Title
Using Inverse Planning and Theory of Mind for Social Goal Inference

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 33(33)

Authors
Tauber, Sean
Steyvers, Mark

Publication Date
2011

Peer reviewed

Using Inverse Planning and Theory of Mind for Social Goal Inference

Sean Tauber (sean.tauber@uci.edu)
Mark Steyvers (mark.steyvers@uci.edu)
Department of Cognitive Sciences, University of California, Irvine
Irvine, CA 92697 USA

Abstract

Previous research shows that people assign latent goals or intentions to simple animated agents based on the motion behavior of these agents. We propose that human observers can infer that an animated agent has a partial state of belief about its environment and that observers use this information – in combination with the agent's observable behavior – to infer its goals. We conducted an experiment that showed that observers used line-of-sight cues – an agent's orientation relative to various objects in the environment, and the presence or absence of visual obstructions – to determine the content of an agent's state of belief about the location of objects. Our results are consistent with the hypothesis that human observers use line-of-sight cues to assign belief states to agents and that these belief states can be used to interpret agent behavior. We found that observer models that incorporated inferences about agents' beliefs outperformed an all-knowing observer model in describing human responses. Additionally, we found that human responses were most consistent with the behavior of a model that incorporates information about both orientation and line-of-sight obstructions.

Keywords: Theory of Mind; ToM; belief states; states of belief; goal inference; social goal inference; inverse planning; perceived animacy.

Introduction

Imagine that you are standing across the street from a bank right before closing time. Suddenly, a car pulls up and four bank robbers get out of the car and charge into the bank. A minute later, another car pulls up and a man jumps out of the car and runs towards the bank entrance. What is he doing? Maybe he is trying to stop the robbery or help the hostages; or maybe he is rushing to cash a paycheck before the bank closes. As it stands, we are missing a key piece of information that would help us understand the man's intentions – whether or not he *knows* that the robbers are in the bank. We are often able to make inferences about the intentions of others based on the context of the situation and their behavior but, as our example shows, sometimes we also need to know something about a person's *state of belief* about the world in order to interpret their actions with any amount of certainty.

Theory of Mind

Much research on *Theory of Mind* (ToM) has focused on the ability (or inability) of animals and human children to represent others as having states of belief about the environment that are different from their own. The general assumption is that most human adults have this ability (Premack & Woodroof, 1978; Doherty, 2008). ToM can play an interesting role in our ability to engage in social interaction. For instance, we have to keep track of the information that individuals know or do not know and combine this with contextual information in order to understand the intentions of others. Others have argued that ToM is much too complex to understand in terms of simply having or lacking the ability to represent other's beliefs and that evidence about the limitations in adult's ToM abilities may provide insight about the cognitive process(es) involved in ToM (Samson & Apperly, 2010). These limitations in adults are only beginning to be explored and may lead us to a better understanding of the process or processes that underlie the phenomenon that has been referred to broadly as ToM.

The perception of animacy

Studying ToM and social goal inference in realistic social contexts is a difficult undertaking with many uncontrollable variables. It is therefore useful to develop controlled experiments that allow us to simulate social interactions that are tractable. Heider and Simmel (1944) were the first to demonstrate that humans perceived simple two-dimensional shapes that were animated on a screen as having latent motives, goals and intentions. The motion of these shapes was designed by an animator – the shapes were not real agents and did not have real latent intentions. Nevertheless, human observers perceive these shapes as agents with "minds." This phenomenon is sometimes referred to as *the perception of animacy* (for a technical description see Feldman and Tremoulet, 2008)

Modern research that employed the perceived animacy phenomenon showed that not only did human observers perceive that the agents had goals; they also appeared to perceive that the agents made inferences about the goals of other agents (Baker, Goodman, & Tenenbaum, 2008; Baker, Saxe, & Tenenbaum, 2009; Ullman et al., 2010). These studies showed that not only can people perceive these shapes as agents with minds, but they can also perceive them as agents who can reason about the minds of other agents.

