Neural correlates of mentalizing-related computations during strategic interactions in humans

Alan N. Hampton[†], Peter Bossaerts[†‡], and John P. O'Doherty[†‡§]

[†]Computation and Neural Systems Program and [‡]Division of Humanities and Social Sciences, California Institute of Technology, 1200 East California Boulevard, M/C 228-77, Pasadena, CA 91125

Edited by Edward E. Smith, Columbia University, New York, NY, and approved February 20, 2008 (received for review November 22, 2007)

Competing successfully against an intelligent adversary requires the ability to mentalize an opponent's state of mind to anticipate his/her future behavior. Although much is known about what brain regions are activated during mentalizing, the question of how this function is implemented has received little attention to date. Here we formulated a computational model describing the capacity to mentalize in games. We scanned human subjects with functional MRI while they participated in a simple two-player strategy game and correlated our model against the functional MRI data. Different model components captured activity in distinct parts of the mentalizing network. While medial prefrontal cortex tracked an individual's expectations given the degree of model-predicted influence, posterior superior temporal sulcus was found to correspond to an influence update signal, capturing the difference between expected and actual influence exerted. These results suggest dissociable contributions of different parts of the mentalizing network to the computations underlying higher-order strategizing in humans.

computational modeling | decision making | functional MRI | neuroeconomics

Humans, like many other primates, live in a highly complex social environment in which it is often necessary to interact with, and compete against, other individuals to attain reward. Success against an intelligent adversary with competing objectives likely depends on the capacity to infer the opponent's state of mind, to predict what action the opponent is going to select in future, and to understand how an individual's own actions will modify and influence the behavior of one's opponent. This ability is often referred to as "mentalizing" and has been linked to a number of specific regions thought to be specifically engaged when processing socially relevant stimuli, and especially when inferring the state of minds of others (1). Neuroimaging studies in humans have implicated a specific network of brain regions including dorsomedial prefrontal cortex (PFC), posterior superior temporal sulcus (STS), and the temporal poles (2, 3) while subjects engage in tasks relevant to mentalizing, such as evaluating false beliefs or social transgressions (4, 5), describing the state of biological movements (6–8), and playing interactive games (9–11). However, although these studies have provided insight into what brain regions may be involved in the capacity to mentalize, the question of how this function is implemented at the neural level has received relatively little attention to date.

The goal of the present study was to build a simple model describing computations underlying the capacity to mentalize (in the context of a strategic game) and to determine whether different components of this model were correlated with neural activity in parts of the mentalizing network. To assess competitive interactions experimentally, we studied pairs of human subjects while they played each other in a two-player strategic game called the "inspection" game (or generalized matching pennies), in which opponents have competing goals (Fig. 1 *A* and *B*). One of the players was being scanned with functional MRI (fMRI), and the opponent was playing outside the scanner. The "employer" could either "inspect" or "not inspect," and the "employee" could either "work" or "shirk." The employer received 100 cents if he/she did not inspect and the employee worked and 25 cents if he/she inspected and caught the employee shirking. Otherwise he/she got zero cents. In contrast, the employee got 50 cents for working when the employer inspected and for shirking when the employer did not inspect, otherwise getting zero cents as well. Both players had competing objectives, in that when one player won in a given trial, the other one lost.

A player can in principle use a number of different strategies to try to win in such a game. Perhaps the simplest strategy is on each trial to simply choose the action that in the recent past gave the most reward. This strategy is referred to as reinforcement learning (RL) and approximates the optimal solution for many different types of decision problem in nonstrategic contexts, even for decision problems with complex higher-order structure whereby such structure can be accommodated by a sufficiently nuanced model of the state-space and transition probabilities (12, 13). However, such a strategy would be devastating for an individual in a competitive scenario because a clever opponent could detect the regularity in the reinforcement learner's choices to work out what action the reinforcement learner is going to choose next and exploit that knowledge by choosing the confounding response.

A more sophisticated approach is to try to predict the opponent's next actions by taking into account the history of prior actions by the opponent and then choosing the best response to that predicted action, a strategy known as "fictitious play" (14–16). A fictive learner is, in contrast to a reinforcement learner, employing an elementary form of mentalizing, because they are engaging a representation of the actions and intentions of their opponent.

However, an even more cognitively sophisticated and Machiavellian strategy a player could use in this game is to not only track the opponent's actions, but also to incorporate knowledge of how one's own actions influence the opponent's strategy. Simply put, this involves a player's building a prediction of what the opponent will do in response to the player's own actions. For example, the more the employer inspects, the higher the probability the employee will work in subsequent trials. The employer can then use this knowledge to make choices with higher expected rewards in subsequent trials, i.e., not inspect. We will term this strategy the "influence" learning model (see Table 1 for a comparison of the different models).

Although the behavioral game theory literature has demonstrated that humans think strategically in one-shot games [they consider what the opponent could possibly believe (17, 18)], the updating of beliefs about the beliefs of the opponent has rarely been incorporated in the analysis of learning in repeated play. Our modeling approach differs from the two exceptions (19, 20) in both

Author contributions: A.N.H., P.B., and J.P.O. designed research; A.N.H. performed research; A.N.H. analyzed data; and A.N.H., P.B., and J.P.O. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[§]To whom correspondence should be addressed. E-mail: jdoherty@hss.caltech.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0711099105/DCSupplemental.

