly affected those subjects who normally would have been capable of more sophisticated play.

Having established a basic behavioral difference between both cTBS conditions that was consistent with our hypothesis, we went on to test whether disrupting the influence update process by rTPJ-rTMS indeed impaired the computation of second-order beliefs, as formalized in the influence update model. We tested this hypothesis by fitting the influence model (Online Methods) to the behavioral data of our participants in a hierarchical Bayesian framework and inspecting the difference in the population-level distributions of the influence update parameter κ (Online Methods). This revealed that compared

to the vertex-cTBS condition, participants in the rTPJ-cTBS condition indeed gave less weight to second-order beliefs as indicated by a lower value of κ ($P_{MCMC} = 0.003$; P_{MCMC} is a P value derived by comparing the posterior distributions of the estimated parameters sampled via Markov chain Monte Carlo; see Online Methods and **Fig. 2c**). This confirmed our central hypothesis. Subsequent exploratory analysis did not reveal any significant differences in the remaining parameters of the model, specifically the first-order learning rate η ($P_{MCMC} = 0.65$) and the softmax inverse temperature ($P_{MCMC} = 0.12$; Online Methods). Moreover, to ensure that the cTBS-related reduction in κ did not reflect changes in the tendency of at least a few subjects to generate random responses, we constructed a mixture model that tested whether cTBS changed influence learning while controlling for possible changes in how much subjects randomized their responses (Online Methods). This model also showed that cTBS led to reductions in the key parameter κ ($P_{MCMC} = 0.03$), whereas it did not change the degree of random responding (Online Methods). All these findings corroborate that the behavioral changes due to cTBS indeed reflected a selective reduction in the ability to account for the influence of one's own choices on opponent choices.

Notably, the results of the model-based and model-free analysis were fully congruent: the influence parameter κ was strongly associated with the logit estimate of how much subjects' choices could be predicted by their previous choices ($t_{54} = 7.39$, $P = 9.5 \times 10^{-9}$), as well as with the switch frequency ($t_{54} = 5.4$, $P = 1.4 \times 10^{-6}$; **Fig. 2d**). Thus, our behavioral data show that rTPJ-cTBS indeed led to a systematic deficit in the ability to incorporate second-order beliefs into choice, suggesting that the influence-update computation implemented in the rTPJ is necessary for sophisticated strategic behavior to emerge.

Local cTBS effects on influence computations in rTPJ

We performed a replication of Hampton *et al.*[2], who found that BOLD activity in bilateral pSTS correlates with the influence update signal. The replication was successful (across both groups: right pSTS region of interest (ROI), peak at Montreal Neurological Institute (MNI) coordinates x, y, $z = 54$, -64, 4; $t_{56} = 6.36$, P(small-volume-corrected for family-wise error; SVC) $= 0$; left pSTS ROI, peak at MNI x, y, $z = -48$, -70, 10; $t_{56} = 5.12$, P(SVC) $= 0.001$; Online Methods). This confirmed that the expected representations of the influence update were indeed present in pSTS for our dataset. Moreover, exploratory whole-brain analysis across both groups revealed bilateral activations in the TPJ extending into STS, bilateral activations in dorsolateral prefrontal cortex extending into dmPFC and activity in visual areas (all P(nonparametric family-wise error corrected; FWE) < 0.05; **Supplementary Table 2**). For just the vertex-cTBS group, we also found pronounced influence-update-related activity in rTPJ extending to pSTS (cluster extent (Ke) = 1,326, degrees of freedom (df) = 28, nonparametric P(FWE) = 0.009; **Fig. 3a** and **Supplementary Table 2**).

Next, we searched for TMS-induced disruptions of the influence-update process by comparing BOLD signals between the vertex-cTBS and rTPJ-cTBS groups. An initial ROI analysis (Online Methods) of the stimulated area showed that, compared to the control group, cTBS indeed significantly reduced the rTPJ BOLD activity associated with the influence update signal (right TPJ ROI, peak at MNI x, y, $z = 39$, -46, 28; $t_{55} = 3.65$, P(SVC) $= 0.002$; **Fig. 3b**). Exploratory whole-brain analysis did not reveal any other effects in brain areas remote from the stimulation site (whole-brain vertex-cTBS > rTPJ-cTBS, all P(FWE) > 0.05).
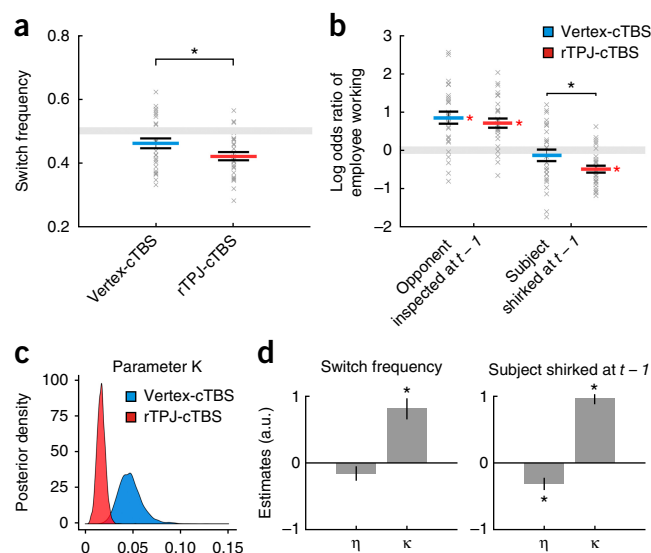

Figure 2 Effects of rTPJ-cTBS on behavior. (**a**) Difference in switch frequency between conditions. Subjects in the rTPJ-cTBS group were less likely to switch than those in the vertex-cTBS group. Black line intervals denote s.e.m. (**b**) We disentangled the effects of the opponent's choice and the subject's choice with a random-effects logit regression. We predicted the Employee's choice at t based on the Employee's past choice at $t - 1$ and the opponent's choice at $t - 1$ for each subject. This revealed that subjects in the rTPJ-cTBS group (red lines) but not the vertex group (blue lines) were more predictable on the basis of their past actions. Black line intervals denote s.e.m. Red asterisks denote that the means of the coefficients are statistically different from 0. (**c**) Model-based estimates of the population-level posterior distributions of κ for both conditions. Subjects in the rTPJ-cTBS condition had lower κ than the control group. (**d**) Regression analysis of how (left) switch frequency and (right) the predictive value of previous actions for subsequent choices depended on the first-order learning rate η and second-order learning rate κ. This shows that κ from the model-based analysis indeed captures similar aspects of strategic choice behavior as do the model-free indices in **a** and **b**. All depicted estimates were derived with multiple regression models that also included a predictor coding for stimulation condition, thereby controlling for baseline differences between stimulation groups. Error bars denote standard errors of the regression coefficients; *$P < 0.05$.

Network cTBS effects on influence-update computations in the dmPFC

It has been postulated that dmPFC plays an important role in updating representation of other's actions in relation to one's own behavior[19,23]. In line with this conjecture, it has been reported that BOLD activity in the dmPFC represents the computations predicted by the influence model and that functional communication between dmPFC and rTPJ peaks at the time of feedback[2]. We therefore tested whether cTBS-perturbation of the rTPJ computations would also affect the related model-representations in the dmPFC. To do so, we calculated the divergence in model likelihood between the influence model and the fictitious play model for each subject given the parameters, as also done in Hampton *et al.*[2]. Next, we included this index as a predictor in a regression analysis examining how the neural representation of the influence signal during the outcome stage varied as a function of how likely a given participant was to employ this model. An ROI analysis (Online Methods) of the dmPFC region identified in the previous study[2] replicated the finding that the greater the difference in fit between the influence model and the fictitious play model, the more the influence signal was represented in the dmPFC (dmPFC ROI, peak at MNI x, y, $z = -3$, 44, 22; $t_{55} = 4.15$, P(SVC) $= 0.014$;

