The Impact of POMDP-Generated Explanations on Trust and Performance in Human-Robot Teams

Ning Wang and David V. Pynadath
Institute for Creative Technologies
University of Southern California
nwang@ict.usc.edu,pynadath@usc.edu

Susan G. Hill
U.S. Army Research Laboratory
Aberdeen Proving Ground, MD
susan.g.hill.civ@mail.mil

ABSTRACT

Researchers have observed that people will more accurately trust an autonomous system, such as a robot, if they have a more accurate understanding of its decision-making process. Studies have shown that hand-crafted explanations can help maintain effective team performance even when the system is less than 100% reliable. However, current explanation algorithms are not sufficient for making a robot's quantitative reasoning (in terms of both uncertainty and conflicting goals) transparent to human teammates. In this work, we develop a novel mechanism for robots to automatically generate explanations of reasoning based on Partially Observable Markov Decision Problems (POMDPs). Within this mechanism, we implement alternate natural-language templates and then measure their differential impact on trust and team performance within an agent-based online testbed that simulates a human-robot team task. The results demonstrate that the added explanation capability leads to improvement in transparency, trust, and team performance. Furthermore, by observing the different outcomes due to variations in the robot's explanation content, we gain valuable insight that can help lead to refinement of explanation algorithms to further improve human-robot interaction.

General Terms

Algorithms

Keywords

Human-robot interaction, POMDPs, explainable AI, trust

1. INTRODUCTION

Robots are increasingly teaming with humans in complex real-world tasks, ranging from search and rescue to space exploration [2, 3]. Although the ever-improving capabilities of robotic systems may lead to improved team capabilities, they also create challenges that need to be overcome before such hybrid partnerships can achieve their full potential [1]. When robots are more suited than humans for a certain task, then we want the humans to trust the robots to perform that task. When the robots are less suited, then we want the humans to appropriately gauge the robots' ability and perform the task themselves. Failure to do so results in *disuse* of robots in the former case and *misuse* in the latter [27]. Real-world case studies and laboratory experiments show that failures in both cases are common [20].

Research has also shown that people are more likely to avoid such failures if they have an accurate understanding of the robot's decision-making process [19]. Unfortunately, as robots gain complexity and autonomy, it is increasingly challenging for humans to understand their decision processes. Successful human-robot interaction (HRI) therefore hinges on the robot's ability to make its decision-making process transparent to the people it works with. Hand-crafted explanations have shown to be effective in providing such transparency [7].

However, such hand-crafted explanations do not scale to the sophisticated reasoning that robots currently perform to handle the uncertainty and conflicting goals within their task environments. Many robotic platforms use Partially Observable Markov Decision Problems (POMDPs) [14], whose quantitative transition probabilities, observation probabilities, reward functions, and decision-making algorithms have proven successful in many robotic domains, such as navigation [4, 18] and HRI [28]. Unfortunately, the quantitative nature of these models and the complexity of their solution algorithms also make POMDP-reasoning opaque to potential human teammates.

In our work, we develop algorithms that can generate natural-language explanations from POMDP-based reasoning. We thus draw inspiration from the aims of prior researchers in "explainable AI" [33, 40], but within the novel context of decision-theoretic reasoning with uncertain beliefs. We build our algorithms on top of a multiagent social simulation framework, PsychSim [21, 29], that includes transparency of the various components of a POMDP model (e.g., beliefs, observations, outcome likelihoods). By grounding explanations in the agent's decision-making process, we can automatically generate a space of possible explanation content and measure their impact on the human-robot interaction (Section 4).

To quantify the effectiveness of different explanation content in achieving the desired transparency, we implemented an experimental testbed to simulate an HRI scenario (Section 5). This virtual human-robot simulation teams a robot with a human counterpart in reconnaissance missions [38]. The robot is modeled as a POMDP, with beliefs and observations of its surroundings, goals (e.g., mission objectives), and actions to achieve those goals. We conducted a study

Appears in: *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2016), J. Thangarajah, K. Tuyls, C. Jonker, S. Marsella (eds.), May 9–13, 2016, Singapore.*

where people interacted with different versions of the robot, where we varied its ability and its explanation content. The empirical results quantify the degree to which the explanations impacted transparency, human-robot trust, and overall team performance (Section 6). By examining people's behaviors over different combinations of the robot's ability and explanation content, we discuss the implications of the results and directions for future work (Section 7).

2. RELATED WORK

Our current work follows a long history of similar explorations of automated explanation mechanisms, especially within the context of expert systems [33]. While most of this work operated on rule- and logic-based systems, there has been more recent work on generating explanations based on Markov Decision Problems (MDPs) [8, 15]. However, there has been no work on explaining POMDP-based policies, where the system operates under uncertainty about the state of the world. Furthermore, there has been little empirical evaluation of the impact of these explanations on human-machine trust, although the existing data suggest that explanations do increase user acceptance of expert systems [40]. This limited evidence is encouraging as to the potential success of applying a general-purpose explanation on top of a robot's decision-making process.

The need for such explanations is evidenced by existing HRI studies that have shown that a human's ability to understand its robot teammate has a clear impact on trust [19]. Explanations have shown to contribute to that understanding in a way that provides transparency and improves trust [7]. Our goal is to create an automated, domain-independent method for generating explanations that has the same impact as the manually crafted explanations used in prior work.

