

# Bayesian nonparametric models characterize social sensitivity in a competitive dynamic game

**Kelsey McDonald (krm58@duke.edu)**

Department of Psychology and Neuroscience, Duke University  
Durham, NC 27708

**Scott Huettel (scott.huettel@duke.edu)**

Department of Psychology and Neuroscience, Duke University  
Durham, NC 27708

**John Pearson (john.pearson@duke.edu)**

Department of Biostatistics & Bioinformatics, Duke University  
Durham, NC 27710

## Abstract

**Previous studies of strategic social interaction in game theory have predominantly used games with clearly-defined turns and limited choices. Yet, most real-world social behaviors involve dynamic, coevolving decisions by interacting agents, which poses challenges for creating tractable models of behavior. We have previously shown that it is possible to quantify the instantaneous dynamic coupling in strategic human game play when paired against both human and artificial opponents. Here, we apply this coupling model to human neuroimaging data. We observe that the rTPJ and dmPFC exhibit increased activation when playing against a human opponent compared to a computer opponent, both immediately before and after game play. Moreover, a network of regions frequently associated with social cognition, including the dlPFC and dmPFC, was found to correlate with player coupling metrics derived from our model for both human and computer opponents. These findings suggest that prefrontal cortex may play a role in tracking the relationship between oneself and other dynamic agents, regardless of whether those agents are perceived to be human.**

**Keywords:** Social Cognition; Gaussian Process; Decision-Making; fMRI

## Introduction

Over the last fifteen years, game theory has been foundational in establishing a neuroscience of strategic decision making (Camerer, 2011). Paradigms like Matching Pennies, the Trust/Ultimatum Game, and Prisoner's Dilemma have used simple choices in highly standardized contexts to rigorously characterize the psychological processes underlying trust, altruism, and inequity aversion, drawing on literature detailing mathematically normative behavior (Camerer, 2011; Mookherjee & Sopher, 1994). Yet many of the strengths of these paradigms—discrete choices, turn-taking, known payouts—run counter to our experience in real-world actions like negotiation, in which participants respond to one another

in real-time, their strategies coevolving amid ambiguously defined incentives.

Here, we leverage a recently published computational modeling framework (McDonald, Broderick, Huettel, & Pearson, 2019) that borrows from recent advances in reinforcement learning (Sutton & Barto, 1998; Silver et al., 2016; Jaderberg et al., 2018) and nonparametric Bayesian modeling (Rasmussen & Williams, 2006; Hensman, Fusi, & Lawrence, 2013) to capture these social dynamics in a more externally valid, large state-action space. Our approach models behavior in a dynamic, competitive motor decision-making task played against both human and computer opponents and is able to capture strategic differences across participants, trials, and even individual moments within trials. This paradigm generates a rich complexity in individuals' behavior that can be succinctly described by individualized, instantaneous policy functions, facilitating analysis at multiple timescales of interest and types of neural data. We conclude by applying our computational model of behavior to neuroimaging data to reveal distinct brain regions that are recruited for strategic decision-making modulated by social identity of one's opponent in the task.

## Experimental Paradigm

We adapted a zero-sum dynamic control task (Iqbal et al., 2019), inspired by a penalty shot in hockey, see Figure 1. The task involved two players: an experimental participant ( $n = 82$ ) who controlled an on-screen circle (the "puck") and another agent who controlled an on-screen bar (the "goalie"). Both players were able to move their avatars using a joystick. The participant controlling the puck attempted to cross a goal line located at the right end of the screen, while the goalie attempted to block the puck. On half of the trials, the experimental participant played against a human; on the other half of trials, the participant played against a computer-controlled goalie. The identity of the goalie opponent (i.e. human or computer goalie) was randomly selected each trial and was disclosed to the participant before each trial began. Our task was incentive-compatible: both the experimental participant and the human goalie were rewarded in monetary bonuses



based on how frequently each player won.