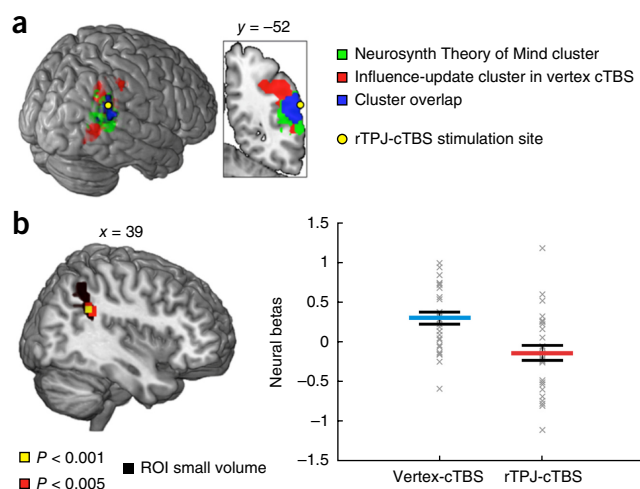**Figure 3** Stimulation site and local neural effects of rTPJ-cTBS. (**a**) The rTPJ surface stimulation point coordinates (yellow dot, MNI x, y, z = 61, −52, 35) were defined as the center of the overlap (blue) of the Neurosynth rTPJ cluster for theory of mind (green) with the representation of the influence-update signal in the vertex-cTBS group (red, shown here at whole-brain P(FWE) < 0.001). (**b**) Left: in line with our hypothesis, we found a reduced representation of the influence update in the rTPJ-cTBS group within the TPJ ROI (black). For display purposes, activations are depicted in red at P(uncorrected) < 0.005 and yellow at P(uncorrected) < 0.001. Right: parameter estimates for the regression of BOLD signals on the influence update, averaged across all yellow ROI voxels. Black lines denote s.e.m.

Figure 4 rTPJ-cTBS effects on neural activity in dmPFC. (**a**) Influence-update modulation in dmPFC. Left: in both conditions, subjects who relied more on the influence model than on the fictitious play model exhibited significantly more activation related to the influence signal in the dmPFC (16 mm spherical ROI in black centered on MNI x, y, z = −3, 51, 25 as reported in Hampton *et al.*[2]). For display purposes, activations are depicted in red at P(uncorrected) < 0.005 and yellow at P(uncorrected) < 0.001. Right: BOLD-signal parameter estimates extracted from the dmPFC cluster (yellow) for both conditions visualize the similar relationship between dmPFC BOLD signals and model reliance. Each dot represents one participant. (**b**) Whole-brain psychophysiological interaction (PPI) analysis. We found reduced connectivity during the feedback epoch between BOLD time-series in the rTPJ and a more dorsal dmPFC region. Graph shows PPI parameter estimates extracted from the dMPFC cluster (left; for all voxels with P < 0.001). Black lines denote s.e.m. For display purposes, activations in dmPFC depicted in red at P(FWE) < 0.005 and in yellow at P(uncorrected) < 0.001.

Fig. 4a). However, this representation did not differ significantly between groups (rTPJ-cTBS < vertex-cTBS; ROI analysis: peak at MNI x, y, z = 9, 56, 28; t_{53} = 1.53, P(SVC) = 0.809; exploratory whole-brain analysis: all P(FWE) > 0.05). Because the relationship between model likelihood and influence update in dmPFC was not significantly affected by the disruption of the computations occurring in rTPJ, we conclude that the influence model representation coded in the dmPFC did not rely exclusively on functional inputs from the signals coded in the rTPJ.

Second, we explored putative changes in the way both regions communicated at the time of feedback. To do so, we performed a psychophysiological interaction analysis comparing how rTPJ connectivity at the time of feedback (versus baseline) differed between the rTPJ-cTBS group and the control group. For the a priori selected dmPFC ROI, this analysis showed a numerical but statistically insignificant cTBS-induced decrease in rTPJ connectivity (dmPFC ROI, peak at MNI x, y, z = −6, 47, 10; t_{55} = 3.15, P(SVC) = 0.17). However, subsequent whole-brain exploratory analysis revealed a significant cTBS-induced decrease in connectivity of the rTPJ with a part of dmPFC (peak at MNI x, y, z = −9, 41, 40) that was more dorsal than the a priori selected ROI (Ke = 332, df = 55, nonparametric P(FWE) = 0.03; **Fig. 4b** and **Supplementary Table 3**).

Taken together, these results suggest that influence-update computations in the TPJ were conveyed to other parts of the mentalizing network, as reflected by the decrease in rTPJ–dmPFC connectivity due to rTPJ disruption by cTBS. This supports the view that the computations occurring in dmPFC and rTPJ are not fully independent but coordinated by means of functional communication between both areas at the time of feedback. Note that the directionality of this communication cannot be inferred from psychophysiological interaction analysis alone. However, because we directly interfered with the computations occurring in rTPJ and observed resulting changes
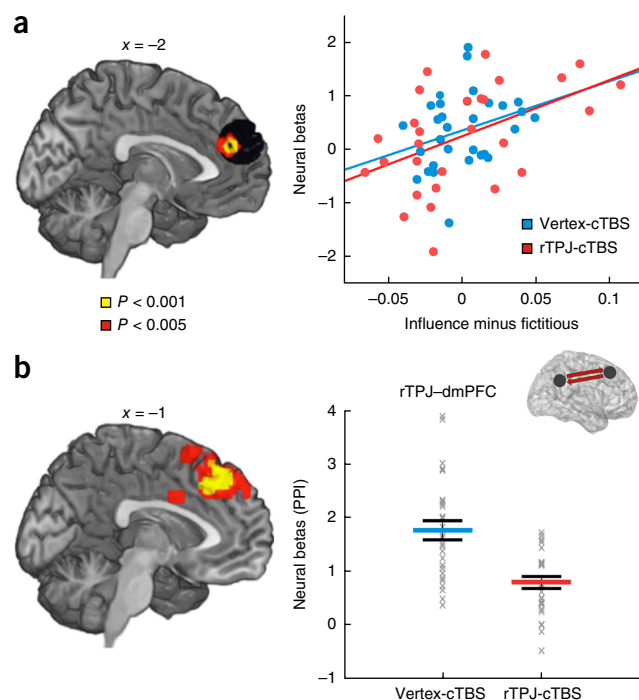
in the spatially remote dmPFC, it is more likely that cTBS affected the forwarding of information from rTPJ to dmPFC rather than the converse.

Network cTBS effects on value computations in vmPFC

Numerous studies have shown that choices in social settings involve value computations in vmPFC[2,15,21,22,24]. We examined how these vmPFC value representations at the time of choice were affected by cTBS-disruption of the TPJ influence update computations, using an ROI analysis (Online Methods) of a vmPFC area known to be involved in value representation[25,26]. When looking for BOLD signals encoding the chosen value, we found numerically congruent but statistically insignificant effects in the vmPFC (vmPFC ROI: peak at MNI x, y, z = −3, 32, −14; t_{56} = 1.12, P(SVC) = 0.84; exploratory whole-brain analysis: all FWE P > 0.05). Similarly, any decrease in value coding for TMS compared to control was also numerically consistent but statistically insignificant (vmPFC ROI: peak at MNI x, y, z = −15, 29, −14; t_{55} = 2.72, P(SVC) = 0.30; exploratory whole-brain analysis: all FWE P > 0.05). Exploratory analysis of the value signals in the two TMS groups by themselves revealed significant vmPFC value coding
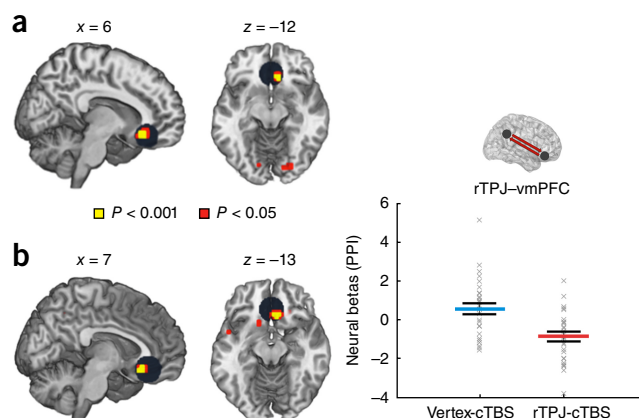
Figure 5 rTPJ-cTBS effects on value coding in vmPFC. (**a**) Value signals in the control group. Representation of the expected value of the chosen option in the vmPFC for the vertex-cTBS group (16-mm spherical ROI in black centered on MNI x, y, $z = 0$, 32, –13 as in previous studies[26,27]). For display purposes, activations are depicted in red at P (uncorrected) < 0.005 and yellow at P (uncorrected) < 0.001. (**b**) PPI analysis seeded in rTPJ. We found evidence that rTPJ-cTBS reduced influence-update-related connectivity between rTPJ and the vmPFC (same black ROI as in **a**) at the time of feedback. Graph shows PPI parameter estimates extracted from the vmPFC cluster (left; for all voxels with $P < 0.001$). Black lines denote s.e.m. For display purposes, activations are depicted in red at P (uncorrected) < 0.005 and yellow at P (uncorrected) < 0.001.

