Descriptive and Prescriptive Visual Guidance to Improve Shared Situational Awareness in Human-Robot Teaming

Aaquib Tabrez*
University of Colorado Boulder
Boulder, Colorado, USA
mohd.tabrez@colorado.edu

Matthew B. Luebbers*
University of Colorado Boulder
Boulder, Colorado, USA
matthew.luebbers@colorado.edu

Bradley Hayes
University of Colorado Boulder
Boulder, Colorado, USA
bradley.hayes@colorado.edu

ABSTRACT

In collaborative tasks involving human and robotic teammates, live communication between agents has potential to substantially improve task efficiency and fluency. Effective communication provides essential situational awareness to adapt successfully during uncertain situations and encourage informed decision-making. In contrast, poor communication can lead to incongruous mental models resulting in mistrust and failures. In this work, we first introduce characterizations of and generative algorithms for two complementary modalities of visual guidance: prescriptive guidance (visualizing recommended actions), and descriptive guidance (visualizing state space information to aid in decision-making). Robots can communicate this guidance to human teammates via augmented reality (AR) interfaces, facilitating synchronization of notions of environmental uncertainty and offering more collaborative and interpretable recommendations. We also introduce a min-entropy multi-agent collaborative planning algorithm for uncertain environments, informing the generation of these proactive visual recommendations for more informed human decision-making. We illustrate the effectiveness of our algorithm and compare these different modalities of AR-based guidance in a human subjects study involving a collaborative, partially observable search task. Finally, we synthesize our findings into actionable insights informing the use of prescriptive and descriptive visual guidance.

KEYWORDS

Shared Mental Models; Human-Robot Collaboration; Explainable AI; Augmented Reality; Reinforcement Learning

ACM Reference Format:
Aaquib Tabrez, Matthew B. Luebbers, and Bradley Hayes. 2022. Descriptive and Prescriptive Visual Guidance to Improve Shared Situational Awareness in Human-Robot Teaming. In *Proc. of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022), Online, May 9–13, 2022*, IFAAMAS, 9 pages.

1 INTRODUCTION

When a team is tasked with solving a problem in an uncertain environment, it is vitally important to keep notions of that uncertainty, as well as the problem-solving strategy, synchronized between teammates as this information changes over time, in order for each teammate to act in a coordinated fashion. In this work, we explore this challenge as it relates to human-robot teaming. Autonomous agents are well-equipped to plan over probabilistic state spaces, updating their probability models in response to new observations, and choosing optimal actions in response to this new information. We hypothesize that visually communicating this knowledge to human teammates efficiently improves team performance.

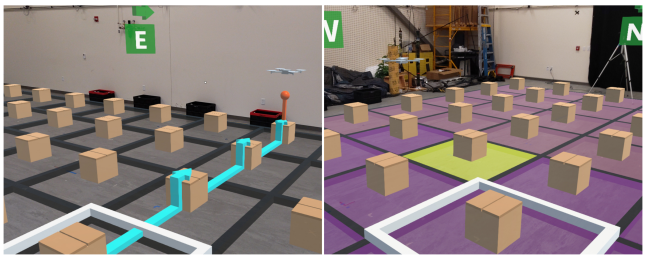

Figure 1: AR-based interfaces for prescriptive (Left) and descriptive guidance (Right) in the Minesweeper domain. In the prescriptive condition, suggested moves are shown as cyan arrows between grid squares, with suggested defuse actions indicated by the orange pin (underneath the virtual drone teammate). In the descriptive condition, grid squares are colored as a heatmap, representing the probability for each square containing a hidden mine as judged by the drone, from dark purple (low) to bright yellow (high).

Consider a search and rescue task with human and robot teammates coordinating to locate a victim: this is an inherently stochastic environment, where the likelihood of finding a victim varies location to location, characterized by a probability mass function (PMF). As the human and robot teammates cover more ground with their search, that PMF continually updates in response to the agents' observations. Since the robot agents are maintaining an up-to-date PMF to plan over, they can also communicate it to their human counterpart to keep them in the loop, a modality we call *descriptive guidance* (synchronizing state space information to aid in human decision making). Additionally, the robots can use that PMF combined with a model of their human counterpart's action space to directly recommend next actions to the human, a modality we call *prescriptive guidance.*

In this work we use a 3D collaborative analogue of the PC game Minesweeper, played using an augmented reality (AR) headset, to serve as an experimental domain reminiscent of real-world spatial navigation and search tasks. For this game, we tasked a human-drone team with locating and defusing a number of mines hidden throughout a grid of cardboard boxes projected onto the floor of an experiment space (Fig. 1). The drone can navigate the environment, taking measurements with a noisy sensor to attempt to determine

*These authors contributed equally to this work.

Proc. of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022), P. Faliszewski, V. Mascardi, C. Pelachaud, M.E. Taylor (eds.), May 9–13, 2022, Online.

whether a box contains a hidden mine. The human must also physically navigate the environment, taking time to search boxes and defuse mines whenever they think they've located one.

To assist the human in their task, we developed an algorithmic framework for multi-agent collaboration under uncertainty, capable of generating prescriptive and descriptive visual guidance for the human teammate as the drone explores the environment. We also developed AR interfaces for each type of drone-provided guidance, with arrows and pins indicating suggested moves under the prescriptive modality, and a heatmap overlaid onto the environment representing the PMF under the descriptive modality (Fig. 1).

We conducted a human subjects study using this collaborative Minesweeper task, varying which modality of guidance participants saw between conditions as they attempted to locate and defuse all hidden mines as quickly as possible: prescriptive guidance (the 'arrow' condition), descriptive guidance (the 'heatmap' condition), and a combination of both (the 'combined' condition). This study served to validate our algorithm in a live human-robot teaming setting with environmental uncertainty, helping to assess the benefits and drawbacks of each type of visual guidance through a variety of objective and subjective measures.

We characterize the core contributions of this work as follows:

- A *characterization of* and *method for* generating AR-based prescriptive and descriptive visual guidance, communicating environmental uncertainty and providing actionable recommendations to human teammates in joint human-robot tasks.
- An empirical validation and analysis of the effectiveness of prescriptive and descriptive visual guidance through a human subjects study involving a collaborative search task with an autonomous robot.