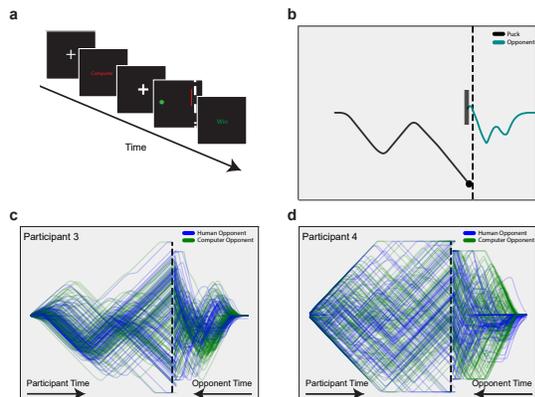


Figure 1: A: Task progression: Following a jittered fixation cue, text indicated the identity of the opponent on the upcoming trial for 2 seconds. Play commenced after a variable delay during which the screen displayed a fixation cue. At the conclusion of each trial, which lasted roughly 1.5 seconds, colored text indicated the winner (green “Win” if the participant won; red “Loss” if the participant lost) for 1.5 seconds. B: Game play on a single trial. The puck moved from left to right at constant horizontal velocity. The bar was only allowed to move vertically, but is depicted as moving from the right side of the screen inward toward the goal line for visualization purposes. C and D: All trajectories for Participant 3 (C) and Participant 4 (D), demonstrating the heterogeneity observed across participants. Subjects exhibited significant variability in both on screen positions’ visited and trajectory shape: Participant 3 is much more consistent in game play, while Participant 4 was more variable. Trials played against the human opponent are displayed in blue. Trials played against the computer opponent are in green.

### Predicting Change Points

As we have previously shown (McDonald et al., 2019), subjects exhibited considerable variability in game play (Figure 1C,D). Despite the fact that participants could produce smooth trajectories by controlling the vertical velocity of the puck, we observed that most trials could be approximated as a sequence of maximal velocity segments separated by change-points, which we defined as either an initial change of the vertical velocity away from 0 or a subsequent change in the sign of vertical velocity. We thus chose to define each trial as defined by the set of such change points. In this approximation, a player’s strategy could be fully characterized by the probability of a change point at each moment.

Viewed through the lens of reinforcement learning, the decision of whether to switch direction at time  $t$  is an action,  $a_t$ , and the probability of this action given a state of the world  $s_t$  is given by the policy function:  $\Pi(a_t, s_t, \omega) = p(a_t | s_t, \omega)$ , where we let  $s_t$  denote a vector of predictors at each time point and

$\omega$  is a binary variable indicating the opponent’s identity (computer = 0, human = 1) (Sutton & Barto, 1998). We define the action space as a single binary variable, with 1 indicating a change in direction and a 0 indicating continuation along the current trajectory. However, the state  $s$  remains continuous and includes 7 predictor variables: the  $x$  and  $y$  positions of the puck, the  $y$  position of the bar, their respective vertical velocities, the time since the occurrence of the last change point (normalized to 1 by dividing by total trial length), and an opponent experience variable that ranged from 0 (first trial) to 1 (last trial) that was specific to each opponent and reflected potential strategic adaptation over the course of the experiment. We fit each subject’s behavioral data with a Gaussian Process (GP) classification model, which offers competitive modeling performance coupled with uncertainty estimation and differentiability, both of which we leverage in our sensitivity analyses.

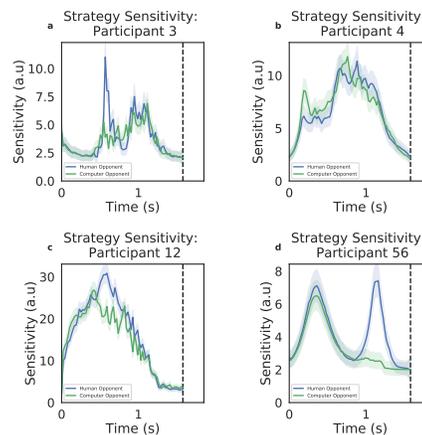


Figure 2: Observed sensitivity to opponent actions in both opponent conditions, for the selected four participants (A,B,C,D). Shaded regions indicate 95% credible intervals. Blue line and shaded region correspond to the human opponent condition, green to the computer opponent.