in the vertex-cTBS condition (vmPFC ROI: peak at MNI x, y, $z = 9$, 29, –11, $t_{28} = 4.11$, $P(\text{SVC}) = 0.04$; **Fig. 5a**) but no such effects in the rTPJ-cTBS condition (vmPFC ROI: peak at MNI x, y, $z = 0$, –32, 17; $t_{27} = 0.43$, $P(\text{SVC}) = 0.86$). Thus, it appears that the absence of statistically robust value coding in the vmPFC may mainly reflect noisy BOLD signals in the rTPJ-cTBS condition.

To investigate the possibility that functional impairment of rTPJ could interfere with the integration of second-order beliefs into vmPFC neural value signals, we performed a second psychophysiological analysis in which we analyzed how rTPJ–vmPFC connectivity at the time of feedback reflected the influence update and how this modulation of connectivity by the influence update was changed by the cTBS. This revealed that rTPJ-cTBS, compared to vertex-cTBS, led to lower influence-update-related rTPJ–vmPFC connectivity (vmPFC ROI: peak at MNI x, y, $z = 9$, 26, –14; $t_{55} = 3.95$, $P(\text{SVC}) = 0.029$; **Fig. 5b**). The peak-to-peak distance between the value representation at the time of choice (**Fig. 5a**) in the vertex control group and the differential connectivity pattern reflecting the influence-update signal (**Fig. 5b**) was 12 mm. This underlined that the functional connectivity disrupted by cTBS arose for a portion of vmPFC that overlapped with the portion found to encode value representations at the time of choice. More broadly, this result suggests that cTBS-induced impairment of TPJ interfered with the integration of second-order beliefs into value computations through alterations in cross-network communication.

Relationship between the neural and behavioral effects caused by TMS

Our analyses identified three main aspects of neural function that were affected by rTPJ-cTBS: first, the influence-update computation in rTPJ; second, the connectivity between dmPFC and rTPJ at the time of feedback; and third, the modulation of connectivity between rTPJ and vmPFC by the influence update signal. This raises the question: which of these neural changes underlie(s) the cTBS-induced behavioral deficit in integrating second-order beliefs into choice?
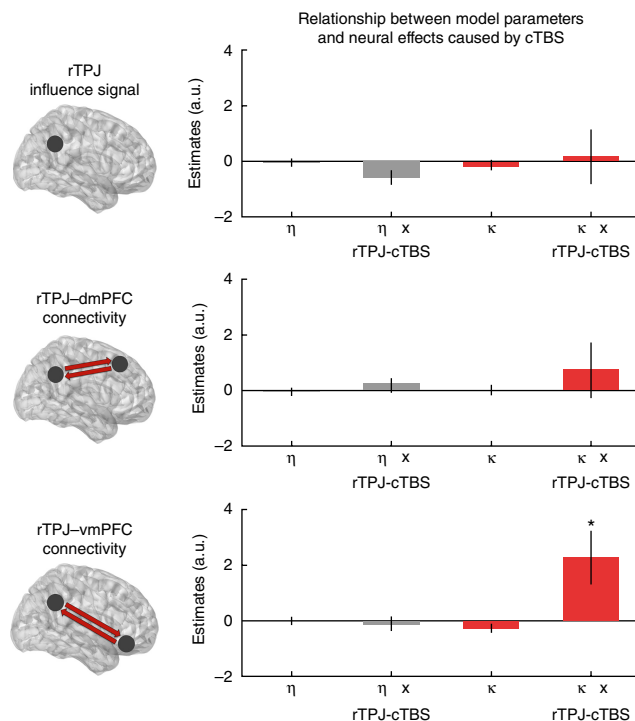

Figure 6 Relationship between neural effects of cTBS and model parameters. Regression analysis of the subject-level neural betas displayed in **Figure 3b** (top), **Figure 4b** (middle) and **Figure 5b** (bottom) on the κ and its interaction with stimulation condition (red). We added the first-order learning rate η and its interaction with stimulation condition as control variables (gray). Error bars indicate robust standard errors. Asterisk indicates the significant interaction of κ and stimulation condition ($t_{51} = 2.34$, *$P = 0.023$) stimulation condition.

To explicitly test this link, we perform three regression analyses in which we related the parameters of our behavioral model to the three neural effects of rTPJ-cTBS (**Fig. 6**; see Online Methods for details on the regressions). This revealed that a stronger cTBS disruption of influence-update-dependent connectivity between rTPJ and vmPFC was associated with a lower use of second-order beliefs (relationship between κ and rTPJ–vmPFC connectivity in the rTPJ-cTBS condition: $t_{25} = 2.84$, $P = 0.009$; Bonferroni $\alpha = 0.05/3$; interaction of κ and cTBS condition: $t_{51} = 2.34$, $P = 0.023$). None of the other two neural effects were significantly related to the behavioral changes induced by cTBS (rTPJ-influence signal: main effect of κ in rTPJ-cTBS condition, $t_{25} = 0.02$, $P = 0.97$; interaction of κ and cTBS condition, $t_{51} = 0.16$, $P = 0.87$; rTPJ–dmPFC connectivity: main effect of κ in rTPJ-cTBS condition, $t_{25} = 0.86$ $P = 0.39$; interaction of κ and cTBS condition, $t_{51} = 0.7$, $P = 0.48$). This suggests that a primary pathway by which rTPJ-cTBS affects subsequent behavior is via the communication of rTPJ influence update signals with vmPFC at the time of feedback.

DISCUSSION
The right temporoparietal junction is causally involved in strategic choice

Our results show that disrupting the neural influence computations in rTPJ by cTBS led to a deficit in the ability to optimally integrate the consequences of one's own actions on the opponent's behavior. The observed deficit is precisely aligned with the theoretical predictions of the model and provides evidence that the influence algorithm validly characterizes the aggregate output of the computations

performed by rTPJ. Failure to incorporate such second-order beliefs into choice means disregarding the opponent's learning, thereby rendering the player's behavior nonstrategic. Our results therefore provide direct evidence that intact rTPJ function is a neurobiological prerequisite for strategic social decision-making because of rTPJ's role in simulating the beliefs of other agents by means of the influence-update computation.

Previous research has ascribed a causal function to rTPJ for self-related processing[27–29] and for processing perpetrator beliefs in moral judgments[30]. However, the interplay between self-generated signals and simulated beliefs, or how one's action affects the beliefs of another agent, has not been addressed in past research. The present study fills this gap by providing causal evidence that the rTPJ is involved in this process.

A recent theory posits that the angular gyrus, which is part of TPJ, supports the construction of a specifically social context when this is relevant for behavior[31]. Another theory proposes that the computations occurring in the TPJ represent domain-general contextual updating of any internal model of the environment, be it social or not[32]. The predictions of both theories therefore differ in the degree of social specificity attributed to the region. The results of our study are in fact congruent with key predictions of both these theoretical frameworks, since they provide causal evidence that neural processes in angular gyrus (an anatomically-defined region comprised in TPJ) were indeed required for updating an agent's model of how strategically sophisticated opponents learn about the agent's own behavior.

Deficits in theory of mind are a defining feature of autism spectrum disorder[33], but the relevant measurement instruments (for example, the false belief task) do not provide quantitative measurements of the underlying neurocognitive operations[34]. In contrast, the strength of the computational approach to theory of mind is that it allows a fine-grained deconstruction of the involved subprocesses and associated neural signals[35]. The present findings suggest that the computational model employed here may be used to characterize putative abnormalities in neural coding associated with autism spectrum disorder and to ultimately contribute to advances in classification and treatment of this disorder.