2 BACKGROUND AND RELATED WORK

Visual Guidance & Augmented Reality Interfaces. Visualization is frequently used in human-robot teaming for tasks such as environmental navigation, search and inspection, and fault recovery [6, 20, 22]. The visualization of task and environment data enables human teammates to develop new insights into the problem being solved and heightens their situational awareness, aiding in decision-making [36]. Gale et al. demonstrated the effectiveness of playbook-based visual interfaces to allocate roles and responsibilities between human-automation systems in an unmanned aircraft system (UAS) swarm support task [15]. Ahmed et al. successfully utilized a visual sketching interface to fuse the data of multiple noisy 'human sensors' in cooperative search missions with autonomous vehicles, further demonstrating the utility of visual information transfer in human-robot teaming [1].

Visualization is particularly useful for communicating uncertainty. Bhatt et al. explored methods for assessing and displaying uncertainty in models, communicating it to stakeholders to assist in trust-building and decision making. [2]. Furthermore, Colley et al. showed that visualizing the internal information of autonomous vehicles improves trust and situational awareness [7]. As these works focus on the communication of internal model-based uncertainty in human-robot teaming, we apply the same concept to external environment-based uncertainty associated with unexplored terrain.

Recent work on augmented reality-based interfaces has shown that providing in-situ visualizations with an AR headset can greatly improve the efficiency of human-robot teaming [30, 31]. Fraune et al. investigated the use of mixed reality interfaces for humans monitoring and commanding drone teams for search and rescue [13]. Kunze et al. show the effectiveness of AR to visually communicate uncertainty during automated driving [23].

Explainable AI & Shared Mental Models. Recent research in model reconciliation and knowledge sharing in human-robot teams has shown the importance of explainability and mental model synchronization to improve trust, transparency, and team performance [5, 38]. Furthermore, explainable AI (xAI) can help complex models become more understandable by human teammates, allowing for faster debugging when unexpected behaviors or failures occur [18, 32]. Visualization is a common modality for presenting explanations through xAI [33]. Visual information presentation is ideally suited to explanations that are complex, long, re-referenced, and which involve uncertainty or noise [10]. Therefore, visualization is often used to aid in the interpretation of complex models, showing how model parameters affect final classification decisions (e.g., in local approximation methods such as SHAP [26], model-agnostic methods such as LIME [32], and saliency map methods such as Grad-CAM [34]).

Other recent studies have utilized case-based explanations as visualizations to expose overconfidence in models and visualize class boundaries [3]. A related technique is visual counterfactuals [25, 27] (showing how an input must change to change the classification of the output). These techniques are typically utilized post-hoc by AI experts to debug models [28, 29]. Our visual guidance methodology on the other hand assumes very little domain knowledge, leverages an AR-based interface for more user friendly visualization, and is usable in live human-robot teaming scenarios.

3 ALGORITHMIC APPROACH

In this section, we introduce a novel algorithm for multi-agent collaboration under uncertainty using min-entropy online reinforcement learning called MARS (Min-entropy Algorithm for Robot-supplied Suggestions).

Our algorithm assumes the presence of two classes of agents: exploration agents (agents who can move through the environment and take observations) and active agents (agents who can directly affect environment state through taking actions). This divide between agents with differing goals and action spaces is typical in human-robot teaming domains. For example, a common search and rescue practice involves an initial search phase conducted by an aerial vehicle, with ground rescue or airlift units deployed to extract targets once their locations are determined. In this work, we explore the case where the active agent is human and the exploration agents are autonomous.

3.1 Multi-Agent Entropy Minimization

The core insight behind this algorithm is that environmental uncertainty over task-relevant variables can be succinctly characterized by probability density distributions, a common practice in search and rescue operations [14, 41, 42]. We use the multivariate probability mass function (PMF), a discrete version of this concept, to

model environmental uncertainty as it changes over time. This PMF serves as a shared utility function between all agents in our formulation for min-entropy collaborative planning, allowing for solving a single Markov Decision Process (MDP) with the PMF as its utility function. Furthermore, this PMF can be communicated to human teammates in order to provide insight into the autonomous agents' policy which we detail in Section 4.

The collaborative task can be formulated as a single MDP M_R, over which one or multiple exploration agents maximize their expected reward. M_R is defined by the 4-tuple: (S, A, T, R):

- S is the finite set of discrete world states consisting of traditional "world features" W (e.g., agent positions) along with "distance features" D that encode pairwise distances between all agents in the collaborative task (including the human teammate), using an appropriate distance metric for the task being solved. A finite set of distance features is given by $D = \{d_{12}, d_{13}, \cdots, d_{(N-1)N}\}$, such that d_{12} represents the distance between agent 1 and 2, and so on. $|D| = \binom{N}{2}$, where N is the total number of agents in the collaborative task.

$$S = \begin{bmatrix} W \\ D \end{bmatrix}, W = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \end{bmatrix} D = \begin{bmatrix} d_{12} \\ d_{13} \\ \vdots \\ d_{(N-1)N} \end{bmatrix}$$

- A is the set containing all N-tuples representing the product of all possible exploration agent joint actions.
- $T : S \times A \rightarrow \Pi(S)$ is the state-transition function describing the model's state transition dynamics.
- $R : S \times A \times S \rightarrow \mathbb{R}$ defines the expected immediate reward gained by the agent for taking an action $a \in A$ in a state $s \in S$ and transitioning into the next state $s' \in S$.

We solve this single MDP M_R via online reinforcement learning to get an optimal policy π_R^* for each autonomous agent using a joint PMF as a reward function given by:

$$R(s, a, s') = \alpha(0.5 - |0.5 - pmf(s')|) + \beta \sum_{n \in N} d_n - 1 \tag{1}$$

In Equation 1, α and β are tunable hyper-parameters, and $pmf(s')$ is the value of the probability mass function at state s', representing the probability that s' contains a desired goal or target. The first term of Equation 1 encourages the exploration of states with higher uncertainty (PMF values close to 0.5), minimizing entropy over time as those states are observed. The second term maximizes distance from other agents, maximizing coverage over the state space for faster learning. Each agent's reward function is affected by the current PMF, which is updated every time agents observe a new state in the environment according to Bayes' rule. Therefore, the MDP should be re-solved whenever the PMF updates, in order to minimize the entropy of the distribution over task-relevant latent state information.

3.2 Generating Assistive Guidance

Here we present our approach for generating assistive guidance for human teammates in uncertain environments. Similarly to section 3.1, we can model a human agent's behavior using an MDP with the PMF as its utility function. The MDP M_H is likewise defined by a 4-tuple (S, A, T, R), where:

- S is the finite set of world states consisting of traditional "world features" W, along with the expected number of goals left ("*goals_left*"), and the latent boolean variable *is_goal* $\in \{0, 1\}$ with *is_goal* $= 1$ indicating a goal is present.

$$S = \begin{bmatrix} W \\ goals\_left \in \mathbb{N} \\ is\_goal \in \mathbb{B} \end{bmatrix}$$

- A is the set of possible task-relevant human actions.
- T and R are similarly defined as seen in Section 3.1.