Inverse planning

Baker, et al. (2008, 2009) showed that a *Bayesian inverse planning* process provided inferences about the latent goals of animated agents that were more similar to human inferences than a simple cue-based model. In general, the idea of an inverse planning process is that humans have access to a generative process in which an agent's behaviors can be generated rationally based on the state of the environment, their own goals, and their inferences about the goals of other agents. Humans infer the goals of another agent by inverting this generative process to infer an agent's goals from its observed behaviors. It is important to note that in the aforementioned experiments, the inverse planning model assumed (and the human observers were instructed) that the agents had complete knowledge of the environment including the position of the other agent(s). This was potentially important for the inferences that humans made because they could assume that an agent had the same knowledge about the other agent's behavior as they had.

In our experiment, we eliminate this assumption so that even though human observers have complete knowledge of the environment, they have the opportunity to take into consideration that the agents have incomplete knowledge of the environment.

States of belief

As demonstrated in our bank robbery example, agents often have incomplete or false beliefs about the state of the environment and this can affect human judgments about the goals of these agents. Our objective for the current study was to build on the inverse planning and perceived animacy literature to include situations in which humans would assign relative or incomplete states of belief to animated agents and combine this information with the agents' observable behavior when inferring their goals. In order to create the perception that agents had different states of knowledge about the environment we instructed observers to assume that agents did not know the location of other agents or objects in the environment unless they "saw" them. We predicted that the perception that agents did or did not "see" portions of the environment would be mediated by the presence or absence of obstructions (such as walls and doors) and by orientation cues that would allow observers to perceive that the agent was "looking" in a certain direction. Taken together, we refer to these as *line-of-sight* cues.

The usefulness of orientation cues was inspired by previous research that indicated that these cues influence the way observers perceive the intentions of agents in perceived animacy experiments (Gao, Newman, & Scholl, 2009). In our case, we hypothesized that observers interpreted the orientation of an agent as the direction in which it was looking. In order for an agent to "see" another agent or object in the environment it must have oriented towards that agent or object and there must not have been any obstructions (closed doors) blocking the line of sight. We predicted that if these two conditions were met then an observer would represent the agent as *knowing with certainty* the location of the other agent or object. If the two conditions were not met – either the agent did not look towards the other agent or object; or it did look but there was an obstruction blocking the line of sight – then an observer would represent the agent as *not knowing with certainty* the location of the other agent or object.

The Challenge of modeling ToM

It is challenging to design ToM experiments that involve the dynamic interaction of multiple agents and are rich enough for observers to perceive the agents as having goals, preferences and states of belief, yet remain tractable for the application of computational modeling.

Previous inverse planning research used a Markov Decision Process (MDP) to model continuous agent behavior as a function of its goals and the state of the environment (Baker, et al., 2008, 2009). One way to extend this framework to account for agents having states of belief is to use Partially Observable Markov Decision Process (POMDP). Both MDPs and POMDPs are complex models of sequentially dependent agent behavior. Because we were more interested in the role of belief states than in action planning, we chose to simplify the generative action process in such a way that we could avoid modeling sequentially dependent action information. Specifically, we decided to reduce agent action sequences down to a single discrete multi-choice decision. The hope was that we would better be able to isolate the effects of the belief state inference process on observers' judgments from the effects of the inverse action planning process.

Another issue that arises when attempting to isolate the effect of different variables in this type of framework is the confounding of goals and priorities. Once an observer has inferred that an agent has a certain belief state and observes the agent's behavior in light of that belief, the observer can attempt to use this information to infer the agent's goal. When there are multiple objects in the environment however, the agent may have multiple goals – some of which may be more important than others. An observer may not be able to infer a unique set of goals/priorities to explain an agent's behavior. We address this issue with our *Cops and Robbers* paradigm by assuming that most people assign the same constant set of goals and priorities to specific agent types. Instead of asking observers to infer an unknown agent's multiple goals and the priority of these goals, we ask observers to identify the *type* of agent they are observing – cop or robber. We eliminate the confounding of goals and priorities by assuming that cops always want to get the robber (primary goal), and robbers want to stay away from the cop (primary goal) and get the loot (secondary goal).