The right temporoparietal junction interacts with dorsomedial and ventromedial prefrontal cortex

Functional connectivity between rTPJ and dmPFC is well documented and both regions have been proposed as belonging to a common functional network[5]. Here we show that perturbing computations occurring in rTPJ causally affected the coupling between both regions, as evidenced by their reduced connectivity at the time of feedback. This result therefore provides direct evidence that dmPFC and rTPJ are part of a functionally interacting network.

It has been proposed that the valuation system integrates decision-relevant information provided by the mentalizing network[2,15,22], but the evidence for this until now has been purely correlative. Here we show that disrupting rTPJ processing caused a change in how the region's connectivity to vmPFC represented the information that coded for the change in the opponent's model. This pattern of disrupted connectivity occurred in the portion of vmPFC that is most sensitive to overall value in our task, suggesting that these network-level perturbations affected the way changes in the opponent's model were integrated into choice value. Taken together, these results support an integrative model of social choice in which TPJ and dmPFC provide crucial inputs to the computation of choice value in vmPFC. More broadly, the effects we observed were generally consistent with findings of TMS-induced interference with

value processing following stimulation of other areas, such as dorsolateral prefrontal cortex[36]. Our data thus generally concur with proposals that the vmPFC may integrate information from multiple brain areas by means of functional coupling to compute value signals driving choices[22,37,38].

Implications for decision neuroscience

The use of computational models in functional neuroimaging has boomed in the past decade[39], with the ambitious aim of providing a mechanistic understanding of how choices are generated by the brain. To date, much of the effort has been focused on using fMRI to discover novel computational signals correlating with choices, but this approach is not without pitfalls. For instance, the correlation of one model's parameters with BOLD signals does not account for the fact that there are multiple plausible instantiations of a model's algorithms[9], and different competing models are rarely directly tested against each other or assessed for stability by direct replications[40]. Moreover, the interdependency between the different computational components coded in distinct functional areas cannot be assessed in the absence of a causal intervention.

In this context, the present results demonstrate two advantages of combined TMS–fMRI studies. First, the ability to localize and interfere with a given model-derived signal allows tests of causal mappings between model components and neural signals, therefore directly validating assumptions of such models that cannot be assessed with BOLD signal correlations alone. Second, combining both methods allows study of how localized perturbations spread through networks and alter neural representations in remote areas, providing a new perspective on how different model-proposed computations are neurally integrated to ultimately drive choice.

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

AUTHOR CONTRIBUTIONS
C.A.H. designed the study, collected data, performed analysis and wrote manuscript. S.S. performed analysis and contributed to the manuscript. R.P. performed analysis and contributed to the manuscript. M.M. contributed to TMS–fMRI procedure, collected data and contributed to the manuscript. J.P.O. designed the study and contributed to the manuscript. C.C.R. designed the study, contributed to data analysis and wrote the manuscript.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Darwin, C. *On the Origins of Species by Means of Natural Selection* (Murray, 1859).
2. Hampton, A.N., Bossaerts, P. & O'Doherty, J.P. Neural correlates of mentalizing-related computations during strategic interactions in humans. *Proc. Natl. Acad. Sci. USA* **105**, 6741–6746 (2008).
3. Devaine, M., Hollard, G. & Daunizeau, J. The social Bayesian brain: does mentalizing make a difference when we learn? *PLoS Comput. Biol.* **10**, e1003992 (2014).
4. Camerer, C., Ho, T.-H. & Chong, J.-K. A cognitive hierarchy model of games. *Q. J. Econ.* **119**, 861–898 (2004).

5. Schurz, M., Radua, J., Aichhorn, M., Richlan, F. & Perner, J. Fractionating theory of mind: a meta-analysis of functional brain imaging studies. *Neurosci. Biobehav. Rev.* **42**, 9–34 (2014).
6. Behrens, T.E.J., Hunt, L.T. & Rushworth, M.F.S. The computation of social behavior. *Science* **324**, 1160–1164 (2009).
7. Bhatt, M.A., Lohrenz, T., Camerer, C.F. & Montague, P.R. Neural signatures of strategic types in a two-person bargaining game. *Proc. Natl. Acad. Sci. USA* **107**, 19720–19725 (2010).
8. Carter, R.M., Bowling, D.L., Reeck, C. & Huettel, S.A. A distinct role of the temporal-parietal junction in predicting socially guided decisions. *Science* **337**, 109–111 (2012).
9. Mars, R.B., Shea, N.J., Kolling, N. & Rushworth, M.F.S. Model-based analyses: Promises, pitfalls, and example applications to the study of cognitive control. *Q J Exp Psychol (Hove)* **65**, 252–267 (2012).
10. Hunt, L.T. *et al.* Mechanisms underlying cortical activity during value-guided choice. *Nat. Neurosci.* **15**, 470–476, S1–S3 (2012).
11. Huang, Y.Z., Edwards, M.J., Rounis, E., Bhatia, K.P. & Rothwell, J.C. Theta burst stimulation of the human motor cortex. *Neuron* **45**, 201–206 (2005).
12. Ruff, C.C., Driver, J. & Bestmann, S. Combining TMS and fMRI: from 'virtual lesions' to functional-network accounts of cognition. *Cortex* **45**, 1043–1049 (2009).
13. Ruff, C.C. *et al.* Concurrent TMS-fMRI and psychophysics reveal frontal influences on human retinotopic visual cortex. *Curr. Biol.* **16**, 1479–1488 (2006).
14. Coricelli, G. & Nagel, R. Neural correlates of depth of strategic reasoning in medial prefrontal cortex. *Proc. Natl. Acad. Sci. USA* **106**, 9163–9168 (2009).
15. De Martino, B., O'Doherty, J.P., Ray, D., Bossaerts, P. & Camerer, C. In the mind of the market: theory of mind biases value computation during financial bubbles. *Neuron* **79**, 1222–1231 (2013).
16. Amodio, D.M. & Frith, C.D. Meeting of minds: the medial frontal cortex and social cognition. *Nat. Rev. Neurosci.* **7**, 268–277 (2006).
17. Clithero, J.A. & Rangel, A. Informatic parcellation of the network involved in the computation of subjective value. *Soc. Cogn. Affect. Neurosci.* **9**, 1289–1302 (2014).
18. Zhu, L., Mathewson, K.E. & Hsu, M. Dissociable neural representations of reinforcement and belief prediction errors underlie strategic learning. *Proc Natl. Acad Sci USA* **109**, 1419–1424 (2012).
19. Nicolle, A. *et al.* An agent independent axis for executed and modeled choice in medial prefrontal cortex. *Neuron* **75**, 1114–1121 (2012).
20. Sack, A.T., Camprodon, J.A., Pascual-Leone, A. & Goebel, R. The dynamics of interhemispheric compensatory processes in mental imagery. *Science* **308**, 702–704 (2005).
21. van den Bos, W., Talwar, A. & McClure, S.M. Neural correlates of reinforcement learning and social preferences in competitive bidding. *J. Neurosci.* **33**, 2137–2146 (2013).
22. Hare, T.A., Camerer, C.F., Knoepfle, D.T. & Rangel, A. Value computations in ventral medial prefrontal cortex during charitable decision making incorporate input from regions involved in social cognition. *J. Neurosci.* **30**, 583–590 (2010).
23. Haroush, K. & Williams, Z.M. Neuronal prediction of opponent's behavior during cooperative social interchange in primates. *Cell* **160**, 1233–1245 (2015).
24. Suzuki, S. *et al.* Learning to simulate others' decisions. *Neuron* **74**, 1125–1137 (2012).
25. Wunderlich, K., Dayan, P. & Dolan, R.J. Mapping value based planning and extensively trained choice in the human brain. *Nat. Neurosci.* **15**, 786–791 (2012).
26. Tricomi, E., Balleine, B.W. & O'Doherty, J.P. A specific role for posterior dorsolateral striatum in human habit learning. *Eur. J. Neurosci.* **29**, 2225–2232 (2009).
27. Tsakiris, M., Costantini, M. & Haggard, P. The role of the right temporo-parietal junction in maintaining a coherent sense of one's body. *Neuropsychologia* **46**, 3014–3018 (2008).
28. Wold, A., Limanowski, J., Walter, H. & Blankenburg, F. Proprioceptive drift in the rubber hand illusion is intensified following 1 Hz TMS of the left EBA. *Front. Hum. Neurosci.* **8**, 390 (2014).
29. Heinisch, C., Dinse, H.R., Tegenthoff, M., Juckel, G. & Brüne, M. An rTMS study into self-face recognition using video-morphing technique. *Soc. Cogn. Affect. Neurosci.* **6**, 442–449 (2011).
30. Young, L., Camprodon, J.A., Hauser, M., Pascual-Leone, A. & Saxe, R. Disruption of the right temporoparietal junction with transcranial magnetic stimulation reduces the role of beliefs in moral judgments. *Proc. Natl. Acad. Sci. USA* **107**, 6753–6758 (2010).
31. Carter, R.M. & Huettel, S.A. A nexus model of the temporal-parietal junction. *Trends Cogn. Sci.* **17**, 328–336 (2013).
32. Geng, J.J. & Vossel, S. Re-evaluating the role of TPJ in attentional control: contextual updating? *Neurosci. Biobehav. Rev.* **37**, 2608–2620 (2013).
33. Baron-Cohen, S. Theory of mind and autism: a review. *Int. Rev. Res. Ment. Retard.* **23**, 169–184 (2000).
34. Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y. & Plumb, I. The "Reading the Mind in the Eyes" test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. *J. Child Psychol. Psychiatry* **42**, 241–251 (2001).
35. Schaafsma, S.M., Pfaff, D.W., Spunt, R.P. & Adolphs, R. Deconstructing and reconstructing theory of mind. *Trends Cogn. Sci.* **19**, 65–72 (2015).
36. Baumgartner, T., Knoch, D., Hotz, P., Eisenegger, C. & Fehr, E. Dorsolateral and ventromedial prefrontal cortex orchestrate normative choice. *Nat. Neurosci.* **14**, 1468–1474 (2011).
37. Smith, D.V., Clithero, J.A., Boltuck, S.E. & Huettel, S.A. Functional connectivity with ventromedial prefrontal cortex reflects subjective value for social rewards. *Soc. Cogn. Affect. Neurosci.* **9**, 2017–2025 (2014).
38. Ruff, C.C. & Fehr, E. The neurobiology of rewards and values in social decision making. *Nat. Rev. Neurosci.* **15**, 549–562 (2014).
39. Johnson, E.J. & Ratcliff, R. Computational and process models of decision making in psychology and behavioral economics. in *Neuroeconomics: Decision Making and the Brain* (eds. Gimlicher, P.W. & Fehr, E.) 358–365 (Academic Press, 2014).
40. Button, K.S. *et al.* Power failure: why small sample size undermines the reliability of neuroscience. *Nat. Rev. Neurosci.* **14**, 365–376 (2013).
41. Spiegelhalter, D.J., Best, N.G., Carlin, B.P. & van der Linde, A. Bayesian measures of model complexity and fit. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **64**, 583–639 (2002).