Looking beyond the AI and HRI literature, we can find a large number of studies that measure the impact of various forms of explanation on people's perceptions of risks and uncertainties when making decisions. A survey of these studies across multiple domains indicates that "People prefer numerical information for its accuracy but use a verbal statement to express a probability to others." [36]. This finding led to a recommendation to include a numeric representation in any communication informing a person of the uncertainties underlying a decision. On the other hand, one of the studies in the survey contrasted a numeric representation of uncertainty with more anecdotal evidence and found that the numeric information carried less weight when both types were present [12]. A study of risk communication in medical trade-off decisions showed that people performed better when receiving numeric expressions of uncertainty in percentage (67%) rather than frequency (2 out of 3) form [39]. This same study also found that people expressed a preference for information "as words" rather than "as numbers". It is therefore clear that both percentage and verbal expressions of uncertainty have value in conveying uncertainty, but it is less clear what form makes the most sense in an HRI context. In translating our robot's reasoning into a human-understandable format, our explanation algorithms use natural-language templates inspired by these various findings in the literature.

3. POMDP MODEL OF AN HRI SCENARIO

We have implemented the explanation algorithms using PsychSim [21, 29], which combines two established agent technologies: decision-theoretic planning [14] and recursive modeling [9]. Decision-theoretic planning provides an agent with quantitative utility calculations that allow it to assess tradeoffs between alternative decisions under uncertainty. Recursive modeling gives the agent a theory of mind, allowing it to form beliefs about the human users' preferences, factor those preferences into its own decisions, and update its beliefs in response to observations of the users' decisions. The combination of decision theory and theory of mind has enabled PsychSim agents to operate in a variety of human-agent interaction scenarios [13, 16, 17, 23, 26].

PsychSim agents generate their beliefs and behaviors by solving POMDPs [6, 14]. In precise terms, a POMDP [14] is a tuple, $\langle S, A, P, \Omega, O, R\rangle$, that we describe here in terms of an illustrative HRI scenario [37]. In it, a human teammate works with a robot in reconnaissance missions to gather intelligence in a foreign town. Each mission involves the human teammate searching eight buildings in the town. The robot serves as a scout, scans the buildings for potential danger, and relays its findings to the teammate. Prior to entering a building, the human teammate can choose between entering with or without equipping protective gear. If there is danger present inside the building, the human teammate will be fatally injured without the protective gear. As a result, the team will have to restart from the beginning and re-search the entire town. However, it takes time to put on and take off protective gear (e.g., 30 seconds each). So the human teammate is incentivized to consider the robot's findings before deciding how to enter the building. In the current implementation, the human and the robot move together as one unit through the town, with the robot scanning the building first and the human conducting a detailed search afterward. The robot has a NBC (nuclear, biological and chemical) weapon sensor, a camera that can detect armed gunman, and a microphone that can listen to discussions in foreign language.

Within the POMDP model of this scenario, the state, S, consists of objective facts about the world, some of which may be hidden from the robot itself, such as the separate locations of the robot and its human teammate, as well as the presence of dangerous people or chemicals in the buildings to be searched. The state also includes feature-value pairs that represent the human teammate's health level, any current commands from the teammate, and the accumulated time cost so far.

The robot's available actions, A, correspond to the possible decisions it can make. Given its search mission, the robot's first decision is where to move to next. We divide the environment into a set of discrete waypoints, so the robot's action set includes potentially moving to any of them. Upon arrival, the robot then makes a decision as to whether to declare a location as safe or unsafe for its human teammate. For example, if the robot believes that armed gunmen are at its current location, then it will want its teammate to take adequate preparations (e.g., put on body armor) before entering. Because there is a time cost to such preparations, the robot may instead decide to declare the location safe, so that its teammates can more quickly complete their own reconnaissance tasks.

We model the dynamics of the world using a transition probability function, P, that captures the possibly uncertain effects of the robot's actions on the subsequent state.

We simplify the robot's navigation task by assuming that a decision to move to a specific waypoint succeeds deterministically. However, we could relax this assumption to decrease the robot's movement ability, as is done in more realistic robot navigation models [4, 18]. The robot's recommendation decision affects the health of its teammate, although only stochastically, as its teammate may not follow its recommendation. Instead, a recommendation that a building is safe (unsafe) has a high (low) probability of decreasing the teammate's health if there are, in fact, chemicals present.

The robot has only indirect information about the true state of the world. Within the POMDP model, this information comes through a subset of possible observations, Ω, that are probabilistically dependent (through the observation function, O) on the true values of the corresponding state features. We make some simplifying assumptions, namely that the robot can observe the location of itself and its teammate with no error (e.g., via GPS). However, it cannot detect the presence of armed gunmen or dangerous chemicals with perfect reliability or omniscience. Instead, it receives a local reading about their presence (or absence) at its current location. For example, if dangerous chemicals are present, then the robot's chemical sensor will detect them with a high probability. There is also a lower, but nonzero, probability that the sensor will not detect them. In addition to such a false negative, we can also model a potential false positive reading, where there is a low, but nonzero, probability that it will detect chemicals even if there are none present. By controlling the observations that the robot receives, we can manipulate its ability in our testbed.

Partial observability gives the robot only a subjective view of the world, where it forms beliefs about what it thinks is the state of the world, computed via standard POMDP state estimation algorithms. For example, the robot's beliefs include its subjective view on the presence of threats, in the form of a likelihood (e.g., a 33% chance that there are toxic chemicals in the farm supply store). Again, the robot derives these beliefs from its local sensor readings, so they may diverge from the true state of the world. By decreasing the accuracy of the robot's observation function, O, we can decrease the accuracy of its beliefs, whether receiving correct or incorrect observations. In other words, we can also manipulate the robot's ability by allowing it to over- or under-estimate the accuracy of its sensors.

We instantiate the human-robot team's mission objectives within the POMDP's reward function, R, which maps the state of the world into a real-valued evaluation of benefit for the agent. The highest reward is earned in states where all buildings have been explored by the human teammate. This reward component incentivizes the robot to pursue the overall mission objective. There is also an increasingly positive reward associated with level of the human teammate's health. This reward component punishes the robot if it fails to warn its teammate of dangerous buildings. Finally, there is a negative reward that increases with the time cost of the current state. This motivates the robot to complete the mission as quickly as possible. By providing different weights to these goals, we can change the priorities that the robot assigns to them. For example, by lowering the weight of the teammate's health reward, the robot may allow its teammate to search waypoints that are potentially dangerous, in the hope of searching all the buildings sooner. Alternatively, lowering the weight on the time cost reward might motivate the robot to wait until being almost certain of a building's threat level (e.g., by repeated observations) before recommending that its teammate visit anywhere. By varying the relative weights of these different motivations, we can manipulate the benevolence of the robot toward its teammate in our testbed.