### Sensitivity Metric

We next sought to quantify how much participants’ switching behavior changed as a function of the opponent’s actions. Because our change point policy model is based on a smooth Gaussian Process, we can quantify this sensitivity using gradients of the GP  $f = \Phi^{-1}(\pi)$  with respect to the opponent’s position and velocity. We then used these gradients to define a moment-by-moment sensitivity index. Since the gradients of the GP measure the degree to which small changes in the current game state affect the participant’s probability of changing course, gradients with respect to the opponent’s position and velocity capture the degree to which the participant’s current behavior is sensitive to the opponent’s actions. For each input variable, we defined a sensitivity as the squared norm of the gradient of the GP along that direction:  $v_i = \|\eta_i^{-1} \nabla_i f\|^2$ , with  $i = 1 \dots 8$  indexing  $(s, \omega)$ ,  $\nabla_i$  the gradient with respect to

the  $i_{th}$  variable, and  $\eta_i^{-1}$  representing the  $i$ th diagonal of the posterior covariance of  $\nabla f$ . Further, to capture overall sensitivity of the puck to the goalie's actions, we combined the sensitivities to goalie position and velocity into a single metric:

$$\zeta = \|L^{-1}\nabla_{\tilde{x}}f(x)\|^2 \quad (1)$$

with  $\tilde{x} \equiv (y_{\text{goalie}}, v_{\text{goalie}})$  and  $L$  the Cholesky factor ( $\Sigma = L^T L$ ) of the covariance ( $\Sigma$ ) of  $\nabla_{\tilde{x}}f$ .

We observed large within-subject heterogeneity with regard to what extent sensitivity to opponent action varies throughout the trial and changes as a function of opponent identity, see Figure 2. With this instantaneous regressor operationalizing the dynamic coupling between opponents, we next sought to apply this behavioral model to neuroimaging data and determine the neural structures that play crucial roles in social cognition and decision-making in our paradigm.

### Neural Structures of Strategic Social Cognition

We wanted to investigate whether strategically playing against the human or computer opponent yielded differences in BOLD activity. To this end, we analyzed fMRI data for (n=72) subjects from the original (n=82) subjects from the behavioral sample that met motion quality thresholds. Our design matrix included the onset and duration of three distinct phases of each trial: 1) the opponent screen, in which participants were notified whom they would be playing against, 2) the game play period, and 3) the outcome screen, in which the result of the game play (win or loss) was displayed. Our design matrix also included the opponent identity of each trial (either the human or computer opponent), and each trial's mean logged opponent sensitivity. According to a GLM analysis, we observed increased activation of the right temporoparietal junction (rTPJ) when the subjects were told they would play the upcoming trial against a human, as opposed to the computer opponent (Figure 3A). In contrast, we observed increased activity for human trials relative to computer trials selectively in the dorsomedial prefrontal cortex (dmPFC) during the outcome screen (Figure 3B).

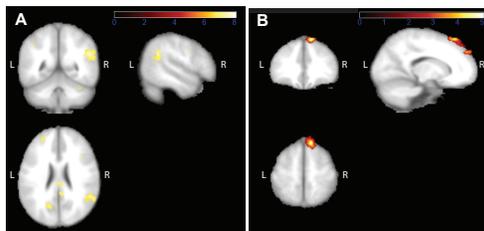


Figure 3: A: GLM results during the Opponent Screen period for the Human - Computer Opponent contrast, which features rTPJ activation (MNI coordinates:  $[x=52, y=-48, z=24]$ ). B: GLM results during the Outcome Screen period for the Human - Computer Opponent contrast, which features dmPFC activation (MNI coordinates:  $[x=14, y=60, z=36]$ ). Color scale corresponds to z-statistic.

