The Robot Olympics: Estimating and Influencing Beliefs About a Robot's Perceptual Capabilities

Matthew Rueben*, Eitan Rothberg†, Matthew Tang‡, Sarah Inzerillo§, Saurabh S. Kshirsagar*, Maansi Manchanda*, Ginger Dudley*, Marlena R. Fraune¶, and Maja J. Matarić*
*Department of Computer Science, University of Southern California, Los Angeles, California, USA.
Emails: {mrueben,skshirsa,maansima,vdudley,mataric}@usc.edu
†The Ohio State University, Columbus, Ohio, USA. Email: rothberg.10@osu.edu
‡University of California Berkeley, Berkeley, California, USA. Email: matthewtang@berkeley.edu
§Clarkson University, Potsdam, New York, USA. Email: inzeris@clarkson.edu
¶Department of Psychology, New Mexico State University, Las Cruces, New Mexico, USA. Email: mfraune@nmsu.edu

***Abstract*— People often hold inaccurate mental models of robots. When such misconceptions regard a robot's *perceptual capabilities*, they can lead to issues with safety, privacy, and interaction efficiency. This work is the first attempt to model users' beliefs about a robot's perceptual capabilities and make plans to improve their accuracy—i.e., to perform *belief repair*. We designed a new domain called the Robot Olympics, implemented it as a web-based game platform for collecting data about users' beliefs, and developed an approach to estimating and influencing users' beliefs about a virtual robot in that domain. We then conducted a study that collected user behavior and belief data from 240 online participants who played the game. Results revealed shortcomings in modeling the participant's interpretations of the robot's actions, as well as the decision making process behind their own actions. The insights from this work provide recommendations for designing further studies and improving user models to support belief repair in human-robot interaction.**

I. Introduction

People often hold inaccurate beliefs about robots. This is partly due to a lack of insider knowledge—users are rarely privy to the robot's programming or the design process that produced it. Instead, they can only observe the "system image"—how the robot looks and behaves—when developing their mental model of the robot to explain and predict its behavior [1]. Included in the user's mental model are their beliefs about what the robot can sense, detect, and recognize about the world—i.e., the robot's *perceptual capabilities*. These beliefs are also prone to inaccuracies, perhaps because of the "perceptual belief problem": that "the perceptual capabilities of robots are diverse and often difficult to infer" [2]. For example, users do not always notice or recognize a robot's sensors, even when they are visible, and it is often unclear when those sensors are recording data [3]. Also, users can struggle to understand robot perceptual capabilities that differ from their own [2].

Research has suggested that, to work well with robot teammates, people need accurate mental models of the robots, including what those robots know and intend to do [4]. Inaccurate mental models of robots can induce two types of problems in human-robot interaction. The first type of problem arises due to the user's misconceptions: the user might command the robot to do things it cannot do, might accidentally reveal personal information to the robot, or might find the robot too unpredictable. The second type of problem arises from the robot's ignorance of the user's misconceptions. The robot might proceed with the interaction under the incorrect assumption that the user knows its capabilities, causing inefficient actions or even safety risks. Robots can use shared mental models to improve the performance of human-robot teams [5]—in particular, a robot might need to estimate the user's beliefs about its perceptual capabilities to avoid these problems. If a belief estimate were used in action planning during a human-robot interaction, the robot could decide whether to work around the user's inaccurate beliefs or address them with teaching actions to improve subsequent interactions.

This work is, to the best of our knowledge, the first attempt to model a user's mental model of a robot's *perceptual capabilities*. It is also the first attempt to use belief estimates about perceptual capabilities to make plans to influence those beliefs through the robot's actions. This paper presents the following contributions. First, we introduce a novel experimental domain—the Robot Olympics—implemented as a web-based game with a virtual robot. Next, we introduce an approach to estimating and influencing the user's beliefs during the game toward improving the combined performance of the human-robot team. Rueben et al. previously reported preliminary findings from a pilot study [6]; this work reports findings from a new study with 240 online participants. The analysis of the study results revealed shortcomings in modeling both the participant's interpretations of the robot's actions and the decision making process behind their own actions. Using those insights, this work presents recommendations for designing further studies and improving user models to support belief repair in human-robot interaction.

II. Related Work

A. Forming Mental Models About Robots

There is relatively little research into how people form mental models of robots, especially of robot perceptual capabilities over longer interactions. Theories of mental

models from cognitive science that have been applied to how people understand, e.g., computer systems [7], have not yet been applied to human-robot interaction. A small number of human-robot interaction studies have looked at mental model formation. Stubbs et al. [8] tracked mental model formation over multiple interactions; they studied how museum employees came to understand a robot that was part of a museum exhibit over more than three months. Several other studies have focused on the development of mental models about the robot's perceptual capabilities in particular. Thellman and Ziemke, for example, found that people struggle to learn about perceptual capabilities that humans lack, and that watching a robot's behaviors over time aids that learning process while verbal descriptions of those behaviors do not [2].

