# Gaussian Process Models Characterize Other-Regarding Strategies Over Multiple Timescales in a Dynamic Social Game

**Kelsey McDonald (krm58@duke.edu)**
Department of Psychology and Neuroscience, Duke University
Durham, NC 27708

**William F. Broderick (billbrod@gmail.com)**
Center for Neural Science, New York University
New York, NY 10003

**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

**A primary aim of computational neuroscience is to produce models of human behavior that meaningfully address population-level variability. Previous approaches to strategic interaction have used games with clearly-defined turns and limited choices, since these problems are amenable to tractable mathematical analysis. However, most real-world decisions are dynamic, involving simultaneous, coevolving decisions by each agent. Here, using a competitive game in which participants control the dynamics of an on-screen avatar against either another human or a computer opponent, we show that it is possible to quantify this dynamic coupling between agents. Despite the complexity of this behavior, modern nonparametric modeling methods can address the challenges posed by high-dimensional decision problems. We used Gaussian Processes to model the joint distributions of players' actions and identities (human or computer) as a function of game state. We show that this approach offers a natural set of metrics for quantifying instantaneous strategies, and that these metrics are linked to the models hyperparameters. Moreover, because opponent identity is part of the joint distribution, we can differentiate between effects due to opponent identity and effects due to game context. Our approach facilitates analysis at multiple timescales and suggests new classes of tractable paradigms for assessing human behavior.**

**Keywords:** Social Cognition; Gaussian Process; Decision

## 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 a vast literature detailing mathematically normative behavior (Camerer, 2011; Delgado, Frank, & Phelps, 2005; 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.

Our interest lies in quantifying individual differences in this dynamic interplay. To this end, we created a competitive task in which experimental subjects played against both a human opponent and a computer opponent in a real-time movement-based game. As we show, though the task generated a rich complexity in individuals' behavior, this complexity can be parsimoniously captured and analyzed through the use of nonparametric modeling methods. An additional benefit of these methods is that they lead to a *natural* quantification of the moment-by-moment coupling between players, as well as a characterization of individual differences in terms of model hyperparameters. We are able to examine these metrics over a hierarchy of time scales: moment-by-moment, across trials, and across subjects. Furthermore, we can decompose the variance of these metrics across levels, offering a characterization of each as more or less trait-like or state-like. We find that metrics related to a player's sensitivity to opponent exhibit more across-subject variance compared to other variables characterizing game state, while metrics pertaining to the shooter's own play result in more within-trial variance.

## Experimental Paradigm

We adapted a zero-sum dynamic control task (Iqbal & Pearson, 2017; Iqbal et al., n.d.), inspired by a penalty shot in hockey, for compatibility in an fMRI scanner, 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, see Figure 1B. 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.

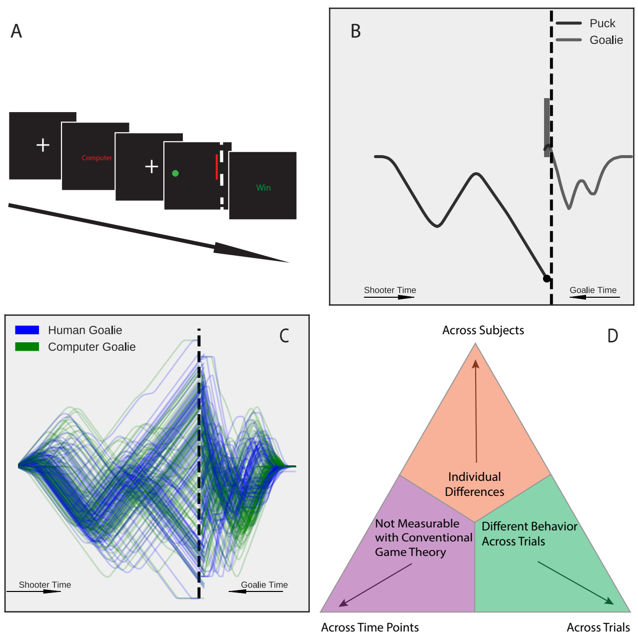


Figure 1: A: Task timeline of each trial. B: Game play on a single trial. The puck moves forward in time from left to right at constant horizontal velocity. The goalie was only allowed to move vertically but is depicted as moving inward toward the goal line over time. C: All trajectories for a representative subject. Trials played against the human goalie are blue. Trials against the computer are green. D: Decomposition of variance for player metrics. Total variance in any quantity defined on a moment-by-moment basis can be decomposed into within-trial, across-trial, and across-subject variance, which must add to 1. All potential metrics thus live within the triangle shown. Metrics that capture only fleeting behaviors lie inside the purple region; those reflecting trial-to-trial adjustments in strategy in green; and those stable within subjects in orange.