The robot can autonomously generate its behavior based on its POMDP model of the world by determining the optimal action based on its current beliefs, b, about the state of the world [14]. Rather than perform an offline computation of a complete optimal policy, π, over all possible beliefs, we instead take an online approach so that the robot computes the optimal decision with respect to only its current beliefs, $\pi(b)$ [31]. The robot uses a bounded lookahead procedure that seeks to maximize expected reward by simulating the dynamics of the world from its current belief state. In particular, the robot will consider declaring a building dangerous or safe (i.e., recommending that its teammate put protective gear on or not). It will combine its beliefs about the likelihood of possible threats in the building with each possible declaration to compute the likelihood of the outcome, in terms of the impact on the teammate's health and the time to search the building. It will finally combine these outcome likelihoods with its reward function and choose the option that has the highest reward.

4. POMDP-GENERATED EXPLANATIONS

By exposing different components of the robot's POMDP model, we can make different aspects of its decision-making transparent to its human teammate. We create natural-language templates to translate the contents of a POMDP model into human-readable sentences:

- A: The robot can make a decision as to whether to declare the building safe or not and communicate its chosen action, e.g., "I think the doctor's office is safe." The string representation of each action, $a \in A$, is a domain-specific string. Upon making its decision, the robot chooses the string corresponding to its current choice, $\pi(b)$.
- S: The robot can also communicate the level of uncertainty underlying its beliefs, e.g., "I am 67% confident about this assessment," if it believed that the probability of the doctor's office being safe was 67%. We use a template that includes a variable indicating which element(s) of the factored state representation the robot should substitute for the probability reference. In this case, the only variable of interest is the robot's $b(\text{safe}_X = \text{True}, \text{my location} = X)$. When generating such an explanation, the robot will compute the indicated belief and insert it into the natural-language template.
- P: The robot can also reveal the relative likelihood of possible outcomes, e.g., "There is a 33% probability that you will be injured if you enter the doctor's office without protective gear." Here, the robot will weigh the possible outcomes by its belief in the hidden state. In particular, it can compute $\sum_s b(s) \Pr(\text{health} \downarrow | s, \text{no protection})$ to instantiate this template.
- Ω: Communicating its observation can reveal information about its sensing abilities, e.g., "My NBC sensors have detected traces of dangerous chemicals." We write domain-specific strings for each possible observation, $\omega \in \Omega$.

- O: Beyond the specific observation it received, the robot can also reveal information about how it models its own sensor capabilities, e.g., "My image processing will fail to detect armed gunmen 30% of the time." In this case, we combine the domain-specific templates for each possible observation, $\omega \in \Omega$, with the appropriate observation function value:

$$\frac{\sum_{\omega \neq \text{gunmen}} \sum_{s|\text{gunmen}_x,\text{my location}=X} \sum_a O(s,a,\omega)}{\sum_{\omega} \sum_{s|\text{gunmen}_x,\text{my location}=X} \sum_a O(s,a,\omega)}$$

.

- R: By communicating the expected reward outcome of its chosen action, the robot can reveal its benevolence (or lack thereof) toward its teammate, e.g., "I think it will be dangerous for you to enter the informant's house without putting on protective gear. The protective gear will slow you down a little." The template here relies on factored rewards, allowing the robot to compute separate expected rewards, $E[R]$, over the goals of keeping its teammate alive and achieving the mission as quickly as possible. We write domain-specific templates for each goal, for both the positive and negative cases. The robot then computes the separate $E[R]$ values and chooses the appropriate template depending on whether the value is positive or negative.

These templates provide a variabilized mechanism for specifying natural-language forms offline that can be instantiated by the robot at runtime based on its current beliefs. These various template formats can be used for any POMDP. We thus ensure that the results can be re-used by other researchers studying other HRI domains as well.

5. EVALUATION

5.1 Simulation Testbed for HRI

We implemented an online version of our HRI scenario to study the impact of these automatically generated explanations on trust and team performance. The testbed can be accessed from a web browser. The testbed's server executes the POMDP to both maintain the state of the simulated mission and to generate decisions for the robot. These are displayed on the participant's web browser, which sends decisions made by the participant back to the testbed's server.

5.2 Participants

We recruited 220 participants from Amazon Mechanical Turk (AMT). The participants had previously completed 500 or more jobs on AMT and had a completion rate of 95% or higher. All participants were located in the USA.

5.3 Design

We used the online testbed to evaluate how the different POMDP-generated explanations from Section 4 impact trust and team performance. We conducted two iterations of the study. For the sake of clarity, we will describe the methodology of the two iterations together. The first iteration of the study primarily focused on whether explanations can build transparency, establish a proper level of trust, and improve task performance. There were two independent variables for the first iteration of the study: *ability* and *explanation*. The *ability* variable had two levels: low and high. The *explanation* variable also has two levels: no explanation, and explanation of two sensor readings. After preliminary analysis of the first iteration of the study (details in Section 6.1), we extended the *explanation* variable to include two additional types of explanations: explanation of three sensor readings, confidence-level explanations. The two iterations of the study all combine to form a 2x4 design.