A study by Rueben et al. was both long-term and focused on perceptual capabilities [9]; it involved a six-week interaction with a piece of mobile robotic furniture deployed in a public setting. Results showed that people can employ sophisticated reasoning to draw conclusions (or decide to hold back from drawing conclusions) about a robot's capabilities based on its behavior. Furthermore, robot behaviors were found to be especially important for inferring perceptual capabilities (sometimes incorrectly) when the robot's sensors were not clearly visible. This work focuses on a similar scenario via a robot—Kuri by Mayfield Robotics—with sensors that are not clearly visible to users, and is also inspired by the authors' recommendation to control the order in which robot behaviors are displayed to users to improve the accuracy of their mental models.

Distinct from previous work, this work is the first attempt for a robot to model a user's belief formation process about its perceptual capabilities. In the absence of existing models or frameworks, we designed a custom domain to help us predict and measure user beliefs throughout the interaction.

B. Mutual Modeling/Adaptation in Human-Robot Collaboration

Two agents maintaining models of each other has been called *mutual modeling* [10]. In the last decade, tools like POMCoP [11] and Cooperative Inverse Reinforcement Learning [12] have enabled an artificial agent to not only consider its estimate of a human teammate's mental model, but also to intentionally gather information to improve that estimate when needed. Nikolaidis and colleagues have developed a game-theoretic approach to this problem wherein the robot's objective is to maximize the team's joint reward over time in light of a human policy that the robot can both learn and influence [13], [14]. Our domain is similar in that it refrains from directly rewarding these learning and influencing processes so that the robot is only incentivized to do them if they improve the team's task performance. Instead of only modeling the user's policies, however, which are domain specific, we explicitly model the user's beliefs about the robot's perceptual capabilities; we model the user's policy separately as a function of those beliefs. Explicitly estimating beliefs allows the robot to identify specific misconceptions that the user holds and take actions to correct them.

Action selection is challenging in such domains because the robot must account for both the immediate rewards of its actions and their effects on the user's future decisions via influencing the user's mental model. Nikolaidis and colleagues have used a Partially Observable Markov Decision Process (POMDP) approach to select optimal actions for the robot. Solving the POMDP quickly becomes intractable, however, as the state and action spaces increase in size. The belief space in our domain is quite large because we are interested in user beliefs about multiple robot skills and combinations of skills, so we use a particle-based approach to state estimation with a Monte Carlo tree search for action selection.

III. Research Question and Scope

This work addresses the following research question: *"How can the robot improve the combined performance of the user-robot team by estimating and influencing the user's beliefs about its (the robot's) perceptual capabilities?"* We focused on using the robot's physical motions to implicitly communicate information about its perceptual capabilities; we did not enable the robot to talk about them explicitly via text-based or spoken dialogue. This work is especially relevant for interactions in which dialogue is problematic: e.g., in noisy environments, when interacting from a distance, when the interaction must be very brief, or even when dialogue would lead the user to overestimate the robot's natural language capabilities. Similarly, we restricted the robot to only using the human user's in-game choices to estimate the user's current beliefs.

IV. Human-Robot Interaction Domain: The Robot Olympics

We designed a simulated multi-turn human-robot collaboration game for our study. Each game board was designed such that the robot must communicate its perceptual capabilities to the user at certain times in the game to earn the highest possible team score. Without this information, the user is likely to make suboptimal choices and lose points for the team. We chose a web-based format with pre-recorded videos of the robot's actions to facilitate collecting a large sample of participants. Also, the user's actions were simple button presses in a web browser, which were easily and reliably detected by the virtual robot.

The user's teammate in the game was a virtual robot based on the Kuri robot by Mayfield Robotics. The name "Denise" was used to refer to the virtual robot throughout the study. Participants were introduced to a picture of a Kuri at the beginning of the game, and then viewed video recordings of the robot performing selected tasks.

The game was called The Robot Olympics, and was themed after the Olympic Games. It consisted of 2–5 "days", and on each "day" there were 4 events that the robot could attempt (see Table I). Of these, two were chosen for the robot to attempt: the user chose the first one, and then the virtual

robot chose the second one. After each event was chosen, a pre-recorded video of the robot attempting that event was displayed along with a message informing the user whether the robot succeeded and earned the points associated with that event. The day was over after the two choices and two videos.