ONLINE METHODS

Participants. 120 healthy volunteers (60 male, 60 female; ages 18–25) took part in the experiment. Sixty participants were scanned with functional magnetic resonance imaging while playing the inspection game in the role of Employee, paired with an anonymous Employer (the other 60 participants) sitting in an adjacent room and playing on a computer connected via network protocols. This sample size was a good compromise between the substantial effort and costs involved in the experiment (fMRI + cTBS with two interacting subjects) and the degree of statistical sensitivity (80% power to detect a large effect size; Cohen's $d = 0.75$) chosen based on the clear a priori hypothesis originating from the previous study[2]. All scanned participants were right-handed nonsmokers, medication-free, with normal eyesight, no history of neurological disorders and in full compliance with safety guidelines for transcranial magnetic stimulation. Half of the scanned participants received cTBS to the vertex while the other half received cTBS to the rTPJ. One participant in the rTPJ-cTBS condition and one participant in the vertex-cTBS condition reported feeling anxious and/or dizzy during the cTBS procedures, leading to the immediate interruption of the experiment. Additionally, fMRI imaging data for one participant in the rTPJ-cTBS condition were lost due to a scanner database error. Two participants in the vertex-cTBS and one in the rTPJ-cTBS provided incomplete answers to the pre- and postexperimental questionnaires (**Supplementary Table 1**).

Ethics compliance and blinding. The Ethics Commission of Canton Zürich approved the experiment. Informed consent was obtained from each subject before the experiment. Data collection and analysis were not performed blind to the conditions of the experiment due to the nature of TMS interventions.

Experimental design and task. Since the focus of our study was on learning the impact of the subject's own actions on the opponent, it was absolutely critical that we tested participants while they had not yet had extensive time to learn about their specific opponent and thus need to update their expectations. We therefore avoided any experimental designs that would enable subjects to gain extensive experience with a specific opponent (for example, within-subject or pre–post designs) and opted for a between-subjects design with minimal training before the TMS intervention.

Between-subject designs make it difficult to explicitly account for any possible pre-existing group differences. We thus took several measures to mitigate the risk that our two groups differed in game performance at baseline. First, our design entailed fully random group assignment, which should lead to matched samples and does not bias the results in a specific way. Second, we acquired—before the TMS intervention—measures of various cognitive functions and preferences known to be critical for choices during strategic social interactions, while precluding any training with the specific opponent in the experimental game. More specifically, we measured ability in working memory capacity with the N-back task[42], depth of strategic thinking with the beauty contest[14,43] and preferences for competition versus cooperation with the SVO (social-value orientation) questionnaire[44]. Working memory was critical in our study task because participants had to maintain representations of their choices and their opponent's choice history. The beauty contest measures how many iterative steps of metalizing-related reasoning a subject may perform and how well the subject can adequately predict the average depth of strategic thinking in the opponent population. Preferences for cooperative versus competitive outcomes may influence choices because subjects may be averse to winning on a majority of trials, therefore allowing themselves to lose on purpose in order to obtain a financial outcome comparable to their opponent. Notably, the participants in both groups did not differ in any of these three measures (**Supplementary Table 1**). This lack of any group difference in key abilities and preferences important for the inspection game strongly suggests that the randomization was effective in matching the groups for baseline ability.

Before the experimental measurements in both the vertex-cTBS and rTPJ-cTBS group, subjects played 10 trials of the inspection game to familiarize themselves with the task structure and response options. Following the training phase, participants underwent cTBS and entered the scanner, where they played 160 trials of the inspection game in two sessions of 80 trials each. Each trial began with a fixation-cross presented for a randomly determined period between 1.25 and 8 s, followed by a decision screen featuring two pictograms describing choice options. Both participants played simultaneously and had up to 2 s to make their decision. Choices were confirmed with a red square, which was displayed for a minimum of 100 ms and up to the time it took for both answers to be recorded. After each player made his or her decision, a fixation cross was presented for a random period between 1.25 and 8 s, after which the feedback screen was displayed for 2.5 s. The average duration of a trial summed to 9 s (**Fig. 1c**). fMRI contrast sensitivity was optimized a priori by simulating effects of task ITI on statistical sensitivity with custom code. In total, the task lasted 26 min.

Data acquisition: transcranial magnetic stimulation. *Stimulation site.* We performed an automated meta-analysis for the keyword "theory mind" using Neurosynth[45], which yielded a synthesis from 140 studies and 6,092 reported activations. From the reverse-inference mask at the default threshold, we extracted all contiguous voxels of the mask for rTPJ. We employed Neurosynth since it is a comprehensive and up-to-date source of activations and has been successfully used in recent TPJ research[33]. For the functional map, we extracted the activation cluster (at $P = 0.001$) correlating with the influence-update signal in the control group. The functional and meta-analysis maps showed a large degree of overlap (72% of the former contained in the latter). We derived a binary cluster corresponding to the overlap of both maps. Next, we calculated the centroid of that cluster. The centroid was then projected onto the nearest surface gyrus and the resulting surface site (MNI x, y, $z = 61$, -52, 35) was used as stimulation site. Anatomically, the stimulation site was located on the anterior portion of Brodmann area 39. For our control group, we selected the vertex as the stimulation site, defined for each participant in their own T1-weighted MRI scan as the intersection of the central sulci from both cerebral hemispheres. Stimulating the vertex is unlikely to have any task-relevant effects for a number of reasons. First, the stimulated vertex site lies at a considerable distance from the TMS coil due to the folding of the two cerebral hemispheres, thereby reducing the effectiveness of the TMS pulse. Second, the site corresponds to the somatosensory and motor representation of the leg. We did not observe any sensations or leg movements during the stimulation, supporting the view that the control stimulation did not result in effective neural stimulation. Last but not least, this area is not known to be involved in the computations underlying complex social behavior as studied in our design (see also Hampton *et al.*[2]).