## Predicting Change Points

As expected, subjects exhibited considerable variability in game play (Figure 1C). Despite the fact that participants could produce smooth motions 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 approximate 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. Data for our player modeling thus consisted of a binary label for each moment in every trial (1 when a change point occurred at the next time point, 0 for all other points), as well as a set of input features characterizing the current state of the game: the position and velocity of each player, the identity of the opponent (1=human, 0=computer), the number of time points that have passed since the last change point, and an opponent-specific experience variable reflecting where in the progression of the task each time point occurred. Our goal was to predict the binary labels from these input variables.

Our model selection was guided by three requirements: First, the model should be flexible enough to capture the rich diversity of player behavior. Second, the model should appropriately handle the sparse number of change-points ($\approx 4.6\%$) with an input space of moderate dimension. And third, the model should avoid overfitting while providing a principled estimate of uncertainty. For these reasons, we fit each subject's data using a Gaussian Process (GP) classification approach. Briefly, a Gaussian Process (GP) is a distribution over functions. In the same way that a sample from a normal distribution is a real number and a sample from a Bernoulli distribution is a binary variable, a sample from a GP is an entire function (e.g., a time course or spatial density). Gaussian processes have the advantage of providing a principled, Bayesian measure of uncertainty over functions while remaining resistant to overfitting and generalizing to unseen data (Rasmussen & Williams, 2006). GPs are widely used in spatial and time series modeling for their combination of flexibility and ability to generalize from even modest data (Gelfand, Diggle, Guttorp, & Fuentes, 2010; Rasmussen & Williams, 2006). As with its finite-dimensional analog, the multivariate Gaussian distribution, a GP is fully defined by its mean (which we take to be 0 in the prior) and the parameters of its covariance function. GP models offer the benefits of a Bayesian approach: resistance to overfitting, modeling of the data distribution, and a principled estimate of uncertainty. Thus, GPs offer competitive modeling performance coupled with uncertainty estimation and differentiability, both of which we leverage in our sensitivity analyses.

In our case, following standard techniques (Hensman, Matthews, & Ghahramani, 2015; Rasmussen & Williams, 2006), the GP parameterized a quantity analogous to the log odds of a change-point at each game configuration:

$$z \sim \text{Bernoulli}(\pi(s,\omega)) \tag{1}$$

$$\Phi^{-1}(\pi(s,\omega)) \equiv f(s,\omega) \sim \mathcal{GP}(0,k) \tag{2}$$

where $s$ is the vector of variables defining the game state, $\omega \in \{0,1\}$ is a binary indicating opponent identity, $\Phi^{-1}$ is the inverse cumulative normal distribution (also called the probit function), and $\mathcal{GP}(0,k)$ is a GP prior on $f$ with mean 0 and kernel function $k$. Because we assume that $f$ is a smooth function of its inputs, we choose the common radial basis function (RBF) kernel (Rasmussen & Williams, 2006):

$$k(x,x') = \sigma^2 \exp\left(\sum_{i=1} \frac{(x_i - x_i')^2}{\lambda_i^2}\right) \quad (3)$$

with $\sigma$ and $\lambda$ tunable parameters setting the overall magnitude of the covariance and the length scale of correlations along each input dimension, respectively. Here, $x$ includes both $s$ and $\omega$. Even though $\omega$ is a discrete parameter, we approximate it as a continuous variable, as is often done in Bayesian optimization using GPs (Snoek, Larochelle, & Adams, 2012).

We found that our GP classification model accurately captured the diverse patterns present in subjects' data (Figure 2). That is, the model increased its predicted probability of a change point in regions of space where change points occurred in the data. This is a direct result both of the nonparametric nature of the GP—the model adapts its complexity to the data—as well as the smoothing effects of the prior. The underlying GP, $f(s,\omega)$ captures in a smooth function the policy of the shooter in response to both game state and goalie identity. Figure 3A illustrates this contrast by plotting the average (across trials) likelihood of a change point as a function of time in trial for a single subject. For this individual, change-points are likeliest early, with a rise during mid-trial, and a very low value near trial end, consistent with a strategy of committing fully to a direction (up or down) late in the trial. Nonetheless, the subject's average change point probabilities evince minimal difference when playing against a human or computer goalie.