Each event consisted of 1–3 tasks (see Table I); if the robot successfully completed all of them, the team earned the number of reward points associated with that event. If the robot failed at any of the tasks, the team earned zero points. The robot attempted the tasks in sequence, so if it failed at a task, users saw the failure as well as any successes that came before it. There were 6 different skills (i.e., perceptual capabilities) required by the tasks in the game. Most tasks required one skill to complete; a few tasks could be completed using either of two skills (i.e., a Boolean OR condition). The robot's success or failure at each task was determined only by its skills: if it had the required skill (or either of two sufficient skills for OR tasks) then it always succeeded; otherwise, it always failed. The skills that the robot *had* were LIDAR navigation, color tracking, face detection, and directional sound tracking; it *did not have* natural language understanding or object recognition. We chose these six perceptual capabilities because they are realistic possibilities for the Kuri robot. The robot knew which skills it had, and which were required to successfully complete each event. The user was not told the robot's actual skills, or the list of possible skills. On each day, the user was shown a cartoon diagram of each event that described what tasks the robot had to complete to earn that event's reward.

Each participant in the study was randomly assigned to one of four game boards. Each game board was a different sequence of the five days shown in Table I. The sequence for each game board is listed in the table's caption. Note that game board AB is composed of game board A followed by game board B, and that game board BA is the same composition in the opposite order.

Each game board was designed so that earning the highest possible team score required the robot to warn the user away from choosing one or two "trap" events. The two "trap" events are impossible for the robot—one requires natural language understanding, the other requires object recognition—but have much higher rewards than the other events on that day, designed to tempt the user. For each trap event, another event was included on a previous day that required the same skill, but with lower stakes: all the events on that day were worth fewer points than on the day of the trap. We thereby made it the optimal strategy for the robot to intentionally fail at the earlier, "teaching" event, scoring zero instead of a low number of points for that day, so that the user, learning from the robot's demonstrated failure, would avoid the trap event and instead score a higher number of points on the day of the trap.

The trap and teaching events can be seen in Table I. In game board A, event 3C is the trap and it costs the fewest points for the robot to intentionally fail at event 1C—the robot attempts and fails the natural language understanding task first, demonstrating to the user that it lacks that skill. In game board B, event 5C is the trap and the robot's only chance to help the user avoid it is to intentionally fail at event 4B, which will show the robot succeeding at the LIDAR navigation task, succeeding at the color tracking task, and then failing at the object recognition task. Game boards AB and BA each have both traps, but in opposite orders.

Day 1	Day 2	Day 3
1A: L & F 1pt	2A: C & N 8pts	3A: C & L 10pts
1B: C or S 3pts	2B: F or N 10pts	3B: L & C 10pts
1C: N & F 2.5pts	2C: S 7.5pts	3C: N 25pts
1D: none 2pts	2D: none 6pts	3D: C 7pts

Day 4	Day 5
4A: L 3pts	5A: C & S 6pts
4B: L & C & O 4pts	5B: O or F 5pts
4C: N & F 5pts	5C: C & O 20pts
4D: none 2pts	5D: none 5pts

TABLE I: Events in the Robot Olympics shown by day, including the robot perceptual capabilities needed to succeed at each event and the possible reward points. Game Board A = Days 1, 2, 3; Game Board B = Days 4, 5; Game Board AB = Days 1, 2, 3, 4, 5; Game Board BA = Days 4, 5, 1, 2, 3. L = LIDAR navigation, C = color tracking, F = face detection, S = directional sound tracking, N = natural language understanding, and O = object recognition.

V. Virtual Robot System: Approach and Implementation

We designed and implemented a system that estimated the user's beliefs about the robot's perceptual capabilities, and selected actions for the virtual robot that could influence those beliefs to maximize the combined score of the user-robot team. This system was our first attempt at producing reward-oriented decision making that is informed by an estimate of the user's beliefs about the robot's perceptual capabilities. The design of the user belief models leveraged our game design: we assumed that participants would play as we expected them to, and that the game would encourage the strategies that we had predicted. We designed the game such that modeling the user's beliefs would be relatively simple, which also kept the robot's actions explainable so we could more easily analyze the results.

A. State Representation and Initial Estimate

The goal of the state estimator was to estimate the user's current beliefs about the robot's perceptual capabilities. We modeled the user as having an independent belief about each of the six skills that the robot could have in the game, even though we never explicitly listed them to the participants. Furthermore, we modeled the user as believing that there are only two possible levels for each skill: the robot either

has the skill, or it does not. This yielded a total of 64 possible skill level combinations: 2 possible skill levels for each of the 6 different skills. Our model did not represent beliefs about intermediate skill levels, occasional malfunctions, or changing skill levels over time.

We modeled the user as considering all 64 possible skill level combinations, and maintaining a probability that each combination is the robot's actual set of skill levels. This was encoded by the virtual robot as a vector of 64 probabilities that always summed to 1. Each probability is the robot's estimate of the user's confidence that the robot actually has that set of skill levels.

This state representation makes the robot's belief space quite large, as the 64 probabilities could each vary independently. Discretizing the user's confidence into 128 tokens to distribute among the 64 possible skill combinations, for example, still yields $2 * 10^{51}$ (64 multichoose 128) belief states for the robot to consider. We chose to avoid this dimensionality challenge in our system by tracking just a single particle representing the robot's best guess of the user's belief state. For the initial value of this particle we assumed that the user starts the game considering all 64 skill combinations to be of equal probability, so we used a uniform probability distribution.