Transcranial magnetic stimulation procedure. Stimulation coordinates were superimposed onto native T1 images for each participant using inverse normalization with trilinear interpolation implemented in SPM8 (Wellcome Trust Centre for Neuroimaging). Passive and active motor thresholds were obtained for each subject by means of neuronavigated single-pulse TMS on the portion of motor cortex controlling movement of the index finger, using the Brainsight interface (Rogue Resolutions Ltd., Cardiff, UK). The passive motor threshold was defined as the single-pulse intensity required to elicit motor evoked potentials (MEP; > 200 mV in amplitude) in at least five of ten pulses as measured by electromyography in Brainsight. The active motor threshold was obtained by having the subject exert constant pressure between the index finger and the thumb (20% of maximum force) while the thresholding procedure described above was repeated. For the rTPJ-cTBS condition, the stimulation intensity was proportionally scaled to the difference between the skull–brain distance of motor cortex and the skull–brain distance of the stimulation site. This experiment employed the standard continuous theta-burst stimulation protocol[11] calibrated at 80% of active motor threshold. Stimulation was applied on the fMRI sliding bed over the coordinates identified with Brainsight with the coil orientation marked on a fixed cap. The induced electromagnetic field was aligned perpendicular to the crown of the angular gyrus for the rTPJ-cTBS condition, with the coil handle pointing backwards and angled accordingly. An fMRI-compatible coil was used to administer the stimulation (Magventure MRi-B91). The cTBS stimulation protocol comprised 600 pulses administered over 40 s in bursts of three pulses at 50 Hz (20 ms), repeated at intervals of 5 Hz (200 ms) in accordance with the guidelines prescribed by the literature[11]. fMRI procedures commenced directly after the stimulation, with the average time between the offset of the stimulation and the onset of the fMRI task corresponding to 223 s (s.d. = 9.94 s).

Data acquisition: fMRI recording and preprocessing. For each session, we acquired 282 T2*-weighted whole-brain echo planar images using a Philips Achieva 3T whole-body scanner (Philips Medical Systems, Best, The Netherlands) equipped with an 8-channel Philips sensitivity-encoded (SENSE) head coil.

Imaging parameters were: 2,600-ms repetition time (TR); 37 slices (transversal, ascending acquisition); 2.6-mm slice thickness; 2.5-mm × 2.5-mm in-plane resolution; 0.65-mm gap; 90° flip angle. After the magnetic field was fully equilibrated, five dummy image excitations were performed and discarded before functional image acquisition started. To enhance BOLD sensitivity throughout the brain, we employed a dual-echo sequence (TE: 17 ms and 44 ms). To combine both images, we used a weighted voxel-wise summation technique[46,47] to generate a single functional image. To do this, we first computed the signal-to-noise ratio for each voxel measured during a resting-state scan. We then weighted each voxel by a function of the voxel-based SNR derived from the resting state. Finally, we acquired a T1-weighted whole brain structural image with $1 \times 1 \times 1$-mm^3 cubic voxel size for each subject. All image preprocessing steps were applied using SPM8 (Wellcome Trust Centre for Neuroimaging) according to standard procedures employed in our group[48]. First, we performed slice-time correction to the middle slice acquisition time. We then realigned the images to account for subject head motion. Each subject's T1 image was co-registered to the functional image and normalized to the standard T1 MNI template with the new-segment procedure in SPM8[49]. Finally, the images were spatially resampled to 3-mm^3 isotropic voxels and smoothed using an 8-mm FWHM Gaussian kernel.

Data acquisition: peripheral measures. To control for the effects of heart rate and breathing on fMRI recordings, we acquired cardiac and respiratory signals using electrocardiography and a breathing belt. The physiological time series were transformed following the RETROICOR[50,51] procedure, which uses Fourier expansions of various orders for the phase and cardiac pulsation (third order), respiration (fourth order) and cardiorespiratory interaction (first order)[51]. The corresponding regressors were created using the Matlab PhysIO Toolbox[52] (open source code available as part of the TAPAS software collection; https://www.tnu.ethz.ch/de/software/tapas.html).

Data analysis: behavior. We employed three distinct computational models based on Hampton et al.[2] to account for our behavioral data. These models were fitted to the behavioral data of the Employees undergoing scanning using a hierarchical Bayesian procedure (see next section). The fitted models were used to perform inference on model parameters[53] and derive subject-level parametric modulators[54] (first-order prediction error, influence update and expected value of the chosen option) for the fMRI analysis.

Model 1: reinforcement learning. Reinforcement learning posits that an agent learns from past rewards which actions (for example, Shirk or Work) are more likely to produce future rewards. We used a simple Rescorla-Wagner based RL model, which updates action values V on trial t based on a prediction error δ reflecting which action was selected and whether or not it was followed by reward.

$$V_{t+1}^{Work} = V_t^{Work} + \eta \delta_t \quad (1)$$

The prediction error δ is calculated on each trial by comparing the experienced reward R_t with the expected reward. The prediction error is weighted by a learning rate η.

$$\delta_t = R_t - V_t^{Work} \quad (2)$$

To choose between two actions (for example, Shirk or Work) the model compares the expected value for both actions and biases the choice probability to favor the action with the greatest expected value, using a logistic sigmoid where β is the inverse temperature.

$$P(work) = \frac{1}{1 + exp^{-\beta(V_t^{work} - V_t^{shirk})}} \quad (3)$$

Model 2: fictitious play. In the fictitious play model, the player tracks their opponent's probability P of choosing an action (for example, Not Inspect) on trial t based on the opponent's history of play, in order to choose the action that maximizes the expected reward (for example, Shirk).

$$p_{t+1}^{Not\,Inspect} = p_t^{Not\,Inspect} + \eta \delta_t^P \quad (4)$$

The prediction error δ_t^P is updated as follows.

$$\delta_t^P = P_t - p_t^{Not\,Inspect} \quad (5)$$

where P_t is the observed action played by the opponent at trial t. It is equal to 1 if the employer played *Not Inspect* and equals 0 if the opponent played *Inspect*. In the fictitious play model, the expected value z of each action is derived from the game's payoff matrix.

For the Employee:

$$z_t^{Work} = 2 - 4p_t^{Not\,Inspect} \quad (6)$$

For the Employer:

$$z_t^{Not\,Inspect} = 5p_t^{Work} - 1 \quad (7)$$

To derive the probability of selecting an action, the expected value z is then fed into a softmax function similar to that given in equation (3), where β is the inverse-temperature.

$$P(work) = \frac{1}{1 + exp^{-\beta z_t^{Work}}} \quad (8) \quad (8)$$

Model 3: influence model. The influence model posits that the decision of an agent at trial t is a function of both the agent's history of choice and the opponent's history of choice. The agent is essentially tracking the fictitious play model of the opponent in addition to their own, expressed for the Employee as

$$\Delta P \approx -\eta 5\beta p^{Not\,Inspect}\left(1 - p^{Not\,Inspect}\right)\left(Q_t - q_t^{Work}\right) \quad (9)$$

where Q_t is the employee's action at trial t (1 if the Employee's action was Work), and q_t^{Work} is the Employer's inferred probability that the employee will Work. In other words, the Employee tracks the second order belief q_t^{Work} and updates it at each trial with his or her own action Q_t.