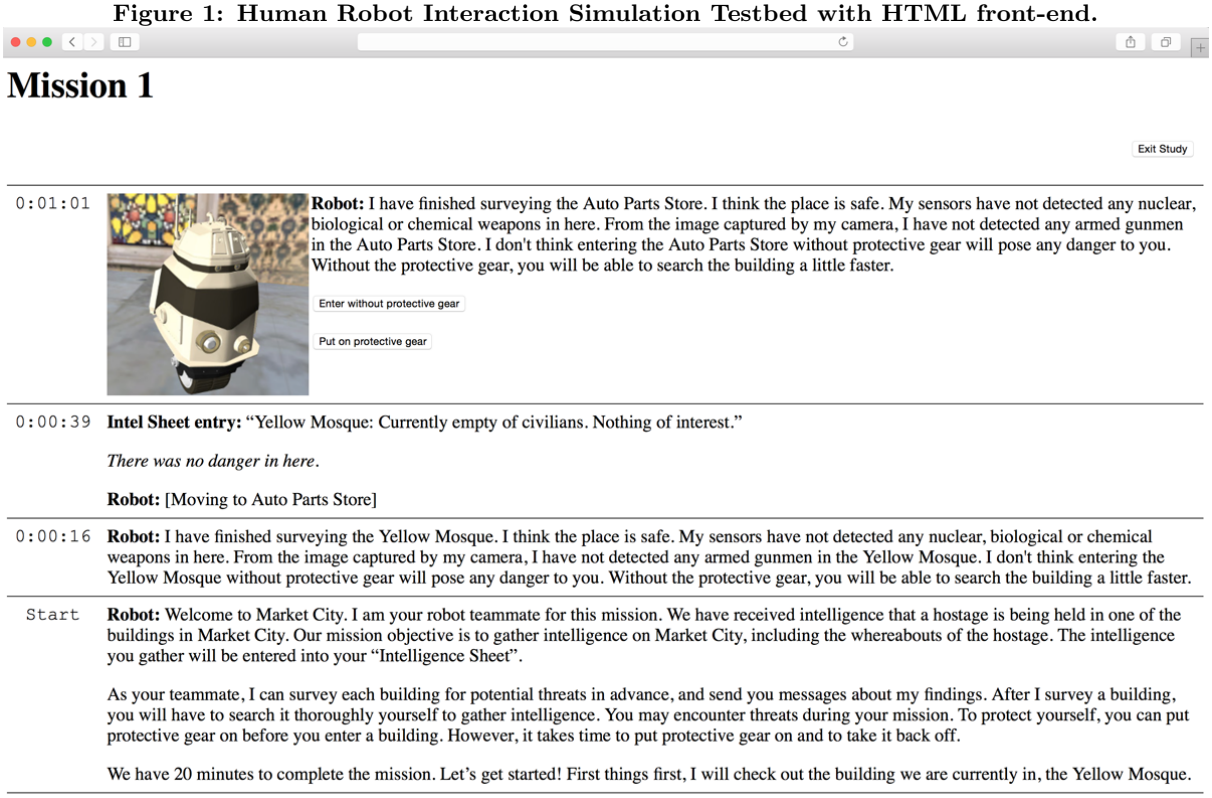
As already mentioned, the *ability* variable has two levels: a low-ability robot vs. a high-ability robot. Regardless of the ability of the robot, the human teammates will learn the correctness of the robot's individual decisions upon entering the buildings themselves.

- **High Ability** The robot with high ability makes the correct decision 100% of the time.
- **Low Ability** The robot with low ability has a faulty camera and makes only false-negative mistakes, e.g., not detecting armed gunmen in the simulation. The other simulated sensors (e.g., NBC weapon detector and microphone) and the robot's decision-making capability remain intact. As a result, the low-ability robot will occasionally give an incorrect "safe" assessment.

The *explanation* variable has a total of four levels: no explanation, explanation of two sensor readings, explanation of three sensor readings, and confidence-level explanation. At all four levels, the robot informs its teammate of its decision, derived from the A template from Section 4 (e.g., "I have finished surveying the doctor's office. I think the place is safe."). Under the conditions where explanations are offered, the robot augments this decision with additional information that should help its teammate better understand its ability (e.g., decision-making and sensing), one of the key dimensions of trust [22].

- **NoExp** In the No Explanation condition, the robot only informs its teammate of its decisions. One such communication from our scenario would be: *"I have finished surveying the Cafe. I think the place is safe."*
- **Exp2Sensor** In the Explanations of Two Sensor Observations condition, these explanations augments the decision message with non-numeric information about the robot's sensing capability. In this case, the sensing capability is limited to the NBC sensor and the camera – the only two sensors used by the PsychSim agents. Section 4's Ω template thus provides the teammate with the robot's observations from these two sensors. One such communication with both decision and explanation from our scenario would be: *"I have finished surveying the Cafe. I think the place is dangerous. My sensors have detected traces of dangerous chemicals. From the image captured by my camera, I have not detected any armed gunmen in the Cafe. I think it will be dangerous for you to enter the Cafe without protective gear. The protective gear will slow you down a little."* Although these explanations can potentially actually help the robot's teammate understand which sensors are working correctly (e.g., the NBC sensor) and which ones are not (e.g., the faulty camera), they do not actually help the teammate decide what to do with sensor readings from the Camera. This is because the robot, particularly the one in the Low Ability condition, has a faulty camera that makes false-negative mistakes. This means that when the robot reports no danger found by the camera, the teammate still doesn't know if they should put on the protective gear or not.

Figure 1: Human Robot Interaction Simulation Testbed with HTML front-end.