B. Transition Model

We modeled the user as understanding which skills are required to complete each task comprising each event from viewing the diagram of the event and the video of the robot attempting it. We also modeled the user as believing that the robot has a 100% success rate when it has the necessary skills to complete a task and a 0% success rate when it does not, which is how the game was programmed. Therefore, whenever the user witnessed the robot succeeding or failing at a task, the probabilities for the skill combinations that would not yield that result were set to zero and all remaining (nonzero) probabilities were normalized to sum to 1. This is a special case of a Bayesian update wherein the likelihood $p(success \mid skill\ combination_i) = either\ 0\ or\ 1$.

C. Observation Model

We modeled the user as understanding from the event diagrams which skills were required for the robot to succeed at the tasks in each event. We could therefore calculate the user's believed likelihood that the robot would succeed at each event by summing the probabilities in the robot's estimate of the user's belief state for all the skill combinations with which the robot would succeed at the event.

The virtual robot planned its actions under the assumption that the user would always play greedily and rationally—i.e., choosing the event with the highest expected value of reward on each day. We assumed that the user did not plan for future days because they were not told how many more days, if any, were left in the game. Assuming that the user used the expected value assumes a certain level of comfort with risk: we multiplied our current estimate of the user's believed likelihood of success for each event by that event's reward without weighting either number.

Because our system only tracked one particle representing the robot's best guess as to the user's belief state, we designed an *ad hoc* adjustment for this estimate when needed. Whenever the user chose a different event besides the one with the highest expected value, the virtual robot used a simple heuristic to adjust its estimate to make more sense of this choice. Specifically, the robot increased the probabilities of all skill combinations with which the chosen event is possible, and also decreased the probabilities of all skill combinations with which events that yield higher rewards are possible. These increases and decreases were all by the same amount: the difference in expected values between the event the user actually chose and the event the robot expected them to choose, multiplied by 0.01.

D. Action Selection

Our first goal for the action selector was for the robot to balance long-term payoffs with immediate rewards. The long-term payoffs would come from improving the accuracy of the user's beliefs to help them earn more points in subsequent days. The immediate rewards were the points from the event chosen by the robot on the current day, if the robot succeeded at the event. We considered using a POMDP to achieve this balance, but the large user belief space made this intractable, and each day of the game also had an abnormal sequence: observation, transition, action, transition.

Our second goal was for each action to be chosen in under five seconds so the game could be played without the user becoming impatient from waiting for the robot to take its turns. Although we pre-computed all of the virtual robot's decisions so the robot's policy would be the same for all participants, our system's online performance was fast enough to meet this requirement.

A Monte Carlo tree search (MCTS) was used to make each of the virtual robot's choices. A MCTS tree was pre-computed for each possible sequence of prior user choices that the virtual robot could encounter at any point in the game. Growing one tree to make one robot choice took an average of 2.6 seconds on a laptop computer. Whenever a leaf node was expanded, upper and lower bounds were calculated for the expected reward from each of the new leaf nodes using a greedy playout. The greedy playout obtained the total team score at the end of the game if the robot were playing greedily and the user were playing rationally. Greedy play for the robot meant always choosing the event that yields the highest reward on each day after first ruling out any that the robot cannot successfully complete. Rational play for the user meant always choosing the event with the highest expected reward, calculated as described above. While traversing the tree up to the leaf node, user actions were sampled according to the probability that the robot currently believes that the user will choose each event, and the robot chooses the event with the highest expected reward from all the tree traversals thus far that include that event.

E. Benchmark Policy for Comparison: "Ignoring Beliefs"

In the rest of this paper we refer to our system for estimating and influencing beliefs about a robot's perceptual capabilities as "our system."

We also implemented a second behavioral policy for the virtual robot to represent its behavior if it were not attempting to estimate or influence user beliefs. We refer to this policy as "ignoring beliefs." This policy implemented greedy play: on each day, the virtual robot chooses the event with the highest reward after first ruling out any that the robot cannot successfully complete. When using this policy, the robot always gets its highest possible rewards, but the user remains likely to choose our trap events (see Section IV above), which would greatly reduce the total team score.

VI. Design of System Evaluation Study

For the study, we used a 4 (game boards: A, B, AB, BA) x 2 (robot policy: our system vs. ignoring beliefs) between-subjects design.

A. Measures

Pre-Game Questionnaire: Before starting the game, participants answered questions about their age, gender, and experience with computers and robots. They also answered three questions about risk-taking behavior when playing games, but we did not analyze those responses for this paper.

During the game we recorded which events were chosen by the participant and the virtual robot, including whether the participant chose a trap event.