Table 1. Model update rules

Model	Update rule
RL	$V^a_{t+1} = V^a_t + \eta(R_t - V^a_t)$
Fictitious	$p^*_{t+1} = p^*_t + \eta(P_t - p^*_t)$
Influence	$p^*_{t+1} = p^*_t + \eta(P_t - p^*_t) - \kappa(Q_t - q^{**}_t)$

The RL model updates the value of the chosen action a with a simple Rescorla–Wagner (35) prediction error ($R_t - V^a_t$) as the difference between received rewards and expected rewards, where η is the learning rate. The fictitious play model instead updates the state (strategy) of the opponent p^*_t with a prediction error ($P_t - p^*_t$) between the opponent's action and expected strategy. The influence model extends this approach by also including the influence ($Q_t - q^{**}_t$) that a player's own action Q_t *has on the opponent's strategy (see Methods)*.

its simplicity (fewer parameters are needed) and its form, which is more relevant for neuroscience, because it disentangles the two components of the prediction error, as we explain later on.

We chose the inspection game for our experimental study because there is no simple best way of playing it (in game-theoretic terms, it does not have a pure-strategy equilibrium): always inspecting makes the employee always work, in which case inspecting is not optimal, etc. Hence, regardless of the learning model implemented, strategies will keep switching, and, consequently, learning never disappears.

Results

Model Fits to Behavior. To address which of the above strategies most closely captured subjects' behavior, we fit each model to behavior separately and compared the goodness of fit of each model. We found that the influence learning model provided a significantly better fit to subjects' behavior ($P < 0.005$, paired t test) than did either the fictitious play rule or the RL rule, even when taking into account the different number of free parameters in each model by performing an out-of-sample test [Fig. 1*C* and supporting information (SI) Fig. S1]. The influence model also fit better than a variation of the experience-weighted attraction (EWA) learning rule, which involves a combination of RL and fictitious play but has the same number of parameters and hence is equal in model complexity to the influence model (15). Fig. 1*D* shows the relationship between the probability of an action being selected as predicted by the influence model, and actual subject choices. These findings suggest that subjects are not only using representations of the opponents' future choices to guide choice, but are also employing representations of the opponents' likely responses to their own actions.

fMRI Correlates of Influence-Based Expectations. We next analyzed the fMRI data from the player being scanned to determine whether we could find evidence of neural signals reflecting the different components of the influence model, and, if so, whether those signals are better accounted for by the influence model than by the fictitious play or simple RL models. A comparison of brain signals associated with the expected reward of the chosen action, as predicted by each model, is shown in Fig. 2*A*. Expected value signals from the influence model were significantly correlated with neural activity in medial orbitofrontal cortex (mOFC), medial PFC (mPFC) [encompassing both ventral and dorsal aspects, significant at $P < 0.05$ corrected for small volume (SVC)], and right temporal pole ($P < 0.05$ SVC). By contrast, only weak correlations with the expected value signals from the fictitious play model were found in mOFC, whereas no significant correlations were found with expected value as computed by the simple RL model.

Comparison of Influence, Fictitious, and RL Model Fits to fMRI Data. We tested for brain regions showing a significantly better fit to the influence model than the RL model. This analysis revealed significant effects extending from mid to dorsal mPFC ($P < 0.05$ SVC; Fig. 2*B*), as well as in the right temporal pole (Fig. S2). The regression fits of the three models are shown in Fig. 2*C* for mPFC, demonstrating the superiority of the influence model in accounting for neural activity in this area. We then binned BOLD activity from mPFC according to the expected reward as predicted by the influence model to illustrate the relationship between evoked fMRI responses and the model predictions (Fig. 2*D*). These data show that the influence model provides a significantly better account of the neural data in mPFC than does a simple RL model. In addition to the voxel-based analysis we performed an ROI analysis by