Our reward function distinguishes between two classes of actions: exploration and goal-centric actions. Exploration actions are geared towards navigating between states to minimize uncertainty or reach a state containing a goal. In comparison, goal-centric actions are conducted within a state and contribute towards task completion (e.g., signaling for pickup in SAR domains).

The reward function for a human agent exploration action is given by:

$$R(s, a, s') = pmf(s') - \beta * is\_goal_s - penalty \tag{2}$$

where,

$$penalty = 1 - \alpha * goals\_left$$

The first term of Equation 2 provides the immediate reward from the next state s', the second term encodes a negative reward for ignoring a goal in the current state s, and the penalty term provides long term incentive to achieve the desired task objectives as quickly as possible. α and β are tunable hyper-parameters. We can expand Equation 2 to get the expected immediate reward as follows:

$$\begin{aligned} \mathbb{E}(R) = (1 - pmf(s)) * (pmf(s') - penalty) \ + \\ pmf(s) * (pmf(s') - penalty - \beta) \end{aligned} \tag{3}$$

The reward function for a human agent to take goal-centric actions is as follows:

$$R(s, a, s') = \beta * is\_goal_s - penalty \tag{4}$$

The first term of equation 4 provides the immediate reward if a goal is present in the current state s, and the rest of the terms are defined the same as in Equation 2. Expanding Equation 4, the expected immediate reward is:

$$\begin{aligned} \mathbb{E}(R) = pmf(s) * (\beta - penalty) \ - \\ (1 - pmf(s)) * penalty \end{aligned} \tag{5}$$

Solutions to this MDP M_H can be used to obtain policy recommendations for a human agent.

3.3 Algorithm

In this section, we outline the details of MARS, as presented in Algorithm 1. We ground the algorithm with an example task inspired by Minesweeper, involving a single human agent and a single robotic drone. The goal of the task is to locate and defuse a number of mines hidden throughout a grid-based environment without unintentionally detonating them. Although only the human teammate is capable of defusing mines, the drone has a noisy sensor capable

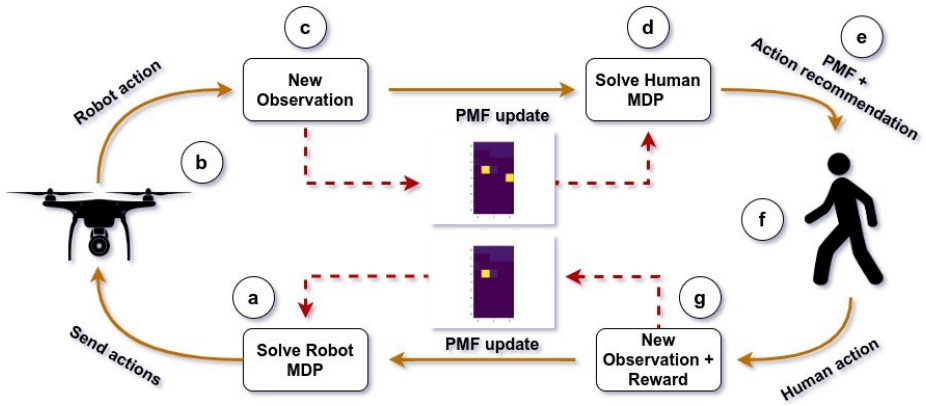

Figure 2: Algorithmic flow: a) the robot's MDP is solved, parametrized by the PMF, and actions are sent to all agents, b) the robot takes an action and c) observes a new potential mine, updating the PMF (the new mine is visible as the righmost yellow square), d) the updated PMF is used to solve the human recommendation MDP, e) the resulting PMF and action recommendations are sent to the human, who f) views the guidance via an AR interface, and takes an action, defusing the mine, g) the new observation and reward update the PMF again (the new mine has been defused, removing the yellow from the heatmap)

Algorithm 1: Min-entropy Algorithm for Robot-supplied Suggestions (MARS)

Input: Robots' MDP M_R (S, A, T, R), Human's MDP M_H (S, A, T, R), R_h, Current Robots State $\overline{S_R} = \{s_1, s_2, \cdots, s_{n-1}\}$, Current Human State s_h, Num. rollout k, Prior P

1 $pmf \leftarrow P$; // Initialize pmf with prior
2 **while** s_h *is not a terminal state* **do**
3 $\quad \pi_R^* \leftarrow solve\_policy(M_R, pmf)$;
4 $\quad \overline{A_R} \leftarrow \pi_R^*(\overline{S_R})$; // Get optimal actions for each robot
5 $\quad \overline{S_R} \leftarrow send\_actions(\overline{A_R})$; // Send optimal actions
6 $\quad pmf \leftarrow update\_pmf(\overline{S_R})$; // Get observations
7 $\quad \pi_H^* \leftarrow solve\_policy\_human(M_H, pmf)$;
8 $\quad \overline{A_H} \leftarrow rollout(\pi_H^*, s_h)[:k]$; // Get actions for human
9 $\quad recommend\_action(\overline{A_H}, pmf)$
10 $\quad s_h, r_h \leftarrow observe\_human\_action()$
11 $\quad pmf \leftarrow update\_pmf(r_h)$

of determining whether the grid square it is currently flying over contains a hidden mine, parameterized by a false positive and false negative rate. If the human teammate leaves a square containing a mine without defusing it, it detonates, providing a substantially negative (non-terminal) reward for the episode.

Before the task begins, the PMF is initialized with a prior to provide an initial heuristic (Line 1). If there is no information with which to seed a prior, a uniform PMF can be used at this step. An optimal policy can then be computed using the prior PMF and the robots' MDP M_R. Based on the learned policy π_R^*, optimal actions are sent to all robots (Lines 3-5). Once the robots execute these actions, they obtain new observations from the environment and update the PMF using Bayes' Rule (Line 6). In the Minesweeper example shown in Figure 2, step c shows the resultant PMF after the robot takes an action and obtains a new observation.