### Neuroimaging of Opponent Sensitivity

We next asked whether any regions in the brain were parametrically modulated by opponent sensitivity. Whole-brain analyses revealed that BOLD activity in the dorsolateral prefrontal cortex (dlPFC) was correlated with opponent sensitivity during game play (Figure 4A). This corroborates findings in the social cognition literature and in monkey physiology that neurons in the dlPFC encode update signals of outcome estimates in a competitive game (Barracough, Conroy, & Lee, 2004; McNamee, Liljeholm, Zika, & O'Doherty, 2015; Tsutsui, Grabenhorst, Kobayashi, & Schultz, 2016). Conversely, activity in the dmPFC before the trial began predicted opponent sensitivity during game play (Figure 4B). This demonstrates that pre-trial activity in the dmPFC predicts how coupled one subjects are with their opponents.

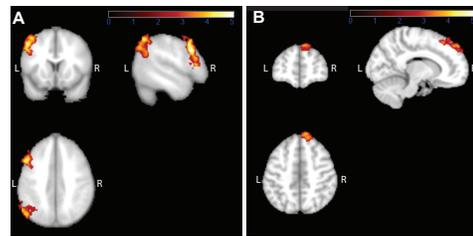


Figure 4: A: Brain regions that are parametrically modulated by opponent sensitivity during the Game Play period, which features the dlPFC and left angular gyrus (MNI coordinates:  $[x=-50, y=14, z=40]$ ). B: dmPFC is predictive of the upcoming trial's opponent sensitivity during the opponent screen (pre-trial) (MNI coordinates:  $[x=12, y=50, z=46]$ ). Color scale corresponds to z-statistic.

Finally, we tested the hypothesis that the rTPJ is representing uniquely social signals when preparing to play against a given opponent. We extracted the activation z-score from over 70 standard regions of interests (ROIs) created from the Harvard-Oxford Cortical and Subcortical atlases during the opponent pre-trial screen. When separating these activation z-scores into human and computer trials, we see that the rTPJ is the brain region that has the highest residual from an orthogonal distance regression line (see Figure 5). This result suggests that rTPJ activity represents information relating to opponent identity in a manner that is preferential to social agents.

### Discussion

Previous studies in social cognitive neuroscience have shown that the rTPJ carries uniquely social signals about the relevance of agents in the current environment (Carter, Bowling, Reeck, & Huettel, 2012; Saxe & Wexler, 2005). We posit that the rTPJ might be instrumental in signaling the presence of a task-relevant social agent. Consistent with this, we find that the rTPJ not only is significantly more active when preparing to play against a social opponent rather than a nonsocial oppo-

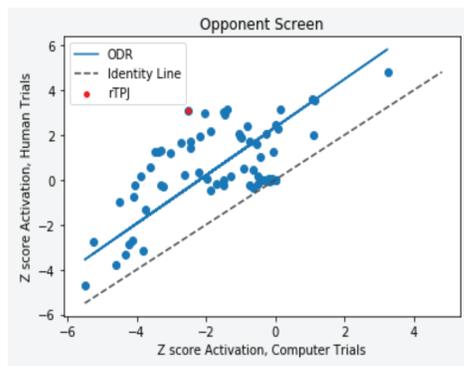


Figure 5: rTPJ has the highest ODR residual when regressing each ROI's human z-score activation against computer z-score activation during both the pre-trial opponent screen.

ment, but also is the single brain region that displays the highest social bias, as defined by the highest residual difference in an orthogonal distance regression plotting beta coefficients from the rTPJ for both social and nonsocial opponents during the pre-trial period. We also find that the dmPFC plays a role in both representing the outcome of trials in our paradigm, but also in predicting future opponent sensitivity. This also accords with existing literature suggesting that the dmPFC tracks a simulated-other's action prediction errors (Lee & Seo, 2016; Suzuki et al., 2012) as well as strategizing during competitive interactions (Rilling & Sanfey, 2011).