Post-Game Questionnaire: After finishing the last day of the game, participants answered several open-ended questions about their understanding of the game. They then answered six questions—one for each skill—assessing their beliefs about each of the six skills used in the game. For example, the question for directional sound tracking read: "[Rate your level of certainty] Denise is able to locate and drive towards the origin of a noise" ["Definitely cannot", "Probably cannot", "Maybe cannot", "Unsure", "Maybe can", "Probably can", and "Definitely can"].

Participants then answered four questions assessing whether they played the game the way we expected them to. Our expected responses are underlined: "Before choosing an event, I studied the four options until I understood them the best I could" ["Never", "Sometimes but not always", "Every time"]; "After I was done looking at the event option, I thought carefully about which one of them to choose" ["Strongly disagree", "Disagree", "Slightly disagree", "Neutral", "Slightly agree", "Agree", "Strongly agree"]; "How many times did you ignore the reward points when choosing an event? For example, choosing an event just to see if Denise could do it." ["Never", "Sometimes but not always", "Every time"]; and, "On average, how risky vs. safe were you when choosing an event? I.e., how much did you *gamble* to try to win more points when you weren't sure about Denise's abilities?" ["1. Safest - when possible, only chose events that I *knew* Denise could complete", "2. Safer", "3. Balanced", "4. Riskier", and "5. Riskiest - always chose the highest possible points if there was any chance that Denise would succeed"].

Attention Checks: To gauge participants' attention and effort, we read each participant's responses to open-ended questions and calculated the time they spent deciding between the four events on each day.

B. Procedure

Participants were recruited via Amazon Mechanical Turk (AMT), and then redirected to our website. After responding to the pre-game questionnaire, each participant watched a 5½-minute long instructional video explaining the game rules and interface. The video tried to encourage participants to play the way our system assumed they would play: learning the robot's capabilities by watching the videos carefully, balancing risk and reward each time they chose an event, and thinking only about rewards on the current day without planning for the future. After watching the instructional video, the participants played the game. A game board and robot policy were randomly assigned to each participant at the beginning of the game. After completing the game and responding to the post-game questionnaire, the participants were given a unique code to receive their payment on AMT. The first 5 participants were paid 9 USD based on an overestimate of the total game duration; the remaining 235 were paid 8 USD. The average duration of participation across all 240 participants (i.e., the AMT-reported "Average Time per Assignment") was about 43 minutes.

VII. Results

In Subsections VII-A to VII-C, we report the main results evaluating our system's performance. In Subsections VII-D to VII-G, we report additional performance metrics and assumption checks.

A. Participants

Two hundred and forty participants (96 women, 1 non-binary, 143 men; 40 aged 20–29, 114 aged 30–39, 48 aged 40–49, 28 aged 50–59, 8 aged 60–69, and 2 aged 70 or older) were recruited to test whether our system improved team performance by estimating and influencing the user's beliefs. Random assignments of these participants to the 8 (i.e., 4x2) experimental conditions are shown in Table II. Fewer than 10% of participants failed our attention checks by giving incoherent responses or taking fewer than 3 seconds to look at the event diagrams and choose one, so we chose not to exclude anybody's data from analysis.

B. Scores from Participant's Choices Only

Although our system attempted to maximize the combined score of the participant-robot team, we only report the participants' scores here to show that our system did not perform better than when ignoring beliefs even when the robot's scores are not lowering the combined team score. The robot's scores were lower for our system than when ignoring beliefs because the robot intentionally failed at certain events to attempt to improve the accuracy of the participant's beliefs.

Participants' game performance with our system vs. the control system that ignored beliefs is reported in Figure 1 and Table II. Table II shows that mean scores for our system were only better by one point or less (cf. the reward distribution within each day in Table I) for game boards A, B, and BA; for game board AB it was more than two points worse. None of these differences was statistically significant.

TABLE II: Mean Scores from Participant's Choices Only with Two-Tailed Student's t-test Results

Game Board	Mean Score from Participant's Choices Only		
	Our System	Ignoring Beliefs	t-test Results
A	**13.1** (n = 26)	**12.4** (n = 39)	$t(63) = 0.59$, $p = .56$
B	**5.2** (n = 34)	**4.2** (n = 35)	$t(67) = 1.20$, $p = .23$
AB	**14.8** (n = 22)	**17.1** (n = 37)	$t(57) = 1.35$, $p = .18$
BA	**19.7** (n = 22)	**19.4** (n = 25)	$t(45) = 0.15$, $p = .88$

TABLE III: Participant Beliefs at the End of the Game, and What Fraction of Participants Chose the "Trap" Events