Given this updated PMF, the human agent's policy π_H^* is computed and a k-step rollout is used to provide action suggestions for the human (Line 7-8). The number of steps k determines how many actions into the future will be recommended to the human teammate, which can be chosen depending on the nature of the task. For the Minesweeper example, we provided suggested actions up to and including the first recommended "defuse" action (step e in Figure 2). These actions $\overline{A_H}$ and the updated PMF are provided to the human agent as guidance (Line 9), the visualization of which is discussed in Section 4. Next, the human action is observed, the reward r_h is recovered from the environment, and the PMF is updated again in response (Lines 10-11).

4 AR-BASED VISUAL GUIDANCE DESIGN

The PMF and action recommendations meant to be communicated to the human agent are particularly well-suited for visual presentation in the Minesweeper domain, but this will vary by task. For the Minesweeper domain, we developed a set of AR visualizations geared toward environment navigation and search tasks. An AR headset-based interface was chosen due to its hands-free nature and its ability to present information in-situ, as holograms projected in environmental context aid in the efficiency of information uptake.

We generalize the proposed AR-based visual guidance into two categories, corresponding to the two data products of Algorithm 1. First is prescriptive guidance, in which sequences of actions are directly suggested to the human based on the algorithm's current recommendations. Second is descriptive guidance, where state space information is presented to the human in the form of the current PMF to support decision making.

4.1 Prescriptive Guidance

The essence of prescriptive guidance is directly suggesting to a human teammate what they should do next. In tasks involving physically navigating through space, like search and rescue or the Minesweeper experimental domain, movement suggestions can be represented as holographic arrows projected onto the ground, extending from the human's current location to their next suggested waypoint (Fig. 1 Left), an AR visualization technique which has shown effectiveness for navigation tasks [16].

This arrow-based guidance is straightforward to understand and requires little mental effort to follow. However, since the recommendations are presented without rationale, they require a degree of trust from the human teammate if they are to be followed, which may or may not be warranted depending on the performance of the autonomous agents under environmental uncertainty. This uncertainty may also lead to frequent changes in the path recommendations, deflecting the arrows and causing confusion on the part of the human teammate as the old guidance is discarded.

4.2 Descriptive Guidance

In contrast to explicit action recommendations, descriptive guidance involves providing state space information with which human teammates can make their own decisions. For spatial navigation tasks like the Minesweeper domain, the current PMF can be projected onto the environment itself, dividing the space into discrete regions and coloring those regions as a heatmap (Fig. 1 Right). In the Minesweeper domain, dark purple is used to represent a low chance of a region containing a goal while bright yellow is used to represent a high chance, with intermediate probabilities colored on a gradient between purple and yellow. Since decision-making in the Minesweeper domain relies more on discrimination between PMF probabilities close to 0 than probabilities close to 1, the heatmap is generated using a logarithmic color scale, a technique used to visually bring out finer distinctions towards the low end of a scale with an uneven distribution [12].

This descriptive guidance acts as a decision support tool, providing the human with information which they can use however they see fit. In contrast to the prescriptive arrows, this type of guidance is highly transparent. On the other hand, it is more cognitively demanding, requiring the human to actively plan ahead, thereby reducing its effectiveness in domains with large and complicated state spaces or domains with time pressure.

5 EXPERIMENTAL VALIDATION

We evaluate the utility of the AR-based visual guidance modalities presented in Section 4 within a partially observable environment involving live human-robot teaming, utilizing the proposed multi-agent entropy minimization algorithm. These results were obtained through a human subjects study using our collaborative Minesweeper-inspired domain.

5.1 Experimental Design

We use a 3×1 within-subjects experiment to evaluate three different varieties of AR-based visual guidance: 1) prescriptive guidance, or the 'arrow' condition, 2) descriptive guidance, or 'heatmap', and 3) a combination of prescriptive and descriptive guidance, or 'combined' (Figure 3). A within-subjects design was chosen to obtain direct, grounded comparisons between visualization types from participants. The guidance was visualized through a Microsoft HoloLens 2, overlaid onto a rectangular grid of cardboard boxes on the floor of the experiment space.

The orderings of the 'arrow' and 'heatmap' conditions were randomized and fully counterbalanced between participants. Since the 'combined' condition relied on the prior introduction of both modalities independently, it was ordered last. As participants played three rounds of the game with differing conditions, three environment maps were created, each with the same number of hidden mines, located on different squares. We blocked participants to match experimental conditions to environment maps using a balanced Latin square design to achieve partial counterbalancing and minimize ordering and learning effects [4, 8]. The Latin square resulted in blocks of size six differing in the ordering of the 'arrow' and 'heatmap' conditions, and in the matching of environment map to condition. Participants were randomly assigned to one of these six permutations.

5.2 Hypotheses

Through a human subjects study, we evaluate five visual guidance hypotheses partitioned into three categories:

H1: Subjective Hypotheses

H1.a: Participants will find the combined guidance to be more trustworthy than descriptive or prescriptive guidance, as transparency of recommendation leads to more trust [35, 37].

H1.b: Participants will find the combined guidance to be more interpretable, informative, and helpful for decision-making compared with the other conditions.

H1.c: Participants will find the combined and prescriptive guidance conditions to be less stressful and demanding compared with descriptive guidance, due to the presence of clear recommendations.

H2: Performance Hypothesis

H2: Participants will take less time to solve the task when given combined or prescriptive guidance compared with descriptive guidance, since they can reduce thinking time by leveraging direct algorithmic guidance.

H3: Independence Hypothesis

H3: Participants will act with more independence and deviate more frequently from the prescribed path in the combined condition compared with solely receiving prescriptive guidance, as they can utilize the added descriptive information to take their own initiative when they perceive suboptimality in robot suggestions.

5.3 Rules of the Game

Each round, participants attempted to solve the Minesweeper puzzle by successfully locating and defusing all four mines hidden throughout the 9×5 grid of cardboard boxes as viewed through the HoloLens headset. Each turn, participants had four options for movement actions: "Go North", "Go South", "Go East", and "Go West", each of which moves a single square in the respective direction. If the participant suspected a square contained a hidden mine, they could take a fifth action: "Defuse", which opened the box on the square they were currently standing on, revealing whether it was empty or contained a mine, which they had now successfully defused (Fig. 3). If they moved from a square containing a mine without defusing, the mine would be unintentionally detonated. Unlike Minesweeper, this did not end the game; participants were simply told beforehand that this would contribute to a low score.