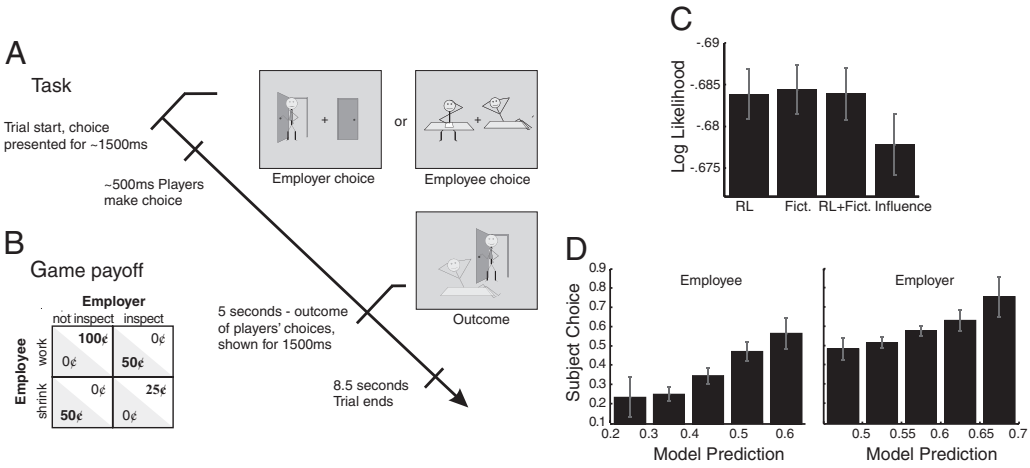

Fig. 1. Inspection game and behavioral results. (*A*) Two interacting players are each individually given two action choices at the beginning of each trial. Players are given 1 second to respond, and their choices are highlighted with a red frame for another second before being covered with a blank screen. Five seconds after the start of the trial, the actions of both players are shown to each player for 1.5 s, with the payoff each one individually receives shown at the top. (*B*) Payoff matrix for the inspection game used in this paper. (*C*) Log likelihood errors for each computational model tested shows that the influence model, which incorporates the effects of players' actions influencing their opponents, has a better fit to subjects' behavior than either the RL or fictitious play models or these two models combined. To account for overfitting and the effects of differences in free parameters between models we used an out-of-sample prediction validation technique, as shown in Fig. S1. Error bars show the SEM of individual log likelihoods. (*D*) Furthermore, the actual probability of a player taking a specific behavioral action is linear with respect to the probability of choosing that action as computed by the influence model. Here, behavior and predictions are shown separately for the employer and employee. Error bars are SEM over subjects.