Experiment

We designed an experiment that used perceived animacy to simulate social interactions in which observers would potentially use line-of-sight cues to track an agent's state of belief about the environment and combine this information with the agent's motion behavior in order to infer the identity of the agent. The idea was that, given the same motion behavior, different line-of -sight cues would affect observers' perception of agents' states of belief, which would in turn affect their inferences about the agents' identities.

Human participants performed a task in which they observed the interactions of two animated agents and had to determine the identity of a particular agent based on its behavior and the state of the environment. We told participants that one agent was a *cop* and the other was a *robber* but they did not know which agent was which on each of the trials. We also told participants that the agents' knowledge of the environment depended on what the agents could or could not see.

We varied the relative motion and orientation of one of the agents with respect to the other agent and the loot; and we varied the positions of the second agent and the loot, as well as whether there were visual obstructions (walls) between the agents and/or between the agents and the loot.

Method

Participants Participants were 28 undergraduate students from The University of California, Irvine that each received partial course credit for their involvement in our experiment.

Stimuli The stimuli consisted of 128 brief animations in which there was an active agent (blue triangle), a static agent (green triangle) and a static object called the loot (a red square). For each trial the static objects were in one of 32 possible configurations (figure 1-a) and the active agent had one of four possible motion sequences (figure 1-b). Participants were instructed that the interior (gray) doors in the environment blocked the sight of the agents when they were closed, but that they always opened when an agent moved towards them**.**

Procedure Participants were provided with a background story for the experiment in which they were told that there were two agents – a cop and a robber. The cop was trying to catch the robber and the robber was trying to get the loot without being caught by the cop. It was not known which agent was the cop and which was the robber. The experimental task was to identify the moving agent as either the cop or the robber for each animation. On each trial participants watched the animation and were presented with a choice about the moving (blue) agent's identity. The options were "Cop", "Robber" and "Don't Know." The order of the trials was randomized for each participant and the order of the options was randomized for each participant on each trial.

Empirical Results

A comparison of several key trials (Fig. 2) demonstrates the relative impact of motion, orientation, and visible obstructions on human judgments and model predictions. We will first outline the results of the human judgments before moving on to the model predictions.

Figure 2-a demonstrates the effect of a wall between the agents in a trial where the active agent moved towards the other agent. Humans overwhelmingly gave *cop* responses when there was no wall between the agents (fig. 2-a-2), whereas the presence of a wall resulted in uncertainty in the human responses (fig. 2-a-1).

Figure 2-b demonstrates the effect of walls and orientation in a trial where the active agent moved away from the other agent. Humans gave mostly *robber* responses when the active agent had a clear line-of-sight to the other agent and then moved away from it (fig. 2-b-3). When there was no line-of-sight because of non-orientation (fig. 2-b-1) or the presence of a wall between the agents (fig. 2-b-2) human responses were more uncertain.