As the participants moved through the grid, a virtual drone teammate concurrently explored the grid autonomously, providing assistive guidance in a format dictated by the experimental condition. After the participant took a turn, they waited briefly for the drone to take theirs. The drone could move faster than the

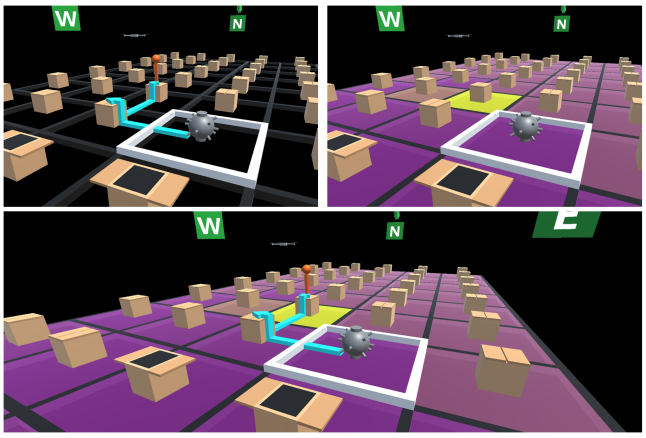

Figure 3: The three experimental conditions. A white square marks the user's current location where they have defused a mine. Top-left: 'arrow' condition, Top-right: 'heatmap' condition, Bottom: 'combined' condition.

human teammate, moving three squares for every human action and using its noisy mine-detection sensor on every square it flew over. However, the drone was incapable of defusing or otherwise interacting with the mines; only the participant could do that. The human and drone teammates alternated turns until all four mines had been successfully defused or unintentionally detonated.

5.4 Study Protocol

Upon providing informed consent, participants were educated on the overall rules of the game through alternating phases of reading an illustrated instruction manual and reviewing it with an experimenter to reinforce the ideas. To minimize potential learning effects, participants were given a brief practice round (without visual guidance) using the HoloLens to ensure that they acclimated to the AR interface and became comfortable exploring the environment and issuing commands, trying every action at least once. Participants were told about their drone teammate, including information about the drone's capabilities and limitations, namely its uncertain sensor. This served to ensure participants would not be overly confused if they saw the drone's guidance change during the experiment round.

Participants began their first experimental round with randomized condition and environment map. They were first shown a page in the instruction manual describing the form of guidance they would be receiving that round. They then donned the HoloLens and played the round, taking actions and navigating the experiment space until all four mines had been defused or unintentionally detonated. After finishing the round, participants removed the HoloLens and returned to the staging area to complete a post-round survey. These steps were repeated twice more for the other experimental conditions. Following the third post-round survey, participants completed a final post-experiment survey and an exit interview.

5.5 Implementation Details

Three environment maps with different locations for the four hidden mines were selected to be of similar difficulty and similar optimal solving time. Each round, the virtual drone's actions were controlled by our algorithm running on a laptop (Intel(R) Core i7-10870H CPU @ 2.20GHz) and broadcasted turn-by-turn via a ROS publisher to the HoloLens. The drone's guidance each round was similarly computed by our algorithm and broadcast to the HoloLens using ROS. Each turn, the drone took three steps to mimic the relative speed of aerial robot navigation over human navigation. The drone observed every square it flew over, even observing some squares more than once, using a simulated noisy sensor with a 10% false-positive rate and a 1% false-negative rate to determine whether a hidden mine is present on that square, adding uncertainty into the drone's recommendations. We chose to use a single drone for our experiment since our domain was small and adding more autonomous agents would lead to quicker convergence towards optimal guidance, causing a more deterministic interaction with participants. The robot's MDP and the human recommendation MDP were solved online each turn using policy iteration.

In the prescriptive 'arrow' condition, our algorithm sent action suggestions every turn up to and including the next suggested "Defuse" action to the AR interface. In the descriptive 'heatmap' condition, our algorithm sent the updated PMF every turn, shown as a heatmap from dark purple for low values to bright yellow for high values, interpolating logarithmically for intermediate values. Each turn, participants selected their action via their choice of voice control (comprising 69.3% of all 1597 recorded moves), or menu-based hand control (30.7% of recorded moves).

In all three environmental maps, there was the possibility for certain scenarios we dub "switchbacks" where participants will turn around and double back on their previous state if they follow the drone's updated prescriptive arrow. These scenarios are an emergent behavior when the participant is located immediately between two potential mine locations, whether they are actual mines or false positives. The drone simply updates its path based on new information and reward maximization, but its behavior is often perceived as suboptimal from the perspective of the human teammate. We observed how participants responded to these switchbacks, especially as they differed based on guidance condition.

5.6 Measurement

We had 19 participants (12 males, 7 females) in our IRB-approved study, ranging in age from 18 to 37 ($M = 25.42; SD = 4.76$). We used a number of subjective and objective measures to evaluate our algorithm and the AR-based visual guidance.

For subjective metrics, we administered post-round questionnaires to participants for each condition to get immediate impressions. These surveys consisted of 7-point Likert-scale items derived from questions from established questionnaires in the robotics and explainable AI community, geared at trust and reliability [19, 24], interpretability and decision-making [19, 40], and stress and workload (NASA-TLX) [17]. From these items, we were able to identify three concepts: *Trust, Interpretability, and Mental Load.*

The *Trust* scale consists of 4 items: confidence, reliability, trust, and intelligence (Cronbach's $\alpha = 0.90$). *Interpretability* consists of 4 items: decision-making power, adaptability, informativeness, and sufficiency (Cronbach's $\alpha = 0.89$). *Mental Load* consists of 2 items: stress and cumbersomeness (Cronbach's $\alpha = 0.84$).

Following the last round of the experiment, participants compared each of the three guidance types they received. Participants ranked each guidance type relative to one another in terms of *trust*, *usefulness*, *helpfulness for decision making*, and *confidence*.

For objective metrics, we recorded the following items for each experiment round: *Total Moves* (the total number of moves needed to solve the puzzle), *Total Time* (the total time needed to solve the puzzle, in seconds), *Time per Move* (the average time per move, in seconds), and *Compliance Rate* (the percentage of moves taken matching the recommendation provided by the system, only applicable for the 'arrow' and 'combined' conditions).

6 RESULTS AND DISCUSSION

6.1 Analysis

6.1.1 Subjective Analysis. We analyzed both the post-round survey scales and post-experiment comparison results to test our subjective hypotheses. The post-round Likert scale data suffered from a significant ceiling effect, where many participants rated all guidance types highly, using primarily 6s and 7s out of a maximum score of 7. For this reason, we transformed the raw Likert scores into rankings, giving for each survey item the participant's preference ordering between the three guidance types, with any ties receiving equal ranks. We analyzed both this ranked scale data and the ranks from the post-experiment survey's comparison questions using a nonparametric Kruskal-Wallis Test with experimental condition as a fixed effect. Post-hoc comparisons used Dunn's Test for analyzing guidance type sample pairs for stochastic dominance.