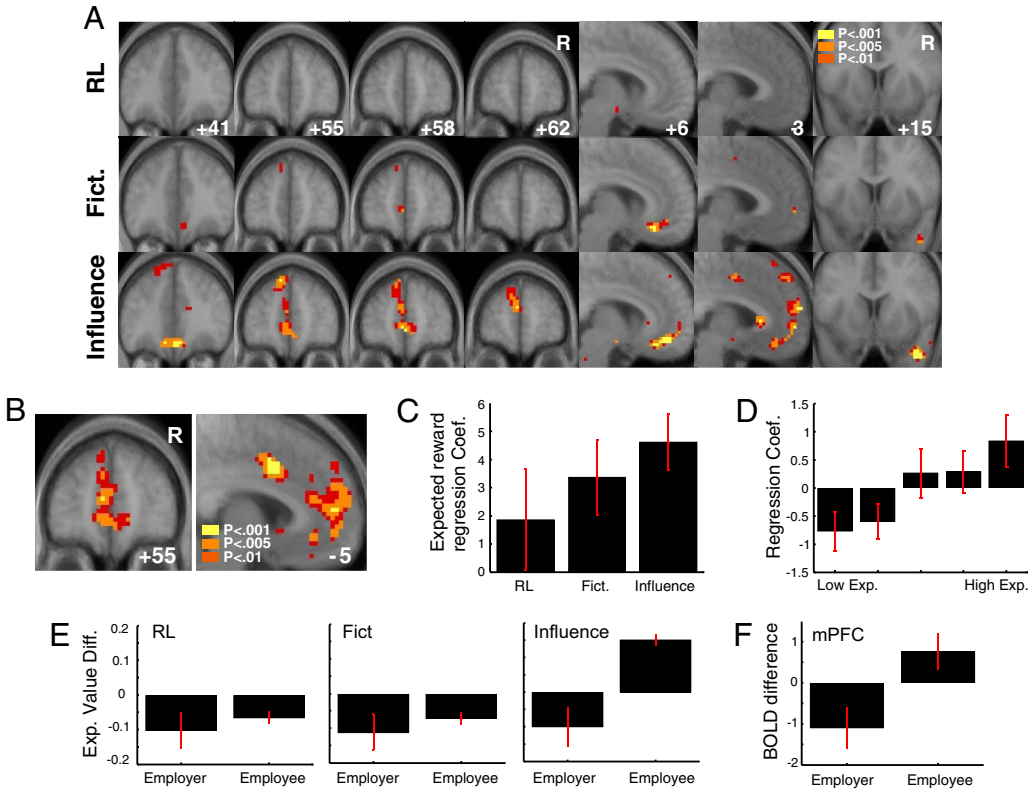

Fig. 2. Expected reward signals. (*A*) At the time of choice, the expected reward of the action selected by a player is shown across the brain as calculated by different computational models. The expected reward signal from the influence model is correlated significantly with BOLD responses in mOFC (0, 36, −21 mm, $z = 3.56$), mPFC (−3, 63, 15 mm, $z = 3.29$), and in the right temporal pole (42, 15, −39, $z = 3.98$), the latter two areas surviving at $P < 0.05$ correction for small volume (SVC) within an 8-mm sphere centered on coordinates from areas implicated in mentalizing (3), whereas only the fictitious play model has significant activity in mOFC (at $P < 0.001$). The RL model had no significant activity correlating with expected reward anywhere in the brain. (*B*) An analysis to test for areas showing neural activity related to expected reward, which is explained significantly better by the influence model than by the RL model, revealed statistically significant effects in mPFC (−3, 57, 12 mm, $z = 3.11$; $P < 0.05$ SVC). (*C*) The average correlation coefficients for each model from the area reported in *B* (extracted from all voxels showing effects at $P < 0.005$ in mPFC). All images shown depict whole-brain voxel-wise comparisons; small volumes are defined only for the purposes of correction for multiple comparisons. (*D*) fMRI activity in mPFC shows a linear relation with binned expected reward probabilities as computed by the influence model (fMRI activity extracted from individual peaks in a 10-mm search radius centered on peak from *B*). (*E*) The computational models tested in this article make distinctly different predictions about the expected reward signals after switching actions (switch) or sticking to the same action (nonswitch) as a consequence of influencing the opponent. Intuitively, the underlying reason is that both RL and fictitious play will most likely ''stay'' after a reward and ''switch'' after a nonreward. However, the influence model has a higher incentive to switch even after receiving a reward. That is, expected reward signals associated with a specific action do not necessarily increase after the receipt of a reward when taking into consideration the influence that specific action exerts on the opponent's strategy. (*F*) fMRI responses in mPFC at the time of choice on switch compared with nonswitch trials show a response profile consistent with the influence model and not the fictitious play models or RL models (the data are extracted from a 10-mm sphere centered on peak from *B*). The difference between the employee and employer was significant at $P = 0.02$.

extracting averaged activity across voxels within an 8-mm sphere in mPFC [centered on mean coordinates from a metaanalysis by Frith and Frith (3)]. We fit the expected reward regressor from the RL and the influence model separately to these data and compared the regression fits across subjects using a paired t test. The expected reward signal from the influence model provided a significantly better fit than the expected reward signal from the RL model at $P < 0.005$ within our mPFC ROI, confirming the conclusions from our voxel-based analysis.

We then aimed to differentiate between the effects of the influence model and the more closely related fictitious play model in this area. For this, we looked specifically at the points in the experiment when the predictions of these two models differ. In particular, the influence model predicts that the expected value after a switch in action choice (i.e., moving from working to shirking or vice versa on successive trials) is on average higher than the expected reward when not switching choice (i.e., taking the same action on successive trials), whereas the fictitious play and indeed RL models predict exactly the opposite (Fig. 2*E*). This effect is greatest for the employee, because behavioral fits indicate that subjects exert more influence on their opponent when playing this role. An analysis of BOLD activity in the mPFC region of interest at the time of choice revealed a positive signal in this area on switch compared with nonswitch trials for the employee, consistent with the predictions of the influence model but not with either the fictitious play or simple RL models (Fig. 2*F*). These results therefore suggest that the influence model does indeed account better for neural activity in mPFC than the fictitious play model.

fMRI Correlates of Updating Signals for Influence Model. At the time of outcome, according to the influence model, a player needs to update his/her expectations of the opponent's strategy using two different components: an influence update signal found only in the influence model and not in either of the other two models, which encodes the magnitude by which the opponent's behavior adapts because of a player's own action; and, in common with both the RL model and the fictitious play model, a prediction error signal that encodes the discrepancy between expected and actual rewards. We found that neural activity in another key component of the mentalizing network, STS (bilaterally), was significantly correlated with

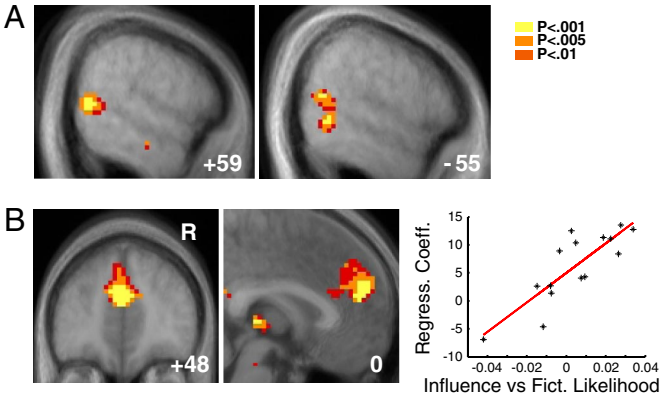

Fig. 3. Influence signals in the brain. (*A*) At the time of outcome, the influence update of the inferred opponent's strategy shows significant correlations with activity in STS bilaterally (−57, −54, 0 mm, $z = 3.32$ and 60, −54, 9 mm, $z = 3.35$; $P < 0.05$ SVC). (*B*) The degree to which a subject thinks he/she is influencing his/her opponent can be measured by taking the difference in log-likelihood fits between the influence and fictitious models on each player's behavior. Likewise, brain regions invoked in computing the influence on the opponent will correlate more strongly with the influence model for subjects invoking this approach when compared with subjects that do not. Influence signals were found to significantly covary with the model likelihood difference (influence − fictitious) across subjects in mPFC (−3, 51, 24 mm, $z = 4.09$; $P < 0.05$ SVC). (*Right*) The relationship between influence regression coefficients and model likelihood differences in mPFC. All images shown depict whole-brain voxel-wise comparisons; small volumes are defined only for the purposes of correction for multiple comparisons.