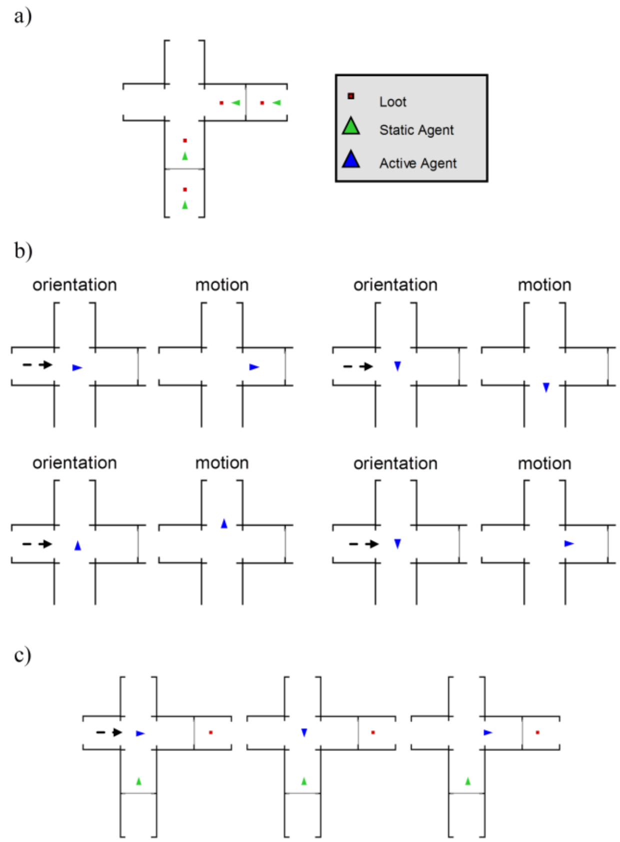

Figure 1. Stimuli: a) shows the four possible positions for the loot (red square) and static agent (green triangle) – which was oriented either towards (as shown) or away from the center room; b) each row demonstrates one of the four possible motion sequences for the active agent (blue triangle) – the three columns depict the active agent's starting position, orientation behavior, and motion behavior; c) A complete example trial as seen by a human observer at 3 different points in time. The blue agent moves from the left-most room into the intersection, "looks" down towards the green agent, and then moves away from the green agent and towards the loot (which is behind a closed door). Gray doors always opened as agent approached them—they obstructed line-of-sight but not motion.

Computational Theory

Graphical models[1] are a useful way to describe the generative process by which human participants respond to

[1]For an introduction to graphical model notation, see Koller, Friedman, Getoor, and Taskar (2007).

experimental stimuli. We develop four graphical models of observer behavior and compare the predictions of these models to the human response data.

We assume that observers use an inverse planning process that reverses an action-planning model to infer the identity of an agent from observations of its actions.

In order to model the agent's goal driven behavior as a single multi-choice decision, we separate each trial into two distinct phases. The *information gathering* phase consists of the agent moving into the center of the maze and its orientation behavior. All of our models assume that this sequence of behavior is not generated by the agent's goal-directed action planning process, but rather by a random information gathering process. This random process allows an observer to infer the agent's state of belief, but does not provide any evidence about the agent's identity. The *decision* phase consists of the agent's movement in one of the three directions. Our models assume that this behavior results from the agent's goal-directed action planning process and therefore provides evidence about the agent's identity. These assumptions allow us to model the agent's belief formation and action planning as two separate processes.

Generative model (agent's perspective) In each trial, the active agent makes a sequence of observations $\boldsymbol{o}$ about the location of objects in the environment. Figure 3-a is a graphical model representing the agent's theory about how these observations are formed from the true locations $\boldsymbol{s}$ of objects in the environment, whether or not the agent oriented towards each location $\boldsymbol{\theta}$ and the location of doors $\boldsymbol{w}$. From the agent's perspective, the true state $\boldsymbol{s}$ of the objects is unobserved and the other variables are observed.

Step one: belief inference In the first step, the agent has a prior belief that there is equal probability that each of the objects is in each of the rooms. The agent then uses the belief model (fig. 3-a) to update the probability that each object is in each room based on its sequence of observations, orientations, and its knowledge of the position of walls. We refer to the posterior distribution of s^k as the agent's *belief state* about the location of object k. For example, $p(s^1 = 4|\boldsymbol{o}^1, \boldsymbol{\theta}, \boldsymbol{w})$ is the agent's belief that the other agent (object $k = 1$) is in location four. Applying Bayes' rule gives the posterior probability (from the agent's perspective) that object k is in location x (Eq. 1). This posterior distribution is proportional to the likelihood of the observations $\boldsymbol{o}^k$, given that x was the true location of object k, multiplied by the prior probability that object k was in location x.

$$p(s^k = x|\boldsymbol{o}^k, \boldsymbol{\theta}, \boldsymbol{w}) \propto \prod_{i=1..L} p(o_i^k|s^k = x, \boldsymbol{\theta}, \boldsymbol{w}) \cdot p(s^k = x) \qquad (1)$$

Step two: action planning The belief state that the agent inferred in step one becomes an observed variable in the action planning model (fig. 3-b). The model assumes that the agent has a goal g^k with respect to each object and a