We found a significant effect in favor of the 'combined' condition over 'arrow' for the *Trust* scale ($H(2) = 8.26, p = 0.016$). Post-hoc analysis with Dunn's Test found that participants consistently preferred 'combined' ($M = 2.68$), $p = 0.017$ over 'arrow' ($M = 2.03$). We also found significant effects in the related post-experiment comparison measures of *trust* ($H(2) = 21.56, p < 0.0001$), and *confidence* ($H(2) = 20.63, p < 0.0001$). Post-hoc analysis for the *trust* comparison found that 'combined' ($M = 2.52$), $p < 0.0001$ and 'heatmap' ($M = 2.16$), $p = 0.0051$ were both ranked significantly higher than 'arrow' ($M = 1.32$). Likewise, post-hoc analysis for the *confidence* comparison also found that 'combined' ($M = 2.58$), $p < 0.0001$ and 'heatmap' ($M = 2.05$), $p = 0.032$ were both ranked significantly higher than 'arrow' ($M = 1.37$). These results all serve to **validate H1.a**.

Many participants shared similar insights in the post-experiment survey, reporting trust in the 'combined' condition over 'arrow' because they could reason about the rationale of the suggestions:

- *"The combination of a "safe" path and heatmap information helped me trust the system because I could compare the assessed path with the sensor information and make my own decision"*

We also found a significant effect in favor of the 'combined' condition over 'arrow' for the *Interpretability* scale ($H(2) = 8.26, p = 0.039$). Post-hoc analysis with Dunn's Test found that participants consistently preferred 'combined' ($M = 2.70$), $p = 0.040$ over 'arrow' ($M = 2.14$). There was an additional significant effect in the related post-experiment comparison measure of *helpfulness for decision-making* ($H(2) = 19.24, p < 0.0001$). Post-hoc analysis found that 'combined' ($M = 2.53$), $p < 0.0001$ and 'heatmap' ($M = 2.11$), $p = 0.0018$ were both ranked significantly higher than 'arrow' ($M = 1.37$). These results serve to **validate H1.b**.

Participants also emphasized how simply following the arrow-based guidance was easy, while noting that they were taking a leap of faith by following the suggestions, a feeling which was alleviated through the addition of the heatmap and its associated transparency.

- *"The arrows were certaintly "easier" to use...The heatmap [guidance] required more thought, but it made me more confident."*
- *"...with the heatmap you could see how confident the system was in its choices... The arrows alone were bad because you couldn't see why the system was changing its mind. "*

Though we found overall significance for the *Mental Load* scale ($H(2) = 6.68, p = 0.036$), there was not enough statistical power to make definitive post-hoc conclusions. Analysis with Dunn's Test found nearly significant effects for 'arrow' ($M = 2.63$) being rated as higher load than both 'heatmap' ($M = 2.24$), $p = 0.062$ and 'combined' ($M = 2.32$), $p = 0.099$. Interestingly, this effect appears to be indicating the opposite of hypothesis H1.c, showing that conditions containing prescriptive guidance are rated as more taxing. However, due to the lack of significance, **H1.c is inconclusive**, and will require more data to definitively address.

Some insight into this effect is visible though in participant reactions to path changes in the 'arrow' condition. Participants felt they needed to follow the guidance given to them since they had no other information, but felt stressed and irritated when they encountered sudden path changes, especially switchbacks.

- *"Arrow advice was frustrating when it kept changing the suggestions. I was not sure why it was happening."*
- *"I would like to be involved in the decision making, rather than being restricted by the guidance system. The arrow system essentially tells the player to trust its decision with no alternative consideration."*

The post-experiment comparison measure of *usefulness* also had significant effect. ($H(2) = 15.98, p = 0.0003$). Post-hoc analysis revealed significant effects for 'combined' ($M = 2.58$) being rated as more useful than both 'arrow' ($M = 1.89$), $p = 0.0003$ and 'combined' ($M = 1.53$), $p = 0.032$. Lastly, in asking which guidance participants would prefer to use in a hypothetical round 4, the significant favorite was also 'combined' based on a one-sample test of proportions (11/19 participants chose 'combined'; a greater proportion than the expected random proportion of 0.33, $p = 0.024$).

6.1.2 Objective Analysis. For measuring the performance of a round, we investigated two measures: *Total Time* and *Time per Move*. The domain was small enough that most participants solved it within a few moves of the optimal solution length. For all objective data analysis, we removed a single round out of the 57 conducted where the experiment was interrupted and the participant removed their HoloLens for an extended period of time, invalidating the data. We analyzed these performance metrics using a one-way analysis of variance (ANOVA) with experimental condition as a fixed effect. Post-hoc tests used Tukey's HSD to control for Type I errors in comparing performance across each guidance type.

The ANOVA revealed significant effects for both total time ($F(2, 53) = 3.91, p = 0.026$), and time per move ($F(2, 53) = 3.78, p = 0.029$). Post-hoc analysis for total time with Tukey's HSD shows

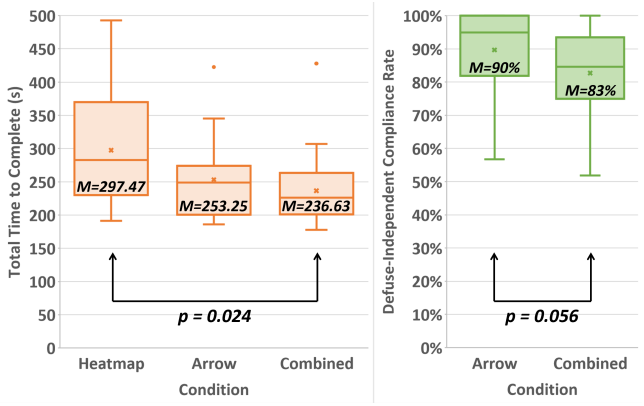

Figure 4: 'Combined' visualization achieves the Total Time performance benefits of 'arrow' while allowing for reduced rigidity in suggested action compliance.

that participants spent significantly less time solving the puzzle in the 'combined' condition ($M = 236.63s$), $p = 0.024$ compared to the 'heatmap' condition ($M = 297.47s$). The 'arrow' condition ($M = 253.25s$) fell in the middle, with no significant effects. Post-hoc analysis for time per move discovered that participants spent significantly less time per move in the 'arrow' condition ($M = 8.59s$), $p = 0.045$ compared to 'heatmap' ($M = 10.41s$), with 'combined' ($M = 8.74s$), $p = 0.066$ nearly achieving significantly lower time per move compared to 'heatmap'. The effects surrounding time and time per move serve to **validate H2**.