- **Exp3Sensor** In the Explanations of Three Sensor Observations, the explanations again augments the decision message with non-numeric information about the robot's sensing capability – the NBC sensor, the camera and the microphone. Section 4's Ω explanation provides the teammate with the robot's observations from these two sensors. One such communication with both decision and explanation from our scenario would be: *"I have finished surveying the Cafe. I think the place is safe. My sensors have not detected any NBC weapons in here. From the image captured by my camera, I have not detected any armed gunmen in the cafe. My microphone picked up a friendly conversation."* These explanations will thus potentially help the robot's teammate understand which sensors are working correctly and which ones are not, and help them decide what to do in case of camera failure. For example, while a faulty camera may not be able to detect armed gunman, the microphone is capable of picking up a suspicious conversation.
- **ExpConf** In the Confidence-Level Explanations condition, the confidence-level explanations augment the decision message with additional information about the robot's uncertainty in its decision. Section 4's S template incorporates the robot's probabilistic assessment of the hidden state of the world (e.g., the presence of threats) on which it bases its recommendation.[1] One example of a confidence-level explanation would be: *"I have finished surveying the Cafe. I think the place is dangerous. I am 78% confident about this assessment."* Because the low-ability robot's one faulty sensor will lead to occasional conflicting observations, it will on those occasions have lower confidence in its erroneous decisions after incorporating that conflicting information into its beliefs.

[1]Probability and confidence are generally different concepts. We used the probability as an approximation of the robot's confidence level.

The study is a between-subject design. Each participant interacted with one of the eight simulated robots. In the first iteration of the study, we assigned 30 participants to each condition. In the second iteration of the study, we assigned 25 participants to each condition.

5.4 Procedure

Each participant first read an information sheet about the study and then filled out the background survey. Next, participants worked with a simulated robot on three reconnaissance missions. After each mission, participants filled out a post-mission survey. Each participant worked with a robot with the same ability and communication (e.g., low ability and communicates with confidence-level explanations) throughout the three missions. Participants were randomly assigned to team up with one of the eight robots. The study was designed to be completed in 90 minutes. Participants were compensated with $10 for their participation.

5.5 Measure

The Background Survey includes measures of the demographic information, education, video game experience, military background, predisposition to trust [24], propensity to trust [25], complacency potential [30], negative attitude to-

wards robots [34] and uncertainty response scale [10]. In the Post-Mission Survey, we have designed items to measure participants' understanding of the robot's decisions and decision-making process. A sample item from this measure is "I understand the robot's decision-making process". We modified items on interpersonal trust to measure trust in the robot's ability, benevolence and integrity [22]. We also included the NASA Cognitive Load Index [11], Situation Awareness Rating Scale [35], trust in oneself and teammate [30], and trust in robots [32]. The Post-Mission survey was filled out after each mission (3 missions total in the study). We have also collected interaction logs from the online testbed. Based on the log data, we compute measures of team performance: mission success, percentage of correct decisions, and compliance (e.g., percentage of the robot's decision adopted by its teammate). The Post-Mission survey data and the log data provide coarse measures of how trust changes over time, which is beyond the scope of this paper. In this paper, the analysis includes measures on team performance, and trust in the robot's ability [22] and the scale we designed on the understanding of the robot's decision-making process presented in the Post-Mission Survey.

6. RESULTS

We excluded data from 18 participants due to incomplete entries (e.g., participants skipped survey questions or left the simulations). Although not part of the experiment manipulation, a closer examination revealed that the incomplete entries only occurred in conditions where explanations were offered. As a result, 202 participants are included in the analysis. The participants average 33.4 years old. 42% of the participants are female and 58% participants are male. 5 participants answered that they had worked with an automated squad member (such as a robot) before. 3 participants had reconnaissance or search and rescue training, and 1 was actually involved in such missions. Only 1 participant was an active service member.

We measured participants' predisposition to trust, propensity to trust, complacency potential, negative attitude towards robots and uncertainty responses. We did not find any significant main or interaction effect of the independent variables on any of these scales. Studying the impact of individual differences on trust is not the focus of this paper. Instead, we will focus on comparing the impact of different explanation algorithms on trust [22] and team performance. In the analysis presented here, we focus on self-reported perceptions of the robot's ability and behavioral measures of task performance (e.g., mission success rates, correct decisions percentage). Self-reported measures are calculated by averaging survey responses gathered after each mission (3 missions total). Behavioral measures are based on log data from all 3 missions as well.

The four dependent variable included in the analysis are:

- **Trust in Robot's Ability** Trust in the robot's ability, benevolence and integrity is measured by modifying an existing scale [22] that measures these three factors of trustworthiness. Each factor of trust is calculated by averaging corresponding Post-Mission Survey items collected after each of the 3 missions. The explanations compared in this paper are designed to influence perceptions of the ability factor of trust, and do not explicitly target the benevolence and integrity factors of trust. So we focus on only the ability component of trust in this paper. The value ranges from 1 to 7.
- **Understand Robot's Decisions** This is measured using 1âĂŞ7 Likert scale items on the understanding of the robot's decision-making process, designed by the researcher. A sample item from this measure is "I understand the robot's decision-making process". The measure is calculated by averaging responses to corresponding survey items in the Post-Mission Survey after each of the 3 missions. The value range from 1 to 7.
- **Mission Success Percentage** This team-performance measure is extracted from a line in the interaction log indicating whether the mission ended in success/failure, then divided by the total number of missions (3) in the study. The value ranges from 0 to 100.
- **Percentage of Correct Decisions** This variable is measured using log data. It is calculated by dividing the total number of the participant's correct decisions (e.g., putting on protective gear when there is danger, and forgo the protective gear when it is safe) by the total number of participant's decisions, across three missions. The value ranges from 0 to 100.
- **Percentage of Decisions that Follow Robot Recommendations** This variable is measured using log data. It is calculated by dividing the total number of the participant's decisions that are the same as the robot's recommended, by the total number of participant's decisions, across three missions. The value ranges from 0 to 100.