Game Board	Mean Post-Game Belief				Two-tailed Student's t-test Results: Beliefs	Point-Biserial Correlation Coefficient test Results: Trap
	% of Participants Who Chose the "Trap" Event					
	Our System		Ignoring Beliefs			
Trap A: Natural Language Understanding						
A	5.31	52%	5.69	61%	$t(63) = 0.96$, $p = .34$	$r_{pb}(122) = .08$, $p = .36$
AB	4.23		5.19		$t(57) = 1.97$, $p = .05$*	
BA	4.55	36%	5.32	28%	$t(45) = 1.38$, $p = .17$	$r_{pb}(45) = -.09$, $p = .55$
Trap B: Object Recognition						
B	3.97	38%	3.49	50%	$t(67) = 1.03$, $p = .30$	$r_{pb}(114) = .13$, $p = .18$
BA	4.41		4.64		$t(45) = 0.41$, $p = .69$	
AB	4.09	59%	4.49	41%	$t(57) = 0.79$, $p = .44$	$r_{pb}(57) = -.18$, $p = .17$

C. Participant Beliefs at the End of the Game

We examined whether experiencing our system resulted in more accurate participant beliefs about the skills that were most highly incentivized by the game boards: natural language understanding for game boards with Trap A (i.e., A, AB, and BA) and object recognition for game boards with Trap B (i.e., B, AB, and BA). The robot lacked both of these skills, so lower beliefs on the 7-point response format are more accurate (see Section VI-A). Table III shows the results. For Trap A, participants who experienced our system had more accurate beliefs about natural language understanding than those in the ignoring beliefs condition for game board AB, but no effect of robot policy was found for game boards A and BA. For Trap B, our system was not found to affect participant beliefs about object recognition for game boards B, BA, or AB. Since we only collected self-reported beliefs at the very end of each game, there are not enough data to determine whether only certain events were not being interpreted as expected, or if there was a more systemic flaw in our assumptions.

D. Which Participants Chose "Trap" Events

Table III shows that no effect was found of our system on how many participants chose each of the two trap events described in Section IV.

A much stronger predictor of which participants would choose the trap events than robot policy was the participants' answer to the post-game question about risk taking (see Section VI-A for item text). The percentage of participants who fell into either of the traps increased monotonically with self-reported riskiness, from 20% of the 23 participants who answered "1. Safest" to 100% of the 16 who answered "5. Riskiest". Self-reported riskiness was also predictive of participants' other event choices, which made it difficult to evaluate whether the events altered participants' beliefs in the expected ways.

E. Accuracy of Belief Estimator

Participants' self-reported beliefs at the end of the game were not as we expected from what the events were designed to teach them. The overall correlation for all 240 participants and all 6 skills between our system's predictions at the end of the game and self-reports was low (Pearson's r = .27, Spearman's ρ = .26).

F. Factor Structure of Participant Beliefs

To analyze whether participants distinguished between the six robot skills, we calculated the intercorrelations between responses to the six post-game skill belief questions. All but two of the intercorrelations were low ($|r| <= .15$): beliefs about natural language understanding and object recognition were positively correlated (r = .47), as were beliefs about face detection and object recognition (r = .42). These results suggest that participants probably thought of the robot as having at least four or five distinct skills.

G. Accuracy of Assumptions About How Participants Choose Events

We also analyzed how well participants met our assumptions about how they would study the event options and make choices. Less than one third—69 (29%) of 240—met our assumptions as defined in Section VI-A by the underlined response options. These participants who met our assumptions did not constitute a large enough sample for statistical analysis, as there were as few as 5 participants in some of the 8 conditions. Instead, we performed a visual comparison of the histograms of scores from participants' choices only for these participants vs. those for all 240 participants. This did not reveal a clear and consistent increase or decrease in our system's performance. Therefore, if our system worked better for participants who met our assumptions, it was likely by only a small amount.

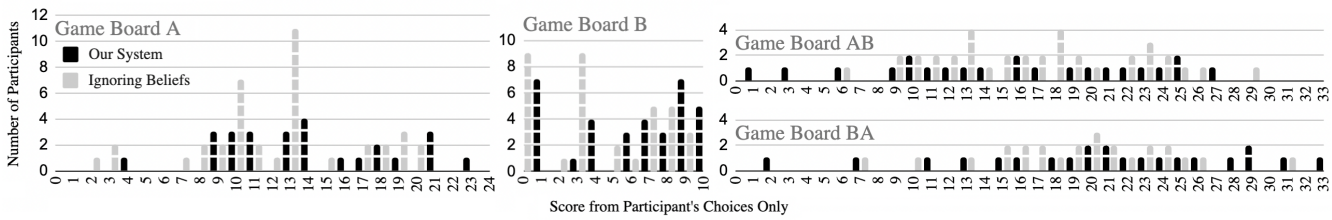

Fig. 1: Histograms of scores from the participant's choices only, separated by game board and robot policy.

VIII. Discussion

Our system mostly failed to achieve its goals: it did not improve the accuracy of participant beliefs about key skills for most game boards, and therefore did not help participants to avoid the trap events and score more points. Here we present recommendations for future systems based on insights from our findings. We also present recommendations for future studies in this and similar domains.