We were also interested in observing how differing compliance rates affected total moves in rounds using the 'arrow' and 'combined' conditions (conditions which contained prescriptive guidance), to see whether straying from the prescribed path led to changes in performance. Using Pearson's correlation coefficient, in the 'arrow' condition, there is a significant negative correlation between compliance rate and total moves (i.e., the more participants follow the guidance, the quicker they solve the puzzle) ($r(18) = -0.49, p = 0.039$). However, there is no such statistically significant correlation between compliance rate and total moves in the 'combined' condition ($r(19) = -0.11, p = 0.64$). This suggests that deviation from the path is a bad strategy when it is not informed, as in the case of 'arrow', but when there is extra information to work with such as the addition of PMF data in 'combined', it may be acceptable to deviate in certain cases.

Interviews from participants who deviated from the system's suggestions paint a similar picture: providing PMF data empowers people to act more independent of the guidance.

- *"It gives specific recommendations which are really just easy to use and follow. But it also gives you the broader understanding of the map to make deviations when they make sense."*

To determine the extent that this strategy was employed by participants, we compare the compliance rates of 'arrow' and 'combined'. Using a one-tailed t test, we measure whether participants strayed from the path more frequently in the presence of the added PMF data. Running this test, no significance was found between 'arrow' ($M = 0.83$) and 'combined' ($M = 0.78$); ($t(35) = -0.84, p = 0.20$). However, a high proportion of noncompliant moves were overly conservative defuse actions, especially early in rounds. By measuring the *defuse-independent* compliance rate between the two conditions, representing the frequency with which participants stayed on the same recommended path, we find a near-significant effect between 'arrow' ($M = 0.90$) and 'combined' ($M = 0.83$); ($t(35) = -1.63, p = 0.056$). This compliance data suggests that the addition of PMF data in 'combined' allows for more independence and injection of beneficial human decision compared to the monolithic 'arrow', and that participants are willing to take advantage of this. These findings **support and nearly validate H3**.

However, from the survey responses, it is evident that many participants altered their search strategy in the 'combined' condition: instead of entirely relying on the system's suggestions, participants started mixing the provided guidance with their own intuition.

- *"With just the arrow guidance, I was forced to follow it always since there was no other way to gather information. With the heatmap and combined (since it includes the heatmap) I was able to incorporate my own decisions as well."*

6.2 Discussion and Key Takeaways

We summarize key takeaways to inform the design of visual guidance systems for human-robot teaming, aligning with findings in the xAI literature where people consider robots to be more helpful and trustworthy when they justify their actions [11, 37].

T1: Prescriptive guidance, in the form of arrow or waypoint based suggestions, can be inherently restrictive. This guidance is easy to follow but puts human teammates in an 'automatic' pattern of thought (also known as system 1 thinking) [21]. In contrast, descriptive guidance forces the user to take more conscious actions (system 2 thinking). By combining both types of guidance, human teammates can leverage the explicit prescriptive guidance to help them reduce their workload, while still maintaining environmental awareness and acting with greater independence.

T2: In the 'arrow' condition, participants initially had a highly variable degree of trust in the system's suggestions. Some people over-trusted the guidance, taking its suggestions to be inherently correct, and some under-trusted the guidance, ignoring the arrow to defuse more conservatively. By providing descriptive data alongside prescriptive suggestions, people's behavior often tended towards a degree of trust somewhere in the middle of the two extremes, as they could see for themselves where a drone was more or less confident. This echoes findings on the ability of interpretable systems to mitigate over- and under-trust [9, 39].

T3: Some participants found it difficult to notice changes in the PMF when the change was not in their field of view. They suggested adding a feature notifying the user when a new high confidence target was found so they could be made aware of it. Additionally, some participants expressed desire to receive an explanation when a highly confident square suddenly becomes less confident.

T4: Participants did not like sudden path changes, viewing the behavior as unconfident. Participants expressed a preference for direct paths, desiring an explanation when a change was necessary.

ACKNOWLEDGMENTS

Special thanks to Aditi Periyannan (Tufts University) for their helpful suggestions on interface design. This work was funded as part of the Army Research Lab STRONG Program (#W911NF-20-2-0083).

REFERENCES

[1] Nisar Ahmed, Mark Campbell, David Casbeer, Yongcan Cao, and Derek Kingston. 2015. Fully Bayesian learning and spatial reasoning with flexible human sensor networks. In *Proceedings of the ACM/IEEE Sixth International Conference on Cyber-Physical Systems*. 80–89.

[2] Umang Bhatt, Javier Antorán, Yunfeng Zhang, Q Vera Liao, Prasanna Sattigeri, Riccardo Fogliato, Gabrielle Melançon, Ranganath Krishnan, Jason Stanley, Omesh Tickoo, et al. 2021. Uncertainty as a form of transparency: Measuring, communicating, and using uncertainty. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*. 401–413.

[3] Serena Booth, Yilun Zhou, Ankit Shah, and Julie Shah. 2020. Bayes-TrEx: a Bayesian Sampling Approach to Model Transparency by Example. *arXiv preprint arXiv:2002.10248* (2020).

[4] James V Bradley. 1958. Complete counterbalancing of immediate sequential effects in a Latin square design. *J. Amer. Statist. Assoc.* 53, 282 (1958), 525–528.

[5] Tathagata Chakraborti, Sarath Sreedharan, and Subbarao Kambhampati. 2020. The Emerging Landscape of Explainable Automated Planning & Decision Making.. In *IJCAI*. 4803–4811.

[6] THJ Collett and Bruce A MacDonald. 2006. Developer oriented visualisation of a robot program. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*. 49–56.

[7] Mark Colley, Benjamin Eder, Jan Ole Rixen, and Enrico Rukzio. 2021. Effects of Semantic Segmentation Visualization on Trust, Situation Awareness, and Cognitive Load in Highly Automated Vehicles. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–11.

[8] David Roxbee Cox and Nancy Reid. 2000. *The theory of the design of experiments*. CRC Press.

[9] Ewart J De Visser, Marieke MM Peeters, Malte F Jung, Spencer Kohn, Tyler H Shaw, Richard Pak, and Mark A Neerincx. 2020. Towards a theory of longitudinal trust calibration in human–robot teams. *International journal of social robotics* 12, 2 (2020), 459–478.