6.1 First Iteration of Study

Preliminary analysis of data collected from the first iteration of the study using ANOVA tests revealed no significant impact of the explanation of two sensor readings on any of the four measures, except for understanding of the robot's decisions, compared to when no explanations were offered, regardless of the robot's ability. Closer examination of the scenario design and explanation of two sensor readings suggests that this could be due to the "usefulness" or "helpfulness" of the explanations. The robot, particularly the low-ability robot, makes only false-negative mistakes due to its faulty camera. This means that even when the participants know the robot's decision may not be correct because of the camera failure, they still do not know what the correct decision to make is, e.g., whether to put protective gear on or not. The sensible yet conservative decision would be putting on protective gear all the time. Thus, we added two additional explanation levels— explanation of three sensor readings and confidence-level explanations— that aimed to help participants diagnose faulty sensors and make decisions on whether to put on protective gear. In these two explanations, we also removed the recommendation to put on protective gear, because it is redundant (e.g., the robot's finding of danger implicitly suggests that one should put protective gear on) and removing it reduces the length of text.

6.2 Main Effect of Ability and Explanations

The subsequent analyses include data from both iterations of the study. Overall, ANOVA tests indicate that participants who worked with a high-ability robot reported trusting the robot more, made better decisions and succeeded in more

Table 1: Comparing the main effect of the robot's ability. Means are shown in the table. Differences on all variables between High and Low ability robot are statistically significant ($p < .05$)

	High Ability	Low Ability
Trust in Robot's Ability	6.46	5.05
Understand Robot's Recommendations	5.98	4.70
Mission Success %	82.8	67.6
% of Correct Decisions	92.6	80.0

missions (Table 1). Surprisingly, participants also felt that they understood the robot's decision and decision-making process more, when the robot's ability is high.

As for the main effect of the explanations offered by the robot, Table 2 shows that not all explanations are created equal (Tukey HSD tests on all possible pairwise contrasts). Overall, explanations that facilitate decision-making (e.g., confidence-level explanations, and explanations of 3 sensors) helped the participants succeed in more missions and made the participants feel that they can trust the robot's ability more. Surprisingly, we did not find any significant impact of explanations on the percentage of correct decisions.

There is a significant interaction between the robot's ability and the explanation offered on trust in the robot's ability ($p < .0001$), and understanding of the robot's decisions and decision-making process ($p < .0001$), mission success rate ($p = .0008$), and percentage of correct decisions ($p < .0001$). We will break down the comparison of the impact of robot's explanation between high and low ability robot in the following sections.

6.3 Explanation and Low-Ability Robot

When the low-ability robot makes a decision (e.g., recommendation) that has unwanted consequences, it not only affects the team performance but also jeopardizes the trust its teammate has in it. When no explanations were offered, its human teammate had no additional information to help him/her understand why the robot's recommendation failed. The goal of the explanations is not to help human teammates trust the low-reliability robot more, but to instead calibrate their trust level appropriately and know when and when not to trust it. As a result, the teammate's decision-making and the team performance can be improved. Results from ANOVA and Tukey's HSD tests (on all possible pairwise contrasts) are presented in Table 3. From the table, we can see that the decision-facilitating explanations (e.g., explanations of three sensor and confidence-level explanations), help the teammate understand the low-ability robot's decision and decision-making process, make better decisions, and succeed in more missions, compared to when no explanations were offered or when the explanation is not helpful towards decision making (e.g., explanations of two sensors). The human teammates also trusted the low-ability robot more when it offered the decision-facilitating explanations.

6.4 Explanations and High-Ability Robot

It may seem counter-intuitive that one would not trust a perfectly reliable robot that makes correct decisions 100% of the time. However, disuse is a realistic and common problem in human-automation interaction [27] and often linked to lack of transparency [5]. So we hypothesize that explanations, even offered by a reliable robot, can help improve the trust relationship and team performance. ANOVA and Tukey's HSD tests revealed that there was no statistically significant impact of the robot's explanations on trust in the robot's ability, understanding of the robot's decision and decision-making process, correct decisions made and mission success rate, when the robot is making correct recommendations 100% of the time.

7. DISCUSSION

In this paper, we discussed the design of POMDP-based algorithms for explaining a robot's decision making to a human teammate. We implemented an online experiment platform that we used to conduct an evaluation of the explanation algorithms, where participants teamed up with a simulated robot with either high or low ability, and offered three different types of explanations or no explanations with its decisions. Results indicate that the robot explanations can potentially improve task performance, build transparency and foster trust relationship. However, only explanations that are designed to facilitate decision-making made much difference. Explanations that left participants ambiguous about how to act on the recommendation and explanations did not achieve such an effect, and were as badly regarded as when no explanations were offered at all. This is particularly true so when the robot's ability is low and makes unreliable recommendations.

Additionally, the decision-facilitation explanation helped improve understanding of the robot's decision, but only in the low-ability robot and not the high-ability one. This could be due to the fact that the high-ability robot is making correct decisions 100% of the time. Participants who interacted with this robot never needed to question the robot's decision. Thus, these participants may have never carefully examined the robot's statement that explained its confidence level or observations. Working with a low-ability robot, on the other hand, requires the teammates to pay close attention to the explanations to gauge when and when not to trust the robot's decisions.

Earlier studies on the impact of hand-crafted explanations on trust [7] show that explanations, even those were provided before the interaction and used in ways similar to "excuses", can draw someone into the pitfall of trusting the robot more, even though the robot is unreliable. The result presented here, particularly the finding on decision-facilitating explanation offered by the low-ability robot and subjective trust, sheds some light on the hidden factors between explanations and trust— the "helpfulness" or "usefulness" of the explanation. Our results show that explanations made participants trust the robot's ability more, but only when the explanations facilitated decision-making and not when the explanations left participants unsure about what decisions to make. Participants distrusted a robot that offered such explanations as much as one that did not offer explanations at all.