the influence update signal ($P < 0.05$ SVC; Fig. 3*A*), suggestive of a role for this region in guiding the update of expected value representations in mPFC. Prediction error signals were found to correlate with neural activity in ventral striatum bilaterally (see Fig. S3), consistent with many previous findings implicating this area in prediction error coding (21–24). Moreover, this analysis revealed significant prediction error effects in mPFC, suggesting that this signal could also contribute to the updating of expectations in this region.

fMRI Correlates of Between-Subject Differences in Influence-Related Strategizing. To further investigate differences between the influence and fictitious play models, we examined between-subject variability in the degree to which the influence model provided a better fit to subjects' behavior than the fictitious play model, by comparing the difference in the likelihoods of the two models and correlating that with neural activity elicited by the influence update signal. This measure can be taken as an assay of the individual differences in the degree of influence-based strategizing within our subject group. We found a significant between-subject correlation in the degree of influence activity and the difference in likelihoods between the influence and fictitious play models in dorsomedial PFC ($P < 0.05$ SVC; Fig. 3*B*). These results suggest that, among subjects who strategize more, the influence-based model correlated better with neural activity in mPFC.

Correlations Between Regions During Task Performance. Given the structure of our computational model, an important implication of our findings is that neural activity in mPFC ought to be predictable from a combination of the signals contained in posterior STS (pSTS) and ventral striatum. To test this hypothesis we computed correlations between activity in mPFC and activity in pSTS and ventral striatum separately for each different time point within a trial. We found a significant increase in correlations between activity in STS and mPFC and between activity in ventral striatum and mPFC after receipt of the outcome (when prediction errors and

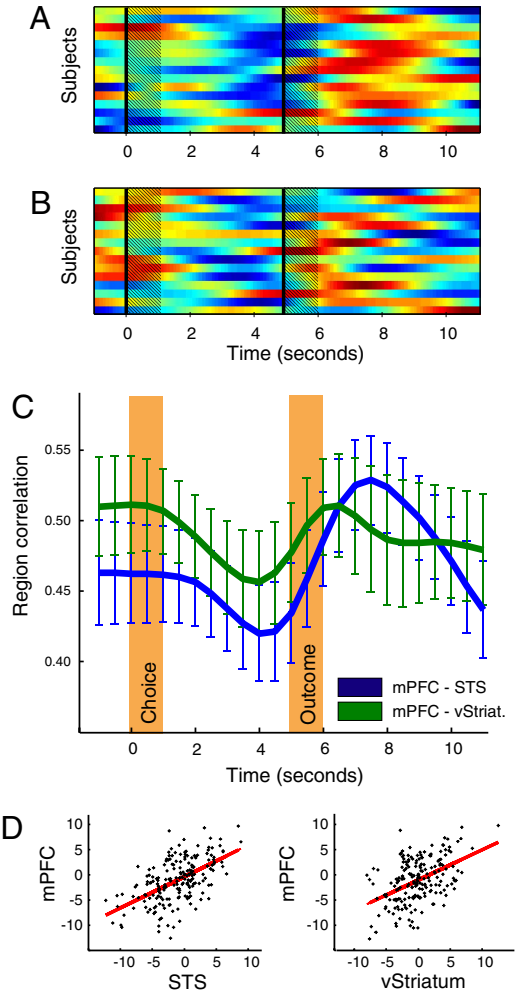

Fig. 4. Interregion correlation analysis. (*A* and *B*) Correlations among mPFC, STS, and ventral striatum were computed for each time point within a trial to determine whether there were significant changes in the correlations between these brain regions at the point the outcome was received (when prediction errors were generated) compared with other time points in the trial. Heat plots of region correlations through time are shown separately for each subject, with correlations between mPFC and STS shown in *A* and correlations between mPFC and ventral striatum (vStriatum) in *B*. Red indicates a high correlation between both regions, and blue indicates a low correlation. Shaded areas indicate the time subjects are given their choices (0 s, time of choice) and the time of outcome (5 s into trial). (*C*) The mean correlation between regions is shown averaged across all subjects for each time point in the trial. A significant increase in correlation after the outcome of a trial was revealed was found at 6.5 s into the trial (significant at $P < 0.01$ compared with the correlation at 4 s; paired *t* test). This supports the idea that information processed at the time of outcome in STS and ventral striatum is being shared with mPFC so as to facilitate updating in the expected reward for a given action. (*D*) Scatter plots of BOLD activity from a typical subject showing the correlation between regions after the trial outcome. Red lines indicates the linear regression fit of STS activity against mPFC activity (*Left*) and of ventral striatum activity against mPFC activity (*Right*).

influence errors are generated) compared with the period before the outcome was delivered (significant at $P < 0.01$ for both regions; Fig. 4). Moreover, activity in mPFC at the time of outcome was significantly better predicted by a linear combination of signals in STS and ventral striatum than by the signals in either of these regions alone (at $P < 10^{-6}$). These findings therefore support the possibility that mPFC, ventral striatum, and pSTS constitute a functionally interacting network underlying computations in strategic game playing, consistent with the tenets of our computational model.