The weight given to these second-order beliefs is captured by the parameter κ, a constant that approximates the parameters of the opponent (learning rate η, temperature β and payoff matrix). The scaling factor $(1 - p^{Not\,Inspect})$ is assumed constant throughout the game, as in Hampton et al.[2], to make it comparable to learning rate parameter η and to stabilize the estimation of parameter κ.

$$\kappa \approx -\eta 5\beta p^{Not\,Inspect}\left(1 - p^{Not\,Inspect}\right) \quad (10)$$

with second-order belief q_t^{Work}

$$q_t^{Work} = \frac{1}{5} - \frac{1}{5\beta}\log\left(\frac{1 - p^{Not\,Inspect}}{p^{Not\,Inspect}}\right) \quad (11)$$

Hence, the influence model for the Employee can be expressed as

$$p_{t+1}^{Not\,Inspect} = p_t^{Not\,Inspect} + \eta\left(P_t - p_t^{Not\,Inspect}\right) + \kappa\left(Q_t - q_t^{Work}\right) \quad (12)$$

The model has the parameter η, reflecting the first-order learning rate and the parameter κ controlling the weight given to second-order beliefs. If $\kappa = 0$, then the model becomes fictitious play. For fictitious play, $p_{t+1}^{Not\,Inspect}$ is fed into the value function outlined in equation (6), which is incorporated into the softmax in equation (8).

Estimation procedure. We employed a Hierarchical Bayesian approach to estimate model parameters and perform model comparison (**Supplementary Fig. 1**). This approach has been recommended[54,55] as an alternative to the pooling of behavioral data (which poses statistical and conceptual issues[54]) and has been shown to improve the precision of subject-level parameter estimates[55].

The model defines the choice process by

$$Ch_{yee(s,(t+1))} \sim \text{dbernouilli}(p_{(s,t)})$$

where $p_{(s,t)}$ is the softmax output of equation (12) Moreover:

$$\eta_{(s)} \sim \text{dbeta}(\mu_\eta K_\eta, (1 - \mu_\eta)K_\eta)$$
$$\kappa_{(s)} \sim \text{dbeta}(\mu_\kappa K_\kappa, (1 - \mu_\kappa)K_\kappa) \text{ and}$$
$$\beta_{(s)} \sim \text{dbeta}(\mu_\beta K_\beta, (1 - \mu_\beta)K_\beta)$$

with hyperparameters

$\mu_\eta \sim \text{dbeta}(1,1)$
$K_\eta \sim \text{dunif}(0.001,100)$
$\mu_\kappa \sim \text{dbeta}(1,1)$
$K_\kappa \sim \text{dunif}(0.001,100)$
$\mu_\beta \sim \text{dbeta}(1,1)$ and
$K_\beta \sim \text{dunif}(0.001,100)$

Mixed-equilibrium model. In addition to our computational learning models, we fitted a mixed-equilibrium model (ME) to test whether our participants may have optimized their play patterns by removing predictability (i.e., by random responding). A mixed strategy for the inspection game posits that employees will work with probability $P(\text{Work})$; the mixed Nash equilibrium for the Employee in the inspection game corresponds to $P(\text{Work}) = 0.5$. This strategy prevents any form of exploitation, as established by theoretical accounts in game theory[56].

A simple hierarchical Bayesian model implementation of a mixed strategy assumes that choices Ch_{yee} are generated by a simple Bernoulli process:

$$Ch_{yee(s,(t+1))} \sim \text{dbernouilli}(\delta_{(s)}) \text{ with}$$
$$\delta_{(s)} \sim \text{dbeta}(\mu_\delta K_\delta, (1 - \mu_\delta)K_\delta)$$

where t denotes the trial and s denotes the subject in the hierarchy. We use mean and precision parameters of the beta distribution dbeta to capture the empirically observable mixed strategy, which can be compared to the mixed Nash Equilibrium of $P(\text{Work}) = 0.5$.

$\mu_\delta \sim \text{dbeta}(1,1)$
$K_\delta \sim \text{dunif}(0.001,100)$

This stochastic model captured the data considerably less well than the influence model (**Fig. 1c**). Moreover, the population-level posterior distribution of μ_δ indicated that participants did not exhibit a matched probability of $P = 0.5$ for playing either option, as would be expected if they employed a mixed Nash equilibrium strategy (**Supplementary Fig. 2**).

Controlling for the use of mixed strategies in both groups. Even though the model comparisons suggest that participants in both groups generally employed influence-update computations to control their behavior, it is nevertheless possible that a few individual subjects may have attempted to randomize their responses. To rule out the possibility that such 'outlier' individuals may have affected our cTBS results, we fitted a mixture model in which we statistically controlled for individual tendencies to randomize responses when analyzing cTBS effects on the influence parameter κ. This model assumes that each subject's behavior was generated by an individually weighted mixture of the influence model and a randomizer model in which $P(\text{Work})$ is defined as the empirical estimate of our ME model. The individual weight given to influence learning versus fully random responding was captured by a parameter λ in an expansion of the softmax choice rule equation (8):

$$P(\text{Work}) = \frac{1}{1 + exp^{-\beta z_t^{Work}}}(1 - \lambda) + \zeta\lambda$$

λ is drawn from a beta distribution:

$$\lambda_{(s)} \sim \text{dbeta}(\mu_\lambda K_\lambda, (1 - \mu_\lambda)K_\lambda) \tag{13}$$

with hyperparameters

$\mu_\lambda \sim \text{dbeta}(1,1)$
$K_\lambda \sim \text{dunif}(0.001,100)$

The parameter ζ represents the empirical mean of the population-level parameter δ in our mixed equilibrium model, averaged over both conditions ($\zeta = 0.42$). Note that this parameter cannot be fitted empirically in the mixture model, as it is not uniquely determined in presence of the mixture-model parameter λ.

The results of this control analysis indicate that the difference in κ between the rTPJ-cTBS and vertex-cTBS condition still holds ($P_{\text{mcmc}} = 0.03$) even when controlling for individual tendencies to randomize responses. Moreover, the λ parameter did not differ between conditions ($P_{\text{mcmc}} = 0.48$), indicating that our causal intervention did not significantly alter our subject's tendency to randomize their choices. Note that the model fits of the mixture model are largely inferior to those of the pure influence model (DIC difference with influence model = 155 for vertex-cTBS and 175 for rTPJ-cTBS). Thus, the influence model alone captures the behavior of our subjects better than the more complex mixture-based variant that also accounts for the use of mixed-strategies.

Bayesian hypothesis testing. We performed inference by inspecting the highest-density interval (HDI)[53]. Simply put, we estimated the influence model separately for vertex-cTBS and rTPJ-cTBS. The Monte -Carlo-based hierarchical approach yields a posterior estimate that represents the empirical distribution of parameter values at the population level (i.e., random effects analysis). To perform inference, we subtract the distribution of $\kappa_{\text{rTPJ-cTBS}}$ from the distribution of $\kappa_{\text{vertex-cTBS}}$ and place bounds on the resulting distribution that capture 95% (for $P > 0.05$) of the distribution density. If the lower bound does not include 0, we can reject the hypothesis that the parameters for both populations are equal. The reported MCMC P values represent the proportion of the distribution that is below the zero line. Note that when none of the HDIs overlaps, we report P values as $P_{\text{mcmc}} \cong 0$.

Non-Bayesian hypothesis testing. Group differences in the relationship between experimental factors and choice behavior (**Figs. 1** and **2**) were analyzed with standard OLS regression. Data distribution was assumed to be normal, but this was not formally tested. The statistics corresponding to the parameter estimates are reported in the results section as t values.

Data analysis: neuroimaging. *fMRI design matrix.* All fMRI data were analyzed with random-effect general linear models (GLMs) using SPM8. Our design matrix included the onset of the decision period with the expected value of the chosen options as parametric modulator and the onset of the feedback epoch with reward, first-order prediction error and influence update as parametric modulators. Orthogonalization was turned off in SPM8 to capture only the variance uniquely explained by each factor. In addition, we included as nuisance regressors all movement parameters and physiological fluctuations related to heart rate and breathing with the procedure described above. We constructed a boxcar epoch function with a duration corresponding to the reaction time of the decision, to optimally account for variability in decision time[57]. The feedback epoch was modeled using stick functions.