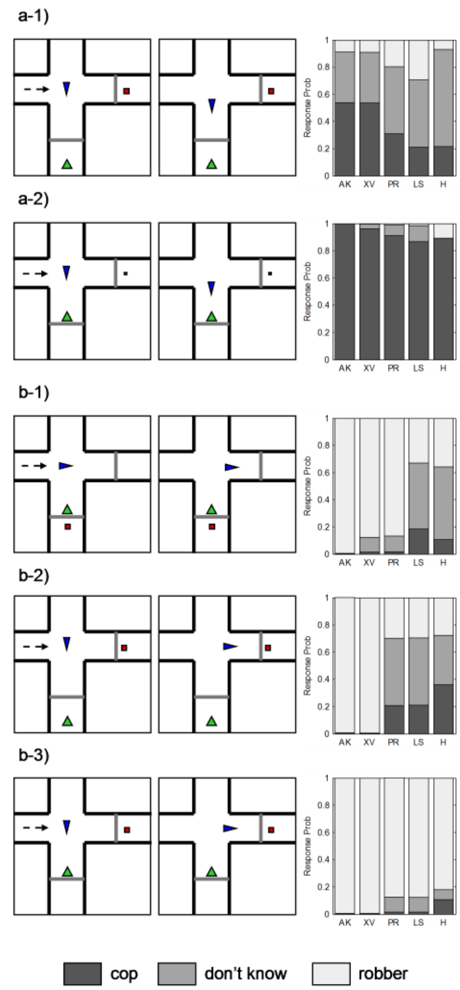

Figure 2. Example trials with human results (H) and model comparisons (LS, PR, XV and AK). For each example trial: The agent always enters the center room from the left (indicated by arrow); first frame shows direction agent oriented after reaching center (information gathering phase); second frame shows agent motion (decision phase).

priority c^k for that object of *primary*, *secondary* or *unimportant*. There were two agent types (cop and robber) and we assumed a constant configuration of goals and priorities for each type. Based on its goals, priorities, and belief state, the agent chooses an action a as a sample from a distribution that is proportional to the expected utility $\boldsymbol{u}$ of the actions (eq. 2).

$$P(a = m|\boldsymbol{s}, \boldsymbol{g}, \boldsymbol{c}) \propto \frac{\boldsymbol{u}(a = m, \boldsymbol{g}, \boldsymbol{c}, \boldsymbol{s})}{\sum_M \boldsymbol{u}(a = m, \boldsymbol{g}, \boldsymbol{c}, \boldsymbol{s})} \qquad (2)$$

Inverse model (observer's perspective) Figure 3-c depicts the inverse planning model from the perspective of the observer.

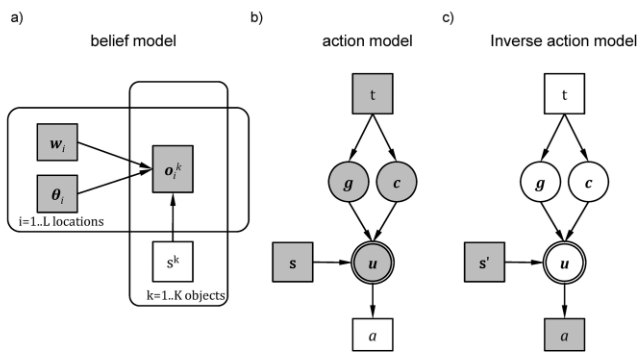

Figure 3. a) Belief model, b) generative action model (agent's perspective), c) inverse-planning model (observer's perspective). Shaded nodes are observed variables and unshaded nodes are unobserved variables.

Step one: belief inference The observer knows the true state of the environment $\boldsymbol{s}$ and infers the agent's belief state using a version of eq. 1 where s^k is replaced with s'^k. We use $\boldsymbol{s}'$ to represent the agent's belief state as inferred by the observer. We do not provide the graphical model for this step because it is identical to fig. 3-a with the exception that s^k is replaced with s'^k.

