RW4T Dataset: Data of Human-Robot Behavior and Cognitive States in Simulated Disaster Response Tasks

Liubove Orlov Savko*
Rice University
Houston, TX, USA
liubove.orlov.savko@rice.edu

Zhiqin Qian*
Rice University
Houston, TX, USA
bill.qian@rice.edu

Gregory Gremillion
Army Research Laboratory
Aberdeen Proving Ground, MD, USA
gregory.m.gremillion.civ@army.mil

Catherine Neubauer
Army Research Laboratory
Aberdeen Proving Ground, MD, USA
catherine.e.neubauer2.civ@army.mil

Jonroy Canady
Army Research Laboratory
Aberdeen Proving Ground, MD, USA
jonroy.d.canady.civ@army.mil

Vaibhav Unhelkar
Rice University
Houston, TX, USA
vaibhav.unhelkar@rice.edu

ABSTRACT

To forge effective collaborations with humans, robots require the capacity to understand and predict the behaviors of their human counterparts. There is a growing body of computational research on human modeling for human-robot interaction (HRI). However, a key bottleneck in conducting this research is the relative lack of data of human internal states – like intent, workload, and trust – which undeniably affect human behavior. Despite their significance, these states are elusive to measure, making the assembly of datasets a challenge and hindering the progression of human modeling techniques. To help address this, we first introduce Rescue World for Teams (RW4T): a configurable testbed to simulate disaster response scenarios requiring human-robot collaboration. Next, using RW4T, we curate a multimodal dataset of human-robot behavior and internal states in dyadic human-robot collaboration. This RW4T dataset includes state, action and reward sequences, and all the necessary data to replay a visual task execution. It further contains psychophysiological metrics like heart rate and pupillometry, complemented by self-reported cognitive state measures. With data from 20 participants, each undertaking five human-robot collaborative tasks, this dataset accompanied with the simulator can serve as a valuable benchmark for human behavior modeling.

CCS CONCEPTS

• **Human-centered computing**; • **Computer systems organization** → **Robotics**; • **Computing methodologies** → Artificial intelligence;

KEYWORDS

Human Modeling, Behavior Prediction, Mental States, Workload, Intent, Trust in Automation

