

# Modeling Cooperation and Competition in the Tiger Task

**Saurabh Amal Kumar (s.kumar@uke.de)**  
Institute of Systems Neuroscience  
University Medical Center Hamburg, Germany

**Tessa Rusch (t.rusch@uke.de)**  
Institute of Systems Neuroscience  
University Medical Center Hamburg, Germany

**Prashant Doshi (pdoshi@cs.uga.edu)**  
Department of Computer Science  
University of Georgia, GA, USA

**Michael Spezio (mspezio@scrippscollege.edu)**  
Psychology and Neuroscience  
Scripps College, CA, USA  
Institute of Systems Neuroscience  
University Medical Center Hamburg, Germany

**Jan Gläscher (glaescher@uke.de)**  
Institute of Systems Neuroscience  
University Medical Center Hamburg, Germany

## Abstract:

Cooperation and competition are fundamental modes of social interaction. It is imperative that we study such behavior to unravel the staggering complexities of the human brain. We aim to develop a modelling framework to disentangle the neural underpinnings of such behavior with a sophisticated design of the iconic tiger task. The task revolves around the nuances of human decision making where the participant is choosing between two doors hiding a tiger or a gold pot and an option of taking a hint. The task becomes more demanding in the multiplayer setting where one needs to either synchronize actions with the other participant (cooperation) or outsmart the other participant (competition) in order to earn maximum reward. We estimate logistic discrete choice models with Bayesian Hierarchical modeling to model the participants' choices in the single and multiplayer versions of the task. The inclusion of the social information in the model for the multiplayer version significantly improves the model fit. As an extension to this descriptive model, we will use I-POMDP that explicitly models the other participant as an intentional agent to investigate the theory of mind of cooperation and competition further.

**Keywords:** Cooperation; Competition; Tiger-task; Bayesian Hierarchical modeling; IPOMDP

## Introduction

Social interactions have been the building block of society. Cooperation and competition are two fundamental modes of interactions between individuals. Both require social reasoning about the mental states of others and to recognize that their beliefs and goals

may be different from one's own (Theory of Mind). This mentalizing often takes the form of a mental model of the other person that can be queried for estimating the beliefs and ensuing actions of others. These models may or may not include a model that the other person builds of ones' own beliefs and actions (recursivity of social reasoning). It remains an open question, whether cooperation and competition place the same or disproportionate demands on mentalizing capabilities. Traditionally, competition is thought to require more elaborate and recursive reasoning about the other players' strategy: one needs a good estimate of the other person's beliefs and action to be able to exploit them. Nevertheless, cooperation may or may not require the same level of reasoning: a similarly precise estimate of the other person's beliefs and actions is necessary for successful coordination and cooperation. In this study we aim to characterize the social reasoning processes in the context of cooperation and competitive decision-making.

To unravel the cognitive and neural underpinnings of these modes of social interaction we use the iconic Tiger Task. It played a crucial role in developing the partially observable Markov decision processes (POMDPs) computational framework by providing a test bed for simulating decision-making of a single agent in an uncertain world. The task mimics the setting of a game show in which the agent is presented with two doors, one of which hides a tiger (incurring a large loss) and the other one hides a pot of gold (incurring



small win). The uncertainty lies in the location of the tiger. POMDPs are an improvement on the traditional MDP's for situations where the present state is uncertain and can be estimated with Bayesian updating over a belief distribution of the states (Kaelbling, Littman, & Cassandra, 1998).

The POMDP framework has been subsequently extended for multi-agent settings resulting in interactive partially observable Markov decision process (*I-POMDP*) (Gmytrasiewicz & Doshi, 2005), in which two or more agent interact in an uncertain world. Following Doshi (Doshi, 2005), we modified the Tiger task into cooperative and competitive environments where two human participants either cooperate or compete to maximize their goals. In this extension we demonstrate that the social factor is crucial for making informed decisions that maximize reward. In addition, we also highlight the basic differences in cooperation and competition and model them using the Bayesian Hierarchical modeling.