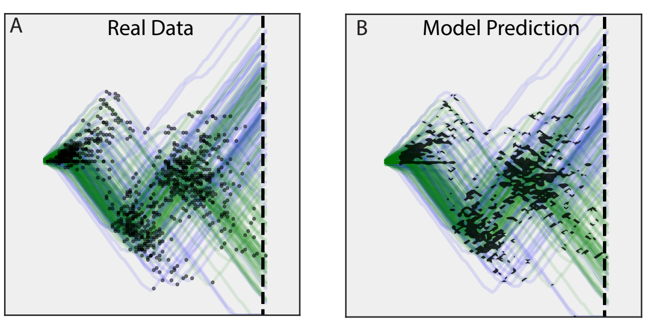


Figure 2: A: Observed data from one subject in the penalty shot task. Trials in which this subject played against the human goalie are colored in blue; trials against the computer goalie are colored in green. The overlaid black points represent change-points (i.e., moments when the subject changed direction). B: Same trajectories from A, this time overlaid with contours illustrating regions in which the odds of a change point increased above the subject's average. The contours are small, indicating a complex policy with numerous local maxima.

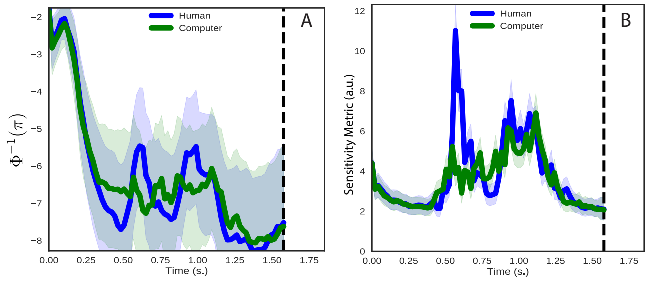


Figure 3: A: Probability of a change point as a function of time for a single subject, averaged across trials. The average for trials played against the human goalie is in blue, against the computer goalie in green. B: The same subject's sensitivity to opponent actions, averaged across time. In both plots, shaded regions represent posterior credible intervals. Whereas the policy does not differ based on opponent, local gradients of the policy, indicative of sensitivity to game state, reveal differences between goalies.

## Sensitivity Metric

Though we observed negligible differences in average change point probability between opponents, we hypothesized that these averages might be obscuring important features of the policy, and that subjects might exhibit differential *sensitivity* to the goalie. To investigate this, we used gradients of the underlying Gaussian Process to define a sensitivity metric for each input variable at each moment in the trial. Because the gradients of the GP measure the degree to which small changes in the current game state affect the shooter's probability of changing course, gradients with respect to the goalie's position and velocity capture the degree to which the shooter's current behavior is sensitive to the goalie's. For each input variable, we defined a sensitivity as the squared norm of the gradient of the GP along that direction: $\nu_i = \|\eta_i^{-1}\nabla_i f\|^2$, with $i = 1\ldots 8$ 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:

$$\varsigma = \|L^{-1}\nabla_{\tilde{x}} f(x)\|^2 \quad (4)$$

with $\tilde{x} \equiv (y_{\text{goalie}}, v_{\text{goalie}})$ and $L$ the Cholesky factor ($\Sigma = L^\top L$) of the covariance ($\Sigma$) of $\nabla_{\tilde{x}} f$. We hypothesized that even when averaged policies were similar, these metrics might differ when subjects played against each opponent.

And indeed, sensitivity to opponent actions varied widely throughout the trial (Figure 3B). Moreover, the difference in sensitivity to the goalie's actions between human and computer opponents was significantly **higher** in those participants who reported the human opponent as requiring more attention than the computer (Mann-Whitney U = 510.0, p = 0.011). Taken together, these results suggest that coarse averages of change point probabilities fail to capture the key differences in subjects' play against the two types of opponents. Rather, it is the local structure of the strategies, as measured by gradients, that differs at temporally localized periods within each trial.