Discussion

In the present study we show that a model that captures an updating strategy in which individuals keep track not only of the actions of the opponent, but also of how opponents are influenced in response to their own actions, provides a good account of behavior during performance of a simple strategic game. We also show that specific computational signals needed for the implementation of such a strategy are correlated with neural activity in different parts of the mentalizing network.

mPFC and pSTS were found to fulfill very different roles in the context of our computational framework. Whereas activity in mPFC was found to track the predicted future reward corresponding to a particular choice given the degree of influence expected, activity in pSTS was found to correspond to an update signal, capturing the difference between the degree of influence expected on a given trial and the actual influence exerted once the outcome had been revealed.

Reward expectations in mPFC were found to be correlated specifically with the predictions of the full influence model, suggesting that these predictions take into account the subjects' expectations of the degree of influence that would be exerted on the opponent given the subject's own inference of how the opponent would respond to his or her own actions. By contrast, reward expectations in this region were not captured well by either a fictitious play model, which simply tracks the actions of the opponent without considering the opponent's reactions to the subject's own actions, or by RL, which tracks only the reward expected given previous choices of the same action. These findings suggest that representations of expected reward in this region take into account inferences about the intentions or beliefs of the opponent toward oneself, often considered a hallmark of the psychological construct of mentalizing.

Another component of the mentalizing network, pSTS, was found to be correlated with the influence that a player's action had on the opponent's strategy. This area has previously been implicated in processing stimuli related to living agents and biologically relevant motion (6, 25). Here we provide evidence that this region is involved in updating an individual's strategy based on computations related to the degree of influence an individual has exerted on their opponent during strategic social interactions. Our computational modeling approach also allowed us to separate out prediction error signals arising from simple RL from those arising from the more complex influence updating mechanism, because the update of expectations in the full influence model is accomplished by a combination of these two signals. When we tested for the presence of prediction error signals arising from the RL component of the model we found significant correlations with those signals in the ventral striatum bilaterally, consistent with many previous reports (21, 26, 27). This signal is independent and dissociable from the influence update signal present instead in pSTS. Taken together these findings suggest that two distinct updating mechanisms are present in the human brain at the same time during strategic interactions: those relating to the difference between the expected and actual rewards (the RL prediction error) and those related to the difference between expected and actual influence exerted.

We also explored functional correlations among the three key brain regions identified as containing signals relevant to our influence model to test an important implication of our model. Namely, activity in mPFC particularly at the time that the outcome is revealed within a trial ought to be correlated with a linear combination of activity in the two regions containing the update signals: ventral striatum and pSTS. Consistent with the predictions of our model, we found a significant increase in correlations between activity in both ventral striatum and pSTS with activity in mPFC at the time the outcome was revealed and the error signals were generated, compared with other time points in the trial. Moreover, activity in mPFC was better predicted by a linear combination of activity in these two regions than by activity in one or the other region alone. These findings support the conclusion that pSTS, ventral striatum, and mPFC constitute a functionally interacting network for implementing the computations relevant to mentalizing.

Although in the present study players understand the effects of influencing the opponent, a key outstanding issue is how they could use that knowledge to alter the opponent's behavior so as to receive bigger future rewards, such as reputation building and teaching (20), or Stackelberg strategies, in which one player commits to a certain strategy and forces the other player to follow suit (28). Furthermore, although in the present study human players always faced real human opponents, an interesting question for further study would be whether similar mechanisms are engaged in these areas when subjects are playing an intelligently adaptive but nonhuman computer, a manipulation often used when probing "theory of mind" areas in human imaging studies (9–11). Another open question is whether other animals besides humans have the capacity for sophisticated strategic computations of this sort, or whether the capacity to engage in such high-level strategies is a uniquely human trait. Although previous studies of strategic game playing in rhesus macaques indicate that these animals do use simple RL and possibly fictitious updating (29–31), it has not yet been addressed whether they are capable of higher-level strategizing as found here in our human subjects.

In this study we have taken the first steps in attempting to characterize the neural underpinnings of mentalizing during strategic interactions in terms of a simple computational model. mPFC and pSTS made distinct functional contributions. Whereas activity in mPFC was found to track the predicted future reward corresponding to a particular choice given the degree of influence expected, activity in pSTS was found to correspond to an influence update signal, capturing the difference between the degree of influence expected on a given trial and the actual influence known to have occurred once the outcome had been revealed. Accordingly, whereas signals in mPFC relating to expectations may be used to guide choice during game performance, signals in pSTS may be used to modulate or change influence expectations on the basis of the actual outcomes experienced. These areas have previously been implicated in mentalizing and in theory of mind but have never been shown to have correlations with distinct computational processes that may potentially underlie such capacities. More generally, the present results show how the application of quantitative computational models to neuroimaging and behavioral data can be used not only to advance knowledge of simple learning situations but also to unlock the complexities of social and strategic interactions (32, 33).