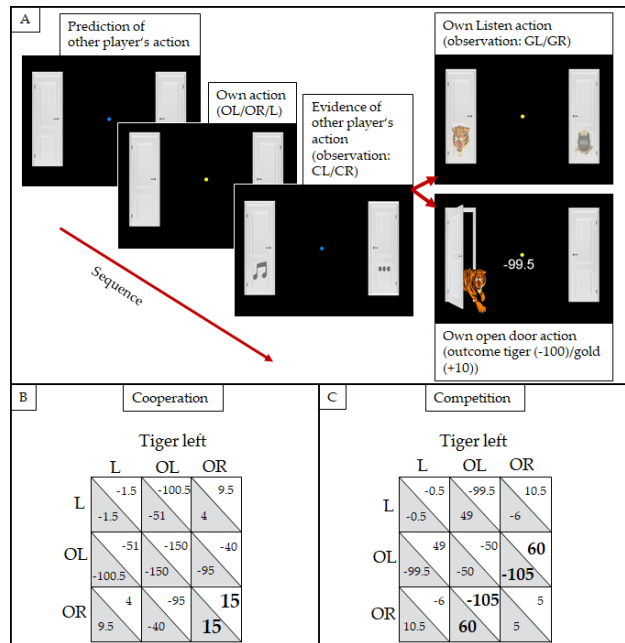
### Task and Hypothesis

The goal of the Tiger Task is to maximize the reward by opening the door hiding the gold (+10 point) and to avoid opening the door with the tiger (-100 points). In each step there are 3 actions available to the participant: open left door (OL), open right door (OR), or listen (L), which results in a probabilistic hint about the location of the tiger (growl left (GL), or growl right (GR)), but also costs 1 point. Thus, participant can accumulate evidence about the tiger location through repeated L actions. After each open action the position of the tiger is reset randomly to one of the two doors (tiger left (TL) or tiger right (TR)).

In the multi-player version, the participants receive an additional probabilistic hint about the actions of the other player: creak left, or creak right (indicating that the other player might have opened one of the doors), or silence (S) indicating that the other player probably listened. Creaks suggest that the location of the tiger might have reset and that currently accumulated beliefs about the tiger location are void. Opening the door reveals the correct location of the tiger and the participant receives the associated reward with additional knowledge of the tiger reset. In our implementation of the Tiger Task participant were also asked to predict the other player's actions at each step before choosing their own action (see Figure 1A for task sequence).

The competitive and cooperative versions differ in the structure of the payoff matrix: while the cooperative version incentivizes concurrent open actions by both players (see Figure 1B bold marking), the competitive version provides the maximum reward, if the correct

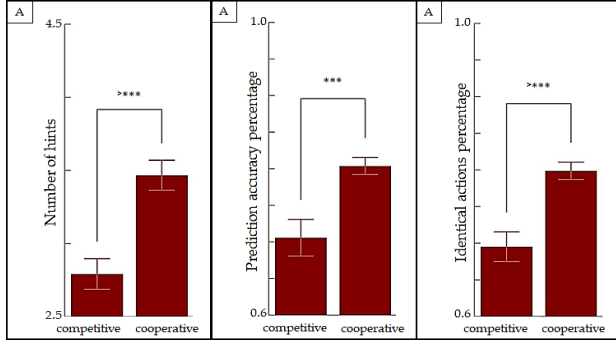
door hiding the gold is opened, while the other player opens the wrong door hiding the tiger (see Figure 1C bold marking). Comparing the two versions, we expected that participants will take more hints to come to reach a consensus in cooperative context to avoid confusing the other player and generate a more predictable behavior. We also expected more identical actions and more accurate predictions of the other player's actions during cooperation.



**Figure 1:** (A) An example of sequence for the multi-player version of the task. The player predicts the action of the other player (indicated by the blue fixation dot) followed by the player's own action choice (indicated by the yellow fixation dot). This is followed by the probabilistic evidence about the other player's action (CL/CR). The next screen is either the probabilistic hint about the tiger location (GL/GR if L was chosen) or the door is opened (for OL or OR actions) revealing the tiger location. (B) The joint payout matrix in the cooperative context is shown when the tiger is on the left side. The bold numbers showing the best choice indicating that the maximum payoff is achieved, when both players open the correct door at the same time. (C) The joint payout matrix in the competitive context is shown for the tiger being on the left side. The bold numbers show the best and worst choice indicating that the best own outcome is achieved if the correct door (with the gold) is opened, while the other player opens the wrong door (with the tiger).