Step two: type inference In the second step (fig. 3-c) the variables that are known to the observer are the agent's belief state $\boldsymbol{s}'$ from step one and its action a. The observer's inference about the agent's type is represented by the posterior joint distribution $P(t, \boldsymbol{g}, \boldsymbol{c} | a, \boldsymbol{s}')$.

$$P(t, \boldsymbol{g}, \boldsymbol{c} | a, \boldsymbol{s}') \propto P(a | \boldsymbol{g}, \boldsymbol{c}, \boldsymbol{s}') P(\boldsymbol{g} | t) P(t) \quad (3)$$

Step three: response The observer chooses a response (*"cop"*, *"robber"*, or *"don't know"*) based on the posterior probability of the agent's type t.

See the online appendix[2] for a more detailed description of the computational theory and modeling – including a description of the agent's utility function and the observer response model.

Modeling

We developed four observer models based on the inverse planning framework—the first model provides a full description of our hypothesis about the mechanism by which human observers assign belief states to agents. Each of the last three models is a version of the full model in which we remove one of the constraints on belief inference.

Line-of-sight (LS) model This observer assumes that an agent's belief about the location of objects is a function of its orientation and the presence or absence of visual obstructions (doors) between the agent and the objects. For an agent to "see" an object it must orient towards the object and there must not be a closed door in front of the object.

Proximity (PR) model This observer does not require orientation for the formation of belief states. It assumes that an agent can "see" an object if it is in an adjacent room and not behind a closed door even if the agent does not orient towards the object.

X-ray vision (XV) model This observer assumes that an agent's belief about the location of objects is a function of its orientation only. For an agent to "see" an object it must orient towards the object – closed doors do not block its line-of sight.

All-knowing (AK) model This model corresponds to an observer that has no ToM. The observer represents the agent as having the same belief state about the environment that it has – in this case, complete and correct knowledge of the environment. It implements only steps two and three of the inverse model where the agent's belief state is equal to the actual state of the environment.

Model Comparison

Figure 4 shows the negative log-likelihood of model predictions for each participant based on a cross validation analysis. We used the responses from all but one participant to optimize a single parameter (σ) that relates to the response mechanism for each of the four observer models. We then used this learned parameter value when predicting the responses of the participant that was held out of the training data. We did this for every participant. The line-of-sight model made the best predictions for every participant, followed by the proximity model, x-ray model and finally the all-knowing model.

Qualitative model comparison Figure 2 provides a comparison of model behavior and human judgments in several illustrative conditions. All of the models tended to correspond to the human responses on trials in which the active agent had a clear line-of-sight to the other agent (figs. 2-a-2 & 2-b-3). When there was not a clear line of sight between the agent and the objects, the LS model, and to some extent the PR model, tended to perform better than the

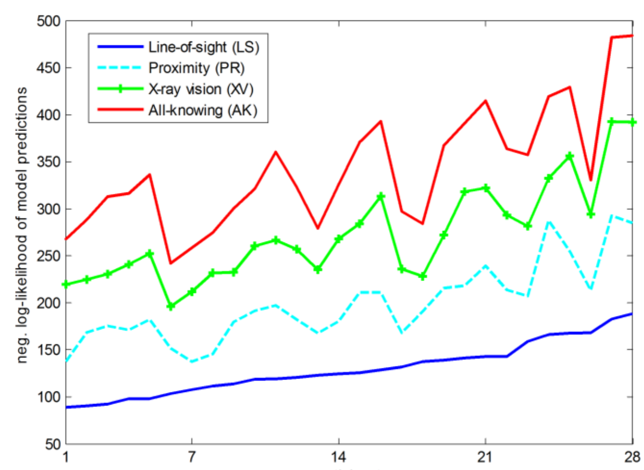

Figure 4. Negative log-likelihood of model predictions.

[2]https://webfiles.uci.edu/stauber/Tauber_Steyvers_CogSci2011_Appendix.pdf

AK and XV models.

Figure 2-a demonstrates the effect of a closed versus open door between the agents when the target agent is approaching the other agent. The LS and PR models, along with the humans, tended to respond with more uncertainty when there was a closed door (fig. 2-a-1) than the AK and XV models did. When there was an open door (fig. 2-a-2), the humans and all of the models gave predominantly *cop* responses.