[10] Bruce H Deatherage. 1972. Auditory and other sensory forms of information presentation. *Human engineering guide to equipment design* (1972), 123–160.

[11] Upol Ehsan, Pradyumna Tambwekar, Larry Chan, Brent Harrison, and Mark O Riedl. 2019. Automated rationale generation: a technique for explainable AI and its effects on human perceptions. In *Proceedings of the 24th International Conference on Intelligent User Interfaces*. 263–274.

[12] Danyel Fisher. 2007. Hotmap: Looking at geographic attention. *IEEE transactions on visualization and computer graphics* 13, 6 (2007), 1184–1191.

[13] Marlena R Fraune, Ahmed S Khalaf, Mahlet Zemedie, Poom Pianpak, Zahra NaminiMianji, Sultan A Alharthi, Igor Dolgov, Bill Hamilton, Son Tran, and ZO Toups. 2021. Developing Future Wearable Interfaces for Human-Drone Teams through a Virtual Drone Search Game. *International Journal of Human-Computer Studies* 147 (2021), 102573.

[14] John R Frost. 1997. *The theory of search: a simplified explanation*. Soza Limited.

[15] Jack Gale, John Karasinski, and Steve Hillenius. 2018. Playbook for UAS: UX of Goal-Oriented Planning and Execution. In *International Conference on Engineering Psychology and Cognitive Ergonomics*. Springer, 545–557.

[16] Renan Luigi Martins Guarese and Anderson Maciel. 2019. Development and usability analysis of a mixed reality gps navigation application for the Microsoft Hololens. In *Computer Graphics International Conference*. Springer, 431–437.

[17] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology*. Vol. 52. Elsevier, 139–183.

[18] Bradley Hayes and Julie A Shah. 2017. Improving robot controller transparency through autonomous policy explanation. In *2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI*. IEEE, 303–312.

[19] Robert R Hoffman, Shane T Mueller, Gary Klein, and Jordan Litman. 2018. Metrics for explainable AI: Challenges and prospects. *arXiv preprint arXiv:1812.04608* (2018).

[20] Haikun Huang, Ni-Ching Lin, Lorenzo Barrett, Darian Springer, Hsueh-Cheng Wang, Marc Pomplun, and Lap-Fai Yu. 2017. Automatic optimization of wayfinding design. *IEEE transactions on visualization and computer graphics* 24, 9 (2017), 2516–2530.

[21] Daniel Kahneman. 2011. *Thinking, fast and slow*. Macmillan.

[22] Tijn Kooijmans, Takayuki Kanda, Christoph Bartneck, Hiroshi Ishiguro, and Norihiro Hagita. 2006. Interaction debugging: an integral approach to analyze human-robot interaction. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*. 64–71.

[23] Alexander Kunze, Stephen J Summerskill, Russell Marshall, and Ashleigh J Filtness. 2018. Augmented reality displays for communicating uncertainty information in automated driving. In *Proceedings of the 10th international conference on automotive user interfaces and interactive vehicular applications*. 164–175.

[24] Michael Lewis, Katia Sycara, and Phillip Walker. 2018. The role of trust in human-robot interaction. In *Foundations of trusted autonomy*. Springer, Cham, 135–159.

[25] Peter Lipton. 1990. Contrastive explanation. *Royal Institute of Philosophy Supplements* 27 (1990), 247–266.

[26] Scott M Lundberg and Su-In Lee. 2017. A unified approach to interpreting model predictions. In *Proceedings of the 31st international conference on neural information processing systems*. 4768–4777.

[27] Tim Miller. 2018. Contrastive explanation: A structural-model approach. *arXiv preprint arXiv:1811.03163* (2018).

[28] Tim Miller. 2019. Explanation in artificial intelligence: Insights from the social sciences. *Artificial intelligence* 267 (2019), 1–38.

[29] Brent Mittelstadt, Chris Russell, and Sandra Wachter. 2019. Explaining explanations in AI. In *Proceedings of the conference on fairness, accountability, and transparency*. 279–288.

[30] Savannah Paul, Christopher Reardon, Tom Williams, and Hao Zhang. 2020. Designing augmented reality visualizations for synchronized and time-dominant human-robot teaming. In *Virtual, Augmented, and Mixed Reality (XR) Technology for Multi-Domain Operations*, Vol. 11426. International Society for Optics and Photonics, 1142607.

[31] Christopher Reardon, Kevin Lee, John G Rogers, and Jonathan Fink. 2019. Communicating via augmented reality for human-robot teaming in field environments. In *2019 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*. IEEE, 94–101.

[32] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. 2016. " Why should i trust you?" Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*. 1135–1144.

[33] Wojciech Samek, Grégoire Montavon, Andrea Vedaldi, Lars Kai Hansen, and Klaus-Robert Müller. 2019. *Explainable AI: interpreting, explaining and visualizing deep learning*. Vol. 11700. Springer Nature.

[34] Ramprasaath R Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, and Dhruv Batra. 2017. Grad-cam: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE international conference on computer vision*. 618–626.

[35] Donghee Shin. 2021. The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *International Journal of Human-Computer Studies* 146 (2021), 102551.

[36] Daniel Szafir and Danielle Albers Szafir. 2021. Connecting Human-Robot Interaction and Data Visualization. In *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*. 281–292.

[37] Aaquib Tabrez, Shivendra Agrawal, and Bradley Hayes. 2019. Explanation-based reward coaching to improve human performance via reinforcement learning. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 249–257.

[38] Aaquib Tabrez, Matthew B Luebbers, and Bradley Hayes. 2020. A Survey of Mental Modeling Techniques in Human–Robot Teaming. *Current Robotics Reports* (2020), 1–9.

[39] Alan R Wagner, Jason Borenstein, and Ayanna Howard. 2018. Overtrust in the robotic age. *Commun. ACM* 61, 9 (2018), 22–24.

[40] Sebastian Wallkotter, Silvia Tulli, Ginevra Castellano, Ana Paiva, and Mohamed Chetouani. 2020. Explainable agents through social cues: A review. *arXiv preprint arXiv:2003.05251* (2020).

[41] Michał Wysokiński, Robert Marcjan, and Jacek Dajda. 2014. Decision support software for search & rescue operations. *Procedia Computer Science* 35 (2014), 776–785.

[42] Lu Yadong and Zhou Ya. 2015. Optimal Search and Rescue Model: Updating Probability Density Map of Debris Location by Bayesian Method. *International Journal of Statistical Distributions and Applications* 1, 1 (2015), 12.