A. Modeling and System Design Recommendations

Regarding the **state representation and initial estimate**, the size of the robot's belief space, which spans all possible user beliefs, remains a challenge. Our system used a 64-dimensional continuous space spanning all possible values of the 64 independent probabilities that comprised the user's belief. One way this dimensionality could be reduced is if the user is assumed to believe that the robot does not use its skills in complicated combinations. In the case where the robot only uses its skills individually, for example, beliefs about N skills could be represented with only N probabilities—i.e., an independent probability for each skill. If beliefs about skill combinations are deemed necessary for the model, independent confidence levels could be tracked for just a subset of them, as it seems unlikely that a user could maintain probabilities for all possible skill combinations at once.

Also, we defined the user's belief space according to the six capabilities we knew the robot to have; another approach would be to discover this structure from interaction data. The factor structure of participants' beliefs from our study showed that this can be important, as there were two pairs of skills that participants did not fully distinguish. Such an analysis could also uncover other, unforeseen elements of the user's state that inform their estimate of the likelihood that the robot's action will succeed.

If the dimensionality of the robot's belief space is not too high, the robot's belief could be handled by a particle filter with a particle for each candidate user belief state.

In the future, the **transition model** should be learned from interaction data. One approach is to make a separate model for each robot action that estimates the likelihood in the user's mind that the robot will succeed given a certain skill combination. If combinations of skills are suspected to be used in the user's interpretation, perhaps in Boolean (e.g., AND, OR) relationships, a decision tree or a neural network could be used to capture this.

It might be important to use a different approach than the one just described if the user's belief updates are not approximately Bayesian, or if the user's interpretation of a robot action is influenced by the history of past robot actions they have observed, thereby breaking our Markov assumption. For example, the transition model may need to account for nonmonotonic reasoning [15] by which people make provisional belief updates that can be revised later in response to additional information.

Also, changes to the user's belief about one capability might itself cause their beliefs about other capabilities to change. For example, manipulating a robot's apparent speech capabilities can change someone's perception of its physical capabilities [16]. This kind of relationship would make it hard to build a model from interaction data of the believed likelihood of the robot's success as suggested above; beliefs that change when a robot action is viewed might not be part of the user's determination of that action's success likelihood.

The **observation model** should use not just the user's current beliefs, but also any additional contextual variables found to influence the user's actions (e.g. in our domain: riskiness, thoughtfulness). Also required is an accurate understanding of how the user will choose their actions in the task domain, which can be difficult to obtain for unstructured or novel scenarios. Lastly, the failure of our observation model to use participants' choices to correct our system's belief estimates could partly be because the scaling parameter for the heuristic adjustment method described in Section V-C was too small; that method would not have been necessary if a full particle filter had been used instead of a single particle.

The Monte Carlo tree search used for **action selection** worked as expected: when using our system, the robot intentionally failed at the teaching events to warn participants away from the trap events. Currently, the MCTS tries to maximize the combined score of the user-robot team; it could also be explicitly incentivized to minimize the error in the user's beliefs relative to the robot's actual capabilities.

Future work on this problem could also benefit from the literature on Intelligent Tutoring Systems. Many such systems improve learning via personalized "curriculum sequencing"—i.e., manipulating the order in which learning materials are presented—using a "student model" of the student's knowledge and learning process [17].

B. Study Design Recommendations

Our results showed that our assumptions about how people would interpret the robot's actions in this domain were

largely wrong. Future work should begin with a more qualitative approach—e.g., using a think-aloud protocol—to discover which factors are important for modeling the user.

Additionally, measurements of the user's beliefs about the robot's skills should be more frequent in future studies. After witnessing each robot action, participants could rate the likelihood that the robot would succeed at a series of hypothetical events. This would still avoid revealing the identities of the six skills during the game, and the hypothetical events could be chosen so as to measure beliefs about all six skills. Using this measure at the beginning of the interaction could establish the robot's initial estimate of the human's beliefs.

Future work should also study when it is better to communicate explicitly about perceptual capabilities via dialogue instead of using implicit communication as in this work.

IX. Conclusions and Outlook

This work was the first attempt to model a user's beliefs about a robot's perceptual capabilities and make plans to influence those beliefs via the robot's actions. We documented challenges with accomplishing these two goals from our study of 240 online participants who used the system we developed. We also presented recommendations for future systems, and for designing future studies to better understand users and to collect better data for training models.

In future work, robots could use dialogue to say what they can sense, or an instruction manual or explanatory video could provide the same information. Alternatively, robots might be designed such that their appearance and behaviors implicitly communicate their perceptual capabilities to users without the need for explicitly modeling user beliefs. Our work is important, however, in the cases where belief *repair* is needed. When a robot's dialogue is misunderstood or ignored, when there is significant variance in how people interpret a robot's actions, or as a safety precaution in case of unforeseen circumstances, robots will need to reason explicitly about how to make their perceptual capabilities clear to the people who need to know.