### Decomposing complex behavior across timescales

We have demonstrated that our GP modeling approach naturally provides a measure of sensitivity to each predictor variable at every time point. Yet complex social decisions are organized over a hierarchy of time scales. With this in mind, and spurred by the results of the previous section, we next asked how sensitivity varied as our instantaneous metrics were aggregated over increasingly coarse timescales. Just like with ANOVA, we can decompose variance by averaging at different timescales and looking at how the data vary around those means. Figure 4 illustrates this decomposition for both the instantaneous probability of switching (as measured by $f(s,\omega)$) and the sensitivity to each input variable. Here, as in Figure1D, each point represents sensitivity to a single input variable in the model, with the location of the point represents the proportion of variance allocated to the within-trial, across-trial, and across-subject levels, with the sum of the three constrained to be 1. Two results immediately stand out: First, no points appear in the lower-right corner, indicating that none of the metrics exhibited unique variance across trials. This suggests that, unlike repeated games, strategy did not meaningfully vary across trials; rather, strategies consisted of a set of complex dynamics to be implemented *within* trial. Second, those sensitivities exhibiting higher levels of stability across subjects (near the pinnacle of the triangle) were variables capturing subjects' sensitivity to the goalie's actions, not their own. In other words, subjects' couplings to their opponents was relatively more "trait-like". For instance, sensitivity to the goalie's position exhibited the majority of its variance (61.9%) across subjects, suggesting that the degree to which players attend and react to the goalie's movements is relatively static within an individual but meaningfully variable across the population.

These results demonstrate the power of nonparametric methods for both modeling and analyzing complex strategic interactions. Our experimental paradigm extends previous approaches to decision-making by incorporating simultaneous movement and coupling between agents, and our model allows us to quantify strategies on a moment-by-moment basis, facilitating analysis at multiple time scales. This is particularly valuable for studying the neuroscientific basis of such behaviors, since complementary methods like fMRI and EEG operate on radically different time scales; our model makes predictions applicable to both. Approaches like ours thus open new possibilities for the study of behaviors like competition and social interaction that have remained resistant to neuroscientific investigation.


## 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.


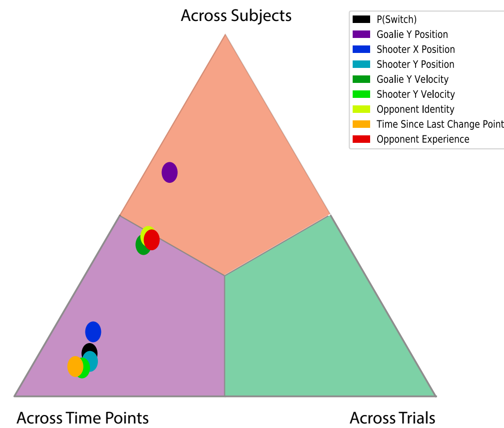


Figure 4: Variance decomposition of player metrics. When within-trial, across-trial, and across-subject variances for a metric are normalized by total variance they must sum to 1, corresponding to a point inside the triangle. Each point thus represents sensitivity to a single input variable in the model, with the location of the point representing the proportion of variance allocated to the within-trial, across-trial, and across-subject levels. The variables that are relatively more trait-like (nearer the pinnacle) represent sensitivity to the opponent (goalie y-position, opponent identity, opponent experience, and goalie y-velocity).

## References


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

Delgado, M. R., Frank, R. H., & Phelps, E. A. (2005). Perceptions of moral character modulate the neural systems of reward during the trust game. *Nature neuroscience*, *8*(11), 1611.

Gelfand, A. E., Diggle, P., Guttorp, P., & Fuentes, M. (2010). *Handbook of spatial statistics*. CRC press.

Hensman, J., Matthews, A. G. d. G., & Ghahramani, Z. (2015). Scalable variational gaussian process classification.

Iqbal, S., & Pearson, J. (2017). A goal-based movement model for continuous multi-agent tasks. *arXiv preprint arXiv:1702.07319*.

Iqbal, S., Yin, L., Drucker, C., Kuang, Q., Gariépy, J.-F., Platt, M. L., & Pearson, J. M. (n.d.). *Latent goal models for dynamic strategic interaction.* (In prep.)

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.

Snoek, J., Larochelle, H., & Adams, R. P. (2012). Practical bayesian optimization of machine learning algorithms. In *Advances in neural information processing systems* (pp. 2951–2959).