Methods

Subjects. Thirty-two healthy normal subjects participated in this study, of which 16 (25 $\pm$ 1 years old, seven female) were scanned while playing a competitive game in pairs with the other 16. Subject pairs were prescreened to make sure that the subjects in each pair did not know each other before the experiment to reduce the possibility of collusion. However, data from one pair of subjects was in fact discarded because of evidence of collusion during the game. The subjects were also preassessed to exclude those with a prior history of neurological or psychiatric illness. All subjects gave informed consent, and the study was approved by the Institute Review Board at California Institute of Technology.

Task. Functional imaging was conducted by using a Siemens 3.0 Tesla Trio MRI scanner to acquire gradient echo T2*-weighted echoplanar images. Each pair of subjects underwent three game sessions. One subject used a computer terminal and keyboard to play the game while the other was in the scanner using goggles as visual input (Resonance Technologies) and a button box to choose an action. The first session, of 50 trials, was for training; the second two sessions, of 100 trials each, are reported in this article. Player roles alternated between the two subjects in each session. Thus, the scanned subjects reported in this paper played both roles in subsequent sessions (employer and employee). We also included randomly intermixed null event trials, which accounted for 33% of the total number of trials in a session. These trials consist of the presentation of a fixation cross for 7 s. Before entering the scanner, subjects were informed that they would receive

what they earned in a randomly selected session (not training), plus an additional $10. The payoff matrix's association with players' roles and actions (Fig. 1 *A* and *B*) remained fixed throughout an experiment but was rotated across experiments to create a balanced design with respect to roles and role actions. However, for convenience, we refer to the roles of employer and employee as having the payoff matrix defined in Fig. 1*B*.

fMRI Scanning Parameters. To optimize functional sensitivity in OFC we acquired the data using an oblique orientation of 30° to the anterior–posterior commissure line. A total of 580 volumes (19 min) were collected during the experiment in an interleaved ascending manner. The imaging parameters were as follows: echo time, 30 ms; field of view, 192 mm; in-plane resolution and slice thickness, 3 mm; TR, 2 s. High-resolution T1-weighted structural scans ($1 \times 1 \times 1$ mm) were acquired for anatomical localization. Image analysis was performed by using SPM2 (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London). Preprocessing included slice timing correction (centered at TR/2), motion correction, spatial normalization to a standard T2* template with a resampled voxel size of 3 mm, and spatial smoothing using an 8-mm Gaussian kernel. Intensity normalization and high-pass temporal filtering (128 s) were also applied to the data (34).

Behavioral Data Analysis. Details of the computational models used in the analysis are provided in *SI Methods* and summarized in Table 1. Model parameters were estimated by maximizing simultaneously the logistic log likelihood of the predicted decision probabilities generated by a model against the actual behavior of all subjects. Parameters for the employer and employee roles were fitted separately to account for any differences. We used the multivariate constrained minimization function (fmincon) of the Optimization Toolbox 2.2 in Matlab 6.5 (Mathworks) for this fitting procedure (see *SI Methods* and Table S1 for more details). All behavioral data shown correspond to the actions of all 30 participants.

fMRI Data Analysis. The event-related fMRI data were analyzed by constructing sets of delta (stick) functions at the time of the choice and at the time of the outcome. Additional regressors were constructed by using the model estimated choice expected values as a modulating parameter at the time of choice and one or more (depending on the model) update signals as modulating parameters at the time of outcome. All of these regressors were convolved with a canonical hemodynamic response function. In addition, the six scan-to-scan motion parameters produced during realignment were included to account for residual motion effects. These were fitted to each subject individually, and the regression parameters were then taken to the random-effects level. All reported fMRI statistics and uncorrected *P* values arose from group random-effects analyses ($n = 15$), and small-volume corrected *P* values were obtained from 8-mm spheres around regions of interest previously implicated in mentalizing from a metaanalysis by Frith and Frith (3). Specifically, coordinates were defined from the center of each of the areas delineated from the metaanalysis: mPFC (0, 56, 19 mm), STS ($\pm$53, $-$51, 10 mm), and the temporal poles ($\pm$46, 11, $-$35 mm). An across-subject model likelihood difference modulator was also fitted at the second level, alongside the mean group level effect, when studying the influence signal in Fig. 3.

We also compared the influence and RL algorithms with each other by fitting both models at the same time against the fMRI data (Fig. 2*B*). To make the models as similar as possible, we normalized all modulating regressors before fitting to the fMRI data. Areas showing significant activity for any regressor of a given model indicated regions that were better explained by that model in comparison to the other. This approach for model comparison, as well as the alternative ROI approach, is explained in more detail in *SI Methods*.

Details of the interregion correlation analysis are also provided in *SI Methods*.

ACKNOWLEDGMENTS. J.P.O. was supported by grants from the Gimbel Discovery Fund in Neuroscience, the Gordon and Betty Moore Foundation, and a Searle Scholarship. P.B. was supported by the Swiss Finance Institute.