ACKNOWLEDGMENT

This material is based upon work supported by the National Science Foundation's National Robotics Initiative, Grant No. IIS-1528121; the USC Viterbi Summer Undergraduate Research Experience (SURE) Program; an Unfettered Research Grant from the Mistletoe Foundation; the New Mexico Space Grant Consortium (NMSGC) under NASA grant No. 80NSSC20M0034; and the New Mexico State University Office of the Vice President of Research and Graduate School (NMSU VPRGS). We are also grateful to Christopher Birmingham and Chris Denniston for help with the modeling work, and to Robert Lancaster for helping with the literature review.

Citation Diversity Statement

Recent work in several fields of science has identified a bias in citation practices: papers from women and other minority scholars are under-cited relative to the number of such papers in the field. To increase awareness of this problem, we state the distribution of citations in this work: 6% were published by a solely female team, 24% by a female/male team with a female lead author, 29% by a male/female with a male lead author, and 41% by a solely male team.

References

[1] D. Norman, *The design of everyday things: Revised and expanded edition*. Constellation, 2013.

[2] S. Thellman and T. Ziemke, "Do you see what I see? Tracking the perceptual beliefs of robots," *Iscience*, vol. 23, no. 10, 2020. [Online]. Available: https://doi.org/10.1016/j.isci.2020.101625

[3] M. K. Lee, K. P. Tang, J. Forlizzi, and S. Kiesler, "Understanding users' perception of privacy in human-robot interaction," in *Proceedings of the 6th International Conference on Human-Robot Interaction*. ACM, 2011, pp. 181–182.

[4] E. Phillips, S. Ososky, J. Grove, and F. Jentsch, "From tools to teammates: Toward the development of appropriate mental models for intelligent robots," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 55, no. 1. SAGE Publications Sage CA: Los Angeles, CA, 2011, pp. 1491–1495.

[5] F. Gervits, D. Thurston, R. Thielstrom, T. Fong, Q. Pham, and M. Scheutz, "Toward genuine robot teammates: Improving human-robot team performance using robot shared mental models." in *AAMAS*, 2020, pp. 429–437.

[6] M. Rueben, M. J. Matarić, E. Rothberg, and M. Tang, "Estimating and influencing user mental models of a robot's perceptual capabilities: Initial development and pilot study," in *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, 2020, pp. 418–420.

[7] S. J. Payne, "Users' mental models: The very ideas," in *HCI models, theories, and frameworks: Toward a multidisciplinary science*, J. M. Carroll, Ed. Morgan Kaufmann, Burlington, MA, 2003, pp. 135–156.

[8] K. Stubbs, D. Bernstein, K. Crowley, and I. R. Nourbakhsh, "Long-term human-robot interaction: The personal exploration rover and museum docents." in *AIED*, 2005, pp. 621–628.

[9] M. Rueben, J. Klow, M. Duer, E. Zimmerman, J. Piacentini, M. Browning, F. J. Bernieri, C. M. Grimm, and W. D. Smart, "Mental models of a mobile shoe rack: Exploratory findings from a long-term in-the-wild study," *ACM Transactions on Human-Robot Interaction (THRI)*, vol. 10, no. 2, pp. 1–36, 2021.

[10] S. Lemaignan and P. Dillenbourg, "Mutual modelling in robotics: Inspirations for the next steps," in *10th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2015, pp. 303–310.

[11] O. Macindoe, L. P. Kaelbling, and T. Lozano-Pérez, "POMCoP: Belief space planning for sidekicks in cooperative games," in *Eighth Artificial Intelligence and Interactive Digital Entertainment Conference*, 2012.

[12] D. Hadfield-Menell, S. J. Russell, P. Abbeel, and A. Dragan, "Cooperative inverse reinforcement learning," *Advances in neural information processing systems*, vol. 29, 2016.

[13] S. Nikolaidis, A. Kuznetsov, D. Hsu, and S. Srinivasa, "Formalizing human-robot mutual adaptation: A bounded memory model," in *11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2016, pp. 75–82.

[14] S. Nikolaidis, S. Nath, A. D. Procaccia, and S. Srinivasa, "Game-theoretic modeling of human adaptation in human-robot collaboration," in *12th ACM/IEEE Intl. Conf. on Human-Robot Interaction (HRI)*. IEEE, 2017, pp. 323–331.

[15] S. Khemlani and P. N. Johnson-Laird, "Why machines don't (yet) reason like people," *KI-Künstliche Intelligenz*, vol. 33, no. 3, pp. 219–228, 2019.

[16] E. Cha, A. D. Dragan, and S. S. Srinivasa, "Perceived robot capability," in *24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 2015, pp. 541–548.

[17] D. Leyzberg, S. Spaulding, and B. Scassellati, "Personalizing robot tutors to individuals' learning differences," in *9th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2014, pp. 423–430.