## Results

We invited 58 participants (30 cooperate, 28 compete) to play the multi-player version of the game. In the model-free analysis we observed that the participants in the cooperative context took more hints than in the competitive context. In addition, prediction accuracy was higher during cooperation. These outcomes were both in line with our expectations. Participants in the competitive version exhibited fewer identical actions when compared to cooperation (Figure 2A-C).



**Figure 2:** (A) In the multi-player version of the task the participants significantly took more hints in the cooperative context when compared to the competitive context. Participants also had significantly higher prediction accuracy and identical actions (showing coordination) in the cooperative context compared to the competitive one in (B) and (C) respectively.

Participants in the Tiger Task form beliefs about the states of the game (TL or TR) based on the probabilistic hints (GL or GR) and – in the multi-agent Tiger Task – the information from the other player (CR or CL). Because there are 3 distinct actions (OL, OR, L) available, we decided to model the action  $a(t)$  at each step  $t$  as an ordered logistic regression model:  $a(t) = \beta_0 + \beta_1 * b(t)$ , where  $b(t)$  is the belief about the location of the tiger.

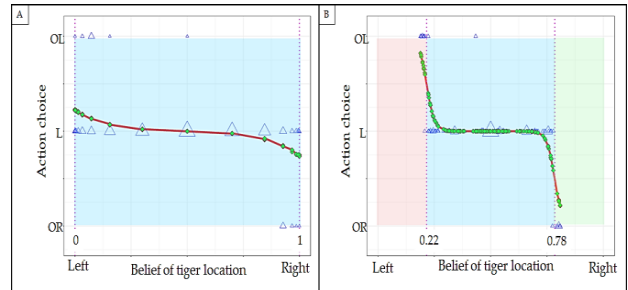
The Tiger Task has only 2 states (TL and TR), which implies a unidimensional belief distribution with both states at the end of the range of possible beliefs. This belief distribution is updated on every step with the observations following the current action. We compared two version of belief updating: a simple “beta-belief” model, which uses the mode of a beta distribution as the point estimate of the belief and is updated by adjusting the parameters of the beta distribution with the observations (the probabilistic hints following L actions). The second model is a Bayesian belief updating model with take the previous belief as the prior and calculates the likelihood based on the observation and transition function. We also tested two versions of the Bayesian updating model without (Eq 1) and with the inclusion of the social information (Eq 2, also see Figure 3A-B as an example):

$$b(t) = \frac{p(gc) * b(t-1)}{p(gc) * b(t-1) + (1-p(gc)) * (1-b(t-1))} \quad (1)$$

Where,  $p(gc)$  is the probability of the hint being correct and  $b(t-1)$  is the previous belief about the tiger location.

$$b(t) = \frac{p(cc) * p(oo)}{p(cc) * p(oo) + (1-p(cc)) * (1-p(oo))} * p(reset) + 1 - \frac{p(cc) * p(oo)}{p(cc) * p(oo) + (1-p(cc)) * (1-p(oo))} * b(t-1) \quad (2)$$

Where,  $p(cc)$  is the probability of the hint about the partners' action being correct and  $p(oo)$  is the probability of the partner opening the door, while  $p(reset)$  is the probability of the tiger being placed after a door is opened (0.5 for a random placement).