Neuroimaging analysis strategy and correction for multiple comparisons. For the neuroimaging analyses, we first describe the results of main effects across groups (to replicate Hampton *et al.*[2]), followed by comparisons between the groups to identify differences brought about by TMS. For the PPI analyses, we only focused on the difference between groups, as these analyses were designed to identify changes in connectivity due to an experimental factor (TMS in our case). For all these analyses, we first present the results in the ROIs selected a priori strictly based on Hampton *et al.*[2], followed by exploratory whole-brain analyses to identify other areas than the a priori ROIs that show similar effects. All whole-brain analyses employed a cluster-level threshold of FWE $P < 0.05$ (cluster-forming threshold $t = 2.66$), corrected for multiple comparisons across the whole brain. We implemented these analyses as nonparametric tests (5,000 permutations, no t-map smoothing) in the software package SnPM (open source; code available at http://warwick.ac.uk/snpm) to optimally correct for type-1 error rates[58].

Analyses of ROIs employed a peak-level threshold of $P(\text{FWE}) < 0.05$, corrected for the ROI search volume. The rTPJ ROI for the analysis of the rTPJ-cTBS effects was defined as the empirical activation cluster obtained from the separate dataset acquired in the control group ($Ke = 487$). The ROIs for the vmPFC, dmPFC and bilateral pSTS were defined as spheres with a radius of 16 mm (twice the smoothing kernel) centered on coordinates reported

previously to correspond to activations for value-based choice (x, y, z = 0, 32, −13 for vmPFC[25,26]) and influence model representations (x, y, z = −3, 51, 25 for dmPFC[2,16]; x, y, z = −57, −54, 0 for left pSTS[2] and x, y, z = 60, −54, 9 for right pSTS[2]).

Psychophysiological analysis (PPI). Using the gPPI approach[59], we included in our design matrix (i) the BOLD time-series extracted from a 16-mm sphere centered on the subject-specific peak representation of the influence-update signal within our stimulated area (defined as all coordinates in a sphere of 8 mm radius around the stimulation site), (ii) all regressors defined in the section "fMRI Design Matrix" above and (iii) the interaction terms of the extracted BOLD time-course for all regressors, to measure the unique effect of the interaction of interest.

Regression models to link cTBS effects on behavior and neural function. We performed three separate analyses in which we regressed the three analyzed aspects of neural activity (beta estimates from the fMRI analyses for the influence update signal, rTPJ–dmPFC connectivity and rTPJ–vmPFC connectivity) on both the TMS intervention and on different aspects of behavior (as captured by model parameters). For all these analyses, we cautiously employed a robust regression to optimally handle the variability present in the neural data. We regressed the beta estimates from these neural cTBS effects on κ and its interaction with the cTBS condition, to capture relationships between behavioral and neural cTBS effects. To further ensure that these results were specific to the affected parameter, we added the first-order learning rate η and its interaction as control variables.

Main effect models for data from the rTPJ-cTBS condition are defined as

$$\text{rTPJ} = \beta_0 + \beta_2\eta + \beta_3\kappa + \varepsilon$$
$$\text{dmPFC} = \beta_0 + \beta_2\eta + \beta_3\kappa + \varepsilon$$
$$\text{vmPFC} = \beta_0 + \beta_2\eta + \beta_3\kappa + \varepsilon$$

Interaction models for the full dataset are defined as

$$\text{rTPJ} = \beta_0 + \beta_1(\text{cTBS}) + \beta_2\eta + \beta_3\kappa + \beta_4(\text{cTBS})\eta + \beta_5(\text{cTBS})\kappa + \varepsilon$$
$$\text{dmPFC} = \beta_0 + \beta_1(\text{cTBS}) + \beta_2\eta + \beta_3\kappa + \beta_4(\text{cTBS})\eta + \beta_5(\text{cTBS})\kappa + \varepsilon$$
$$\text{vmPFC} = \beta_0 + \beta_1(\text{cTBS}) + \beta_2\eta + \beta_3\kappa + \beta_4(\text{cTBS})\eta + \beta_5(\text{cTBS})\kappa + \varepsilon$$

where 'rTPJ' are the beta estimates of the neural effects of cTBS on rTPJ presented in **Fig. 3b**, 'dmPFC' are the PPI beta estimates presented in **Fig. 4b** and 'vmPFC' are the PPI beta estimates presented in **Fig. 5b**. 'cTBS' is a dummy coded variable (0 for vertex-cTBS; 1 for rTPJ-cTBS); η denotes the first-order learning rate and κ denotes the influence parameter. Because we had no a priori reason to believe the relationship should be present for one of these neural effects but not the other, we employed Bonferroni corrections for multiple comparisons

to adjust our α to the number of tests we performed. Hence, our α threshold for inference was set to $P = 0.05/3$.

Data and code availability. Raw behavioral data, Matlab and R code for reconstruction of all figures, computational models and statistical analysis (including fMRI t-maps and ROIs) are available online (https://doi.org/10.5281/zenodo.808428). Raw fMRI data and processing pipelines are available upon reasonable request. A **Supplementary Methods Checklist** is available.

42. Owen, A.M., McMillan, K.M., Laird, A.R. & Bullmore, E. N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. *Hum. Brain Mapp.* **25**, 46–59 (2005).

43. Bosch-Domènech, A., Montalvo, J.G., Nagel, R. & Satorra, A. One, two, (three), infinity, …: Newspaper and lab beauty-contest experiments. *Am. Econ. Rev.* **92**, 1687–1701 (2002).

44. Murphy, R.O., Ackermann, K.A. & Handgraaf, M. Measuring Social Value Orientation. *Judgm. Decis. Mak.* **6**, 771–781 (2011).

45. Yarkoni, T., Poldrack, R.A., Nichols, T.E., Van Essen, D.C. & Wager, T.D. Large-scale automated synthesis of human functional neuroimaging data. *Nat. Methods* **8**, 665–670 (2011).

46. Posse, S. *et al.* Enhancement of BOLD-contrast sensitivity by single-shot multi-echo functional MR imaging. *Magn. Reson. Med.* **42**, 87–97 (1999).

47. Schmiedeskamp, H. *et al.* Improvements in parallel imaging accelerated functional MRI using multiecho echo-planar imaging. *Magn. Reson. Med.* **63**, 959–969 (2010).

48. Grueschow, M., Polania, R., Hare, T.A. & Ruff, C.C. Automatic versus Choice-Dependent Value Representations in the Human Brain. *Neuron* **85**, 874–885 (2015).

49. Ashburner, J. & Friston, K.J. Unified segmentation. *Neuroimage* **26**, 839–851 (2005).

50. Harvey, A.K. *et al.* Brainstem functional magnetic resonance imaging: disentangling signal from physiological noise. *J. Magn. Reson. Imaging* **28**, 1337–1344 (2008).

51. Hutton, C. *et al.* The impact of physiological noise correction on fMRI at 7T. *Neuroimage* **57**, 101–112 (2011).

52. Kasper, L. *et al.* Cardiac artefact correction for human brainstem fMRI at 7T. *Neuroimage* **47**, S100 (2009).

53. Kruschke, J.K. What to believe: Bayesian methods for data analysis. *Trends Cogn. Sci.* **14**, 293–300 (2010).

54. Daw, N.D. in *Decision Making, Affect, and Learning. Attention and Performance XXIII* (eds. Delgado, M.R., Phelps, E.A. & Robbins, T.W.) 555 (Oxford University Press, 2011).

55. Ahn, W.Y., Krawitz, A., Kim, W., Busmeyer, J.R. & Brown, J.W. A model-based fMRI analysis with hierarchical Bayesian parameter estimation. *J. Neuroscience Psychology Econ.* **4**, 95–110 (2011).

56. Fudenberg, D. & Tirole, J. *Game Theory* (MIT Press, 1991).

57. Grinband, J., Wager, T.D., Lindquist, M., Ferrera, V.P. & Hirsch, J. Detection of time-varying signals in event-related fMRI designs. *Neuroimage* **43**, 509–520 (2008).

58. Eklund, A., Nichols, T.E. & Knutsson, H. Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proc. Natl. Acad. Sci. USA* **113**, 7900–7905 (2016).

59. McLaren, D.G., Ries, M.L., Xu, G. & Johnson, S.C. A generalized form of context-dependent psychophysiological interactions (gPPI): a comparison to standard approaches. *Neuroimage* **61**, 1277–1286 (2012).