Together, these fMRI results suggest that brain regions in the social cognition network including the rTPJ, dmPFC and dlPFC play distinct and dissociable roles during evolving decision contexts, with the rTPJ preferentially signaling when a subject is informed they will interact with a social as opposed to nonsocial agent, while the dmPFC is engaged with both predicting opponent sensitivity before game play and preferentially signaling when a win or loss outcome occurs after playing with a social opponent, rather than a nonsocial opponent. This suggests a potential model in which investigators can leverage the power of nonparametric methods for both modeling the computational and neural mechanisms of dynamic decision-making against multiple opponent types, both social and nonsocial in nature.

### Acknowledgments

This work is supported by a BD2K Career Development Award (K01-ES-025442) to JP, a NIMH R01-108627 to SH as well as the National Science Foundation Graduate Research Fellowship under NSF GRFP DGE-1644868 to KM.

### References

Barracough, D. J., Conroy, M. L., & Lee, D. (2004). Prefrontal cortex and decision making in a mixed-strategy game. *Nature neuroscience*, 7(4), 404.

Camerer, C. F. (2011). *Behavioral game theory: Experiments in strategic interaction*. Princeton University Press.

Carter, R. M., Bowling, D. L., Reeck, C., & Huettel, S. A. (2012). A distinct role of the temporal-parietal junction in predicting socially guided decisions. *Science*, 337(6090), 109–111.

Hensman, J., Fusi, N., & Lawrence, N. D. (2013). Gaussian processes for big data. *arXiv preprint arXiv:1309.6835*.

Iqbal, S. N., Yin, L., Drucker, C. B., Kuang, Q., Gariépy, J.-F., Platt, M. L., & Pearson, J. M. (2019). Latent goal models for dynamic strategic interaction. *PLoS computational biology*, 15(3), e1006895.

Jaderberg, M., Czarnecki, W. M., Dunning, I., Marris, L., Lever, G., Castaneda, A. G., ... others (2018). Human-level performance in first-person multiplayer games with population-based deep reinforcement learning. *arXiv preprint arXiv:1807.01281*.

Lee, D., & Seo, H. (2016). Neural basis of strategic decision making. *Trends in neurosciences*, 39(1), 40–48.

McDonald, K. R., Broderick, W. F., Huettel, S. A., & Pearson, J. M. (2019). Bayesian nonparametric models characterize instantaneous strategies in a competitive dynamic game. *Nature communications*, 10(1), 1808.

McNamee, D., Liljeholm, M., Zika, O., & O'Doherty, J. P. (2015). Characterizing the associative content of brain structures involved in habitual and goal-directed actions in humans: a multivariate fmri study. *Journal of Neuroscience*, 35(9), 3764–3771.

Mookherjee, D., & Sopher, B. (1994). Learning behavior in an experimental matching pennies game. *Games and Economic Behavior*, 7(1), 62–91.

Rasmussen, C. E., & Williams, C. K. (2006). *Gaussian process for machine learning*. MIT press.

Rilling, J. K., & Sanfey, A. G. (2011). The neuroscience of social decision-making. *Annual review of psychology*, 62, 23–48.

Saxe, R., & Wexler, A. (2005). Making sense of another mind: the role of the right temporo-parietal junction. *Neuropsychologia*, 43(10), 1391–1399.

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., ... others (2016). Mastering the game of go with deep neural networks and tree search. *nature*, 529(7587), 484.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. MIT press Cambridge.

Suzuki, S., Harasawa, N., Ueno, K., Gardner, J. L., Ichinohe, N., Haruno, M., ... Nakahara, H. (2012). Learning to simulate others' decisions. *Neuron*, 74(6), 1125–1137.

Tsutsui, K.-I., Grabenhorst, F., Kobayashi, S., & Schultz, W. (2016). A dynamic code for economic object valuation in prefrontal cortex neurons. *Nature communications*, 7, 12554.