Interestingly, we did not find any significant differences on the measures we analyzed between the two decision-facilitating explanations e.g., confidence-level explanations and explanations of three sensors. Both types of explanations are useful in helping the human teammate decide when to trust the robot. For example, a teammate could potentially learn

Table 2: Compare the main effect of explanations offered by the robot. Means are shown in the table. A pair of $*$ or † or ♣ or ♠ means the difference between the two variables are statistically significant ($p < .05$).

	NoExp	Exp2Sensor	Exp3Sensor	ExpConf
Trust in Robot's Ability	$5.37^{*\dagger}$	$5.44^{\clubsuit\spadesuit}$	$6.27^{*\clubsuit}$	$6.17^{\dagger\spadesuit}$
Understand Robot's Recommendation	$4.75^{*\dagger}$	5.36	5.93^{*}	$5.51^{\dagger}$
Mission Success %	$65.0^{*\dagger}$	$58.9^{\clubsuit\spadesuit}$	$92.3^{*\clubsuit}$	$96.0^{\dagger\spadesuit}$
% of Correct Decisions	82.8	84.1	89.6	90.0
% of Decisions Follow Robot's Recommendation	83.9	85.6	86.6	86.0

Table 3: Compare the main effect of explanations offered by the *Low Ability* robot. Means are shown in the table. A pair of $*$ or † or ♣ or ♠ means the difference between the two variables are statistically significant ($p < .05$).

	NoExp	Exp2Sensor	Exp3Sensor	ExpConf
Trust in Robot's Ability	$4.31^{*\dagger}$	$4.28^{\clubsuit\spadesuit}$	$6.07^{*\clubsuit}$	$6.15^{\dagger\spadesuit}$
Understand Robot's Recommendation	$3.66^{*\dagger}$	$4.46^{\clubsuit\spadesuit}$	$5.71^{*\clubsuit}$	$5.51^{\dagger\spadesuit}$
Mission Success %	$52.2^{*\dagger}$	$43.0^{\clubsuit\spadesuit}$	$93.7^{*\clubsuit}$	$97.0^{\dagger\spadesuit}$
% of Correct Decisions	$71.9^{*\dagger}$	$74.0^{\clubsuit\spadesuit}$	$87.0^{*\clubsuit}$	$91.9^{\dagger\spadesuit}$
% of Decisions Follow Robot's Recommendation	74.2	77.0	81.5	84.6

his/her own heuristics that if the robot's confidence level is below (for example) 75%, then do not follow the robot's decision. Similarly, a teammate could diagnose from the observation explanations that if the camera reports no signs of danger, but the robot's microphone picks up unfriendly conversations, then it is time to be cautious and put protective gear on, regardless of the robot's overall assessment of safety. It is concerning that participants who received confidence-level explanations also felt that they understood the robot's decision-making process, even though such explanations did not reveal any of the robot's inner workings. While confidence-level explanations may help teammates make decisions just as well as with observation explanations, they will not help teammates diagnose or repair the robot (e.g., the participants will not know that it is the camera that caused the robot to make wrong decisions).

One of the limitations of the current work is that the understanding of the robot's decisions is measured via self-report. In other words, it is unclear whether the participants actually understood the decisions, as they claimed. Future work can include measures to test participants' knowledge of the robot, e.g., its capability, or allow it to be inferred more directly and specifically from the subsequent decisions that participants made, e.g., ask participants to choose MOPP gear vs. body armor. Another limitation of the current work is that the measures are aggregated from participants' responses after each of the 3 missions. More fine-grained analysis of data collected from each mission can be conducted to study how trust evolves over time. These future analyses can lead to further refinements of our explanation algorithms that can increase the positive impact already exhibited by the current implementation on human-robot trust.

Acknowledgment

This project is funded by the U.S. Army Research Laboratory. Statements and opinions expressed do not necessarily reflect the position or the policy of the United States Government, and no official endorsement should be inferred.

REFERENCES

[1] B. Adams, L. Bruyn, S. Houde, and P. Angelopoulos. Trust in automated systems: literature review. Technical Report DRDC-TORONTO-CR-2003-096, Defence Research Reports, 2003.

[2] W. Bluethmann, R. Ambrose, M. Diftler, S. Askew, E. Huber, M. Goza, F. Rehnmark, C. Lovchik, and D. Magruder. Robonaut: A robot designed to work with humans in space. *Autonomous Robots*, 14(2-3):179–197, 2003.

[3] J. L. Burke, R. R. Murphy, M. D. Coovert, and D. L. Riddle. Moonlight in Miami: Field study of human-robot interaction in the context of an urban search and rescue disaster response training exercise. *Human-Computer Interaction*, 19(1-2):85–116, 2004.

[4] A. R. Cassandra, L. P. Kaelbling, and J. A. Kurien. Acting under uncertainty: Discrete Bayesian models for mobile-robot navigation. In *IROS*, volume 2, pages 963–972, 1996.

[5] H. Cramer, V. Evers, S. Ramlal, M. Van Someren, L. Rutledge, N. Stash, L. Aroyo, and B. Wielinga. The effects of transparency on trust in and acceptance of a content-based art recommender. *User Modeling and User-Adapted Interaction*, 18(5):455–496, 2008.

[6] P. Doshi and D. Perez. Generalized point based value iteration for interactive pomdps. In *AAAI*, pages 63–68, 2008.

[7] M. T. Dzindolet, S. A. Peterson, R. A. Pomranky, L. G. Pierce, and H. P. Beck. The role of trust in automation reliance. *International Journal of Human-Computer Studies*, 58(6):697–718, 2003.

[8] F. Elizalde, L. E. Sucar, M. Luque, J. Diez, and A. Reyes. Policy explanation in factored markov decision processes. In *Proceedings of the European WS on Probabilistic Graphical Models*, pages 97–104, 2008.

[9] P. J. Gmytrasiewicz and E. H. Durfee. A rigorous, operational formalization of recursive modeling. In *ICMAS*, pages 125–132, 1995.