Figure 2-b shows the effects of orientation and doors on human and model behavior when the target agent moves away from the other agent. When there is an open door but no orientation (fig. 2-b-1) the LS model and humans are uncertain, and the other three models give primarily *robber* responses. When the target agent orients towards the other agent but there is a closed door between them, the humans, LS and PR model responded with uncertainty. The AK and XV models gave *robber* responses. When the target agent oriented towards the other agent and there was an open door between them, the humans and all of the models gave primarily *robber* responses.

Discussion

We propose that observing an agent's actions in the context of the true state of its environment does not always provide enough information for an observer to infer its goals. Often, an observer needs to know something about the agent's state of belief in order to interpret its actions. We designed an experiment where observers watched a series of animations – in each of which it appeared that an agent moved in a certain direction in order to achieve its goals. Even though there were sets of multiple trials that had equivalent environmental states and agent actions, observers interpreted the agents' actions differently depending on what they thought the agent knew about the environment at the time that it made its decision.

Our results are consistent with the hypothesis that human observers infer an agent's belief state by using information about whether it has a clear line-of-sight to relevant aspects of the environment; and that these inferred belief states affect observers' interpretations of the agent's behavior.

We developed four graphical models that each make predictions about the structure of the process that humans use to infer the identity of agents in our experiment. We found that observer models that incorporated inferences about agents' beliefs outperformed an all-knowing observer model in describing human responses. Additionally, we found that all human responses were most consistent with the predictions of a line-of-sight model that required agents to both orient and have an obstruction-free line of sight towards a location in order to observe it. The correspondence between model predictions and human data was progressively worse when we 1) assumed agent's observed adjacent locations without orienting towards them (proximity model); 2) assumed visual obstructions did not impede observations (x-ray vision model); 3) assumed agents had complete knowledge of the environment (all-knowing model).

Our assumptions about the independence of the information gathering and decision phases simplified the model process. However, it is reasonable to argue that in more realistic situations the information gathering process would depend on the agent's goals and priorities. In this case an agent's information gathering behavior depends on where it has already looked, what it saw, and what its goals are.

Finally, there is a growing body of empirical evidence suggesting that ToM abilities may involve a combination of processes that are each used more or less effectively by human children and adults in different situations (Samson & Apperly, 2010). A new direction for future research is the development of a computational description of the cognitive process(es) involved in ToM that accounts for the wide range of failures and successes on ToM tasks by children and adults described in the empirical literature.

References

Baker, C. L., Goodman, N. D., & Tenenbaum, J. B. (2008). Theory-based social goal inference. In *Proceedings of the Thirtieth Annual Conference of the Cognitive Science Society* (pp. 1447-1452).

Baker, C.L., Saxe, R., & Tenenbaum, J.B. (2009). Action understanding as inverse planning. *Cognition*, 113, 329-349.

Doherty, M. J. (2008). *Theory of mind: how children understand others' thoughts and feelings*. Hove, UK: Psychology Press.

Feldman, J., and Tremoulet, P. D. (2008). The attribution of mental architecture from motion: towards a computational theory. Technical Report RuCCS TR-87, Department of Psychology, Rutgers University.

Gao, T., Newman, G. E., & Scholl, B. J. (2009). The psychophysics of chasing: A case study in the perception of animacy. *Cognitive Psychology*, 59(3), 154-179.

Heider, F., & Simmel, M. A. (1944). An experimental study of apparent behavior. *American Journal of Psychology*, *57*, 243-249.

Koller, D., Friedman, N., Getoor, L., and Taskar, B. (2007). Graphical models in a nutshell. In L. Getoor & B. Taskar (Eds.), *Introduction to Statistical Relational Learning.* Cambridge, MA: MIT Press.

Premack, D. and Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *The Behavioral and Brain Sciences*, 4:515–526.

Samson, D., and Apperly, I. A. (2010). There is more to mind reading than having theory of mind concepts: New directions in theory of mind research. *Infant and Child Development*, 19, 443-454.

Ullman, T.D., Baker, C.L., Macindoe, O., Evans, O., Goodman, N.D., & Tenenbaum, J.B. (2010). Help or hinder: Bayesian models of social goal inference. *Advances in Neural Information Processing Systems* (Vol. 22, pp. 1874-1882).