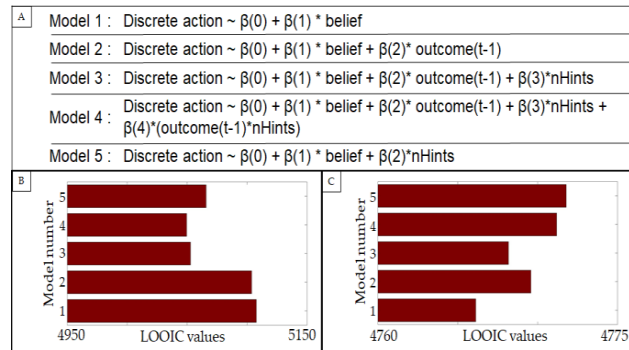


**Figure 3:** (A) An example model behavior of the multi-player version of the tiger task without the social information (see Eq 1) is shown here. The bold red line is the model prediction, while the blue triangles are the actual participant action choices given their computed beliefs. Green dots, which always lie on the red curve show the model predictions of the data (blue triangles). The light-blue area shows the belief region where the ordered logistic model predicts the listen action. In the red and green areas the ordered logistic model predicts Open Left and Open Right action respectively. The absence of these areas in this model suggests that the model without the social information fails to predict the observed open left/right actions. (B) This model behavior shows the prediction made with the social information (Eq 2). This model predicts most of the OL actions (red area) and OR actions (green area) correctly demonstrating the importance of the social information (CR/CL) for correctly predicting the observed data.

Models were estimated using the Stan software package that implements a hierarchical Bayesian workflow. Formal model comparison using LOOIC (Leave-one-out information criterion) revealed that the Bayesian belief update model resulted in a better fit than the beta-belief model (LOOIC (Bayesian belief) = 5107.75, LOOIC (Beta belief) = 8530.70). In control analysis, we expanded the set of predictors in the ordered-logistic model with additional task variables like the number of hints taken, previous outcome and an interaction between them (Model 2-5), but found the simpler model with just the belief as a predictor (Model 1) outperforms these more comprehensive predictor sets (Figure 4A-B). Furthermore, we compared the Bayesian belief update without the social information (Eq 1) to the update with the social information added (Eq 2) and concluded that the social information adds a

significant improvement in the model prediction (see the scales of LOOIC values in Figure 4A and 4B).

Finally, in the Bayesian updating model with the social information we observed that the decision thresholds of



**Figure 4:** (A) Different models compared of the multi-player version of the task. All the models in (B) without the social information perform worse compared with the LOOIC values of the models with the social information added in (C) (see different scales in (B) and (C)). The simplest model with just the belief update (model number 1) in (C) performed better when compared to extensions of number of hints taken, previous outcome and an interaction of them.

the ordered logistic regression in the competitive version were narrower than in the cooperative version ( $p < 0.045$ ) indicating the participants in the latter version compensated for the uncertainty of other player's beliefs about the tiger location before opening the door. In contrast, in the competitive version they were willing to take more risks in order to make to a decision faster than their opponent.

## Outlook

We used an ordered logistic discrete choice model with Bayesian belief updating for modeling the behavioral data in the multi-agent Tiger Task and demonstrated that including the social information is providing a much better model fit to the data. This suggests that participants in the multi-agent Tiger Task do incorporate the information from the other player into their valuation process. However, our Bayesian belief model falls short of an important feature that is likely shaping strategic social decisions: it treats the information from the other players as just another piece of information from the environment and not as an intentional agent that processes the information in a similar way.

I-POMDPs are a computational framework that explicitly computes the beliefs of the other player as an intentional agent as part of the model of the first player. Thus, it is an ideal framework for modeling Theory of Mind of another player in a quantitative way (his goals, intentions, and beliefs). Following our Bayesian belief model, we will also model the Tiger Task within the I-

POMDP framework and compare belief computations of the other player in the competitive and cooperative version of the task.

## Acknowledgments

J. G. was supported by the Bernstein Award for Computational Neuroscience (BMBF 01GQ1006) and J.G. and M.S. were supported by a Collaborative Research in Computational Neuroscience (CRCNS) grant (BMBF 01GQ1603; NSF 1608278). T.R. was supported by a PhD scholarship from the German National Merit Foundation.

## References

- Doshi, P. (2005). *Optimal sequential planning in partially observable multiagent settings*. Ph.D. dissertation, University of Illinois at Chicago.
- Gmytrasiewicz, P. J., & Doshi, P. (2005). A framework for sequential planning in multi-agent settings. *Journal of Artificial Intelligence Research*. doi:10.1613/jair.1579
- Kaelbling, L. P., Littman, M. L., & Cassandra, A. R. (1998). Planning and acting in partially observable stochastic domains. *Artificial Intelligence*. doi:10.1016/S0004-3702(98)00023-X