1. Amodio DM, Frith CD (2006) Meeting of minds: The medial frontal cortex and social cognition. *Nat Rev Neurosci* 7:268–277.
2. Gallagher HL, Frith CD (2003) Functional imaging of ''theory of mind.'' *Trends Cognit Sci* 7:77–83.
3. Frith U, Frith CD (2003) Development and neurophysiology of mentalizing. *Philos Trans R Soc London B* 358:459–473.
4. Greene JD, Sommerville RB, Nystrom LE, Darley JM, Cohen JD (2001) An fMRI investigation of emotional engagement in moral judgment. *Science* 293:2105–2108.
5. Berthoz S, Armony JL, Blair RJ, Dolan RJ (2002) An fMRI study of intentional and unintentional (embarrassing) violations of social norms. *Brain* 125:1696–1708.
6. Puce A, Allison T, Bentin S, Gore JC, McCarthy G (1998) Temporal cortex activation in humans viewing eye and mouth movements. *J Neurosci* 18:2188–2199.
7. Castelli F, Happe F, Frith U, Frith C (2000) Movement and mind: A functional imaging study of perception and interpretation of complex intentional movement patterns. *NeuroImage* 12:314–325.
8. Schultz RT, *et al.* (2003) The role of the fusiform face area in social cognition: Implications for the pathobiology of autism. *Philos Trans R Soc London B* 358:415–427.
9. McCabe K, Houser D, Ryan L, Smith V, Trouard T (2001) A functional imaging study of cooperation in two-person reciprocal exchange. *Proc Natl Acad Sci USA* 98:11832–11835.
10. Gallagher HL, Jack AI, Roepstorff A, Frith CD (2002) Imaging the intentional stance in a competitive game. *NeuroImage* 16:814–821.
11. Rilling JK, Sanfey AG, Aronson JA, Nystrom LE, Cohen JD (2004) The neural correlates of theory of mind within interpersonal interactions. *NeuroImage* 22:1694–1703.
12. Sutton RS, Barto AG (1998) *Reinforcement Learning* (MIT Press, Cambridge, MA).
13. Doya K, Samejima K, Katagiri K, Kawato M (2002) Multiple model-based reinforcement learning. *Neural Comput* 14:1347–1369.
14. Fudenberg D, Levine DK (1998) *The Theory of Learning in Games* (MIT Press, Cambridge, MA).
15. Camerer C, Ho TH (1999) Experience-weighted attraction learning in normal form games. *Econometrica* 67:827–874.
16. Montague PR, King-Casas B, Cohen JD (2006) Imaging valuation models in human choice. *Annu Rev Neurosci* 29:417–448.
17. Costa-Gomes M, Crawford VP, Broseta B (2001) Cognition and behavior in normal-form games: An experimental study. *Econometrica* 69:1193–1235.
18. Wilcox NT (2006) Theories of learning in games and heterogeneity bias. *Econometrica* 74:1271–1292.
19. Stahl DO (2000) Rule learning in symmetric normal-form games: Theory and evidence. *Games Econ Behav* 32:105–138.
20. Camerer CF, Ho TH, Chong JK (2002) Sophisticated experience-weighted attraction learning and strategic teaching in repeated games. *J Econ Theory* 104:137–188.
21. McClure SM, Berns GS, Montague PR (2003) Temporal prediction errors in a passive learning task activate human striatum. *Neuron* 38:339–346.
22. O'Doherty JP (2004) Reward representations and reward-related learning in the human brain: Insights from neuroimaging. *Curr Opin Neurobiol* 14:769–776.
23. O'Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ (2003) Temporal difference models and reward-related learning in the human brain. *Neuron* 38:329–337.
24. Seymour B, *et al.* (2004) Temporal difference models describe higher-order learning in humans. *Nature* 429:664–667.
25. Grezes J, *et al.* (2001) Does perception of biological motion rely on specific brain regions? *NeuroImage* 13:775–785.
26. Lohrenz T, McCabe K, Camerer CF, Montague PR (2007) Neural signature of fictive learning signals in a sequential investment task. *Proc Natl Acad Sci USA* 104:9493–9498.
27. Kim H, Shimojo S, O'Doherty JP (2006) Is avoiding an aversive outcome rewarding? Neural substrates of avoidance learning in the human brain. *Plos Biol* 4:1453–1461.
28. Fudenberg D, Tirole J (1991) *Game Theory* (MIT Press, Cambridge, MA).
29. Barraclough DJ, Conroy ML, Lee D (2004) Prefrontal cortex and decision making in a mixed-strategy game. *Nat Neurosci* 7:404–410.
30. Dorris MC, Glimcher PW (2004) Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron* 44:365–378.
31. Lee D, McGreevy BP, Barraclough DJ (2005) Learning and decision making in monkeys during a rock-paper-scissors game. *Cognit Brain Res* 25:416–430.
32. Glimcher PW, Rustichini A (2004) Neuroeconomics: The consilience of brain and decision. *Science* 306:447–452.
33. Camerer CF, Fehr E (2006) When does ''economic man'' dominate social behavior? *Science* 311:47–52.
34. Friston KJ, *et al.* (1995) Spatial registration and normalization of images. *Hum Brain Mapp* 3:165–189.
35. Rescorla RA, Wagner AR (1972) in *Classical Conditioning II: Current Research and Theory*, eds Black AH, Prokasy WF (Appleton-Century-Crofts, New York), pp 64–99.