[10] V. Greco and D. Roger. Coping with uncertainty: The construction and validation of a new measure. *Personality and individual differences*, 31(4):519–534, 2001.

[11] S. G. Hart and L. E. Staveland. Development of NASA-TLX (task load index): Results of empirical and theoretical research. *Advances in Psychology*, 52:139–183, 1988.

[12] L. Hendrickx, C. Vlek, and H. Oppewal. Relative importance of scenario information and frequency information in the judgment of risk. *Acta Psychologica*, 72(1):41–63, 1989.

[13] W. L. Johnson and A. Valente. Tactical language and culture training systems: Using AI to teach foreign languages and cultures. *AI Magazine*, 30(2), 2009.

[14] L. P. Kaelbling, M. L. Littman, and A. R. Cassandra. Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, 101:99–134, 1998.

[15] O. Z. Khan, P. Poupart, and J. P. Black. Minimal sufficient explanations for factored Markov Decision Processes. In *ICAPS*. Citeseer, 2009.

[16] J. M. Kim, J. Randall W. Hill, P. J. Durlach, H. C. Lane, E. Forbell, M. Core, S. Marsella, D. Pynadath, and J. Hart. BiLAT: A game-based environment for practicing negotiation in a cultural context. *IJAIED: Special Issue on Ill-Defined Domains*, 19(3):289–308, 2009.

[17] J. Klatt, S. Marsella, and N. Krämer. Negotiations in the context of AIDS prevention: An agent-based model using theory of mind. In *IVA*, 2011.

[18] S. Koenig and R. Simmons. Xavier: A robot navigation architecture based on partially observable Markov decision process models. In D. Kortenkamp, R. P. Bonasso, and R. R. Murphy, editors, *AI Based Mobile Robotics: Case Studies of Successful Robot Systems*, pages 91–122. MIT Press, 1998.

[19] J. Lee and N. Moray. Trust, control strategies and allocation of function in human-machine systems. *Ergonomics*, 35(10):1243–1270, 1992.

[20] J. D. Lee and K. A. See. Trust in automation: Designing for appropriate reliance. *Human Factors*, 46(1):50–80, 2004.

[21] S. C. Marsella, D. V. Pynadath, and S. J. Read. PsychSim: Agent-based modeling of social interactions and influence. In *ICCM*, pages 243–248, 2004.

[22] R. C. Mayer, J. H. Davis, and F. D. Schoorman. An integrative model of organizational trust. *Academy of Management Review*, 20(3):709–734, 1995.

[23] R. McAlinden, A. Gordon, H. C. Lane, and D. Pynadath. UrbanSim: A game-based simulation for counterinsurgency and stability-focused operations. In *Proceedings of the AIED WS on Intelligent Educational Games*, 2009.

[24] D. H. McKnight, V. Choudhury, and C. Kacmar. Developing and validating trust measures for e-commerce: An integrative typology. *Information Systems Research*, 13(3):334–359, 2002.

[25] S. L. McShane. Propensity to trust scale, 2014. http://highered.mheducation.com/sites/0073381225/student_view0/chapter7/self-assessment_7_4.html.

[26] L. C. Miller, S. Marsella, T. Dey, P. R. Appleby, J. L. Christensen, J. Klatt, and S. J. Read. Socially optimized learning in virtual environments (SOLVE). In *ICIDS*, 2011.

[27] R. Parasuraman and V. Riley. Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, 39(2):230–253, 1997.

[28] J. Pineau, M. Montemerlo, M. Pollack, N. Roy, and S. Thrun. Towards robotic assistants in nursing homes: Challenges and results. *Robotics and Autonomous Systems*, 42(3):271–281, 2003.

[29] D. V. Pynadath and S. C. Marsella. PsychSim: Modeling theory of mind with decision-theoretic agents. In *IJCAI*, pages 1181–1186, 2005.

[30] J. M. Ross. *Moderators of trust and reliance across multiple decision aids*. PhD thesis, University of Central Florida, 2008.

[31] S. Ross, J. Pineau, S. Paquet, and B. Chaib-Draa. Online planning algorithms for POMDPs. *JAIR*, 32:663–704, 2008.

[32] K. E. Schaefer. *The perception and measurement of human-robot trust*. PhD thesis, University of Central Florida Orlando, Florida, 2013.

[33] W. R. Swartout and J. D. Moore. Explanation in second generation expert systems. In *Second generation expert systems*, pages 543–585. Springer, 1993.

[34] D. S. Syrdal, K. Dautenhahn, K. L. Koay, and M. L. Walters. The negative attitudes towards robots scale and reactions to robot behaviour in a live human-robot interaction study. *Adaptive and Emergent Behaviour and Complex Systems*, 2009.

[35] R. Taylor. Situational awareness rating technique (SART): The development of a tool for aircrew systems design. *AGARD, Situational Awareness in Aerospace Operations*, 1990.

[36] V. H. Visschers, R. M. Meertens, W. W. Passchier, and N. N. De Vries. Probability information in risk communication: a review of the research literature. *Risk Analysis*, 29(2):267–287, 2009.

[37] N. Wang and D. V. Pynadath. Building trust in a human-robot team. In *I/ITSEC*, 2015.

[38] N. Wang, D. V. Pynadath, K. Unnikrishnan, S. Shankar, and C. Merchant. Intelligent agents for virtual simulation of human-robot interaction. In *Virtual, Augmented and Mixed Reality*, pages 228–239. Springer, 2015.

[39] E. A. Waters, N. D. Weinstein, G. A. Colditz, and K. Emmons. Formats for improving risk communication in medical tradeoff decisions. *Journal of health communication*, 11(2):167–182, 2006.

[40] L. R. Ye and P. E. Johnson. The impact of explanation facilities on user acceptance of expert systems advice. *MIS Quarterly*, 19(2):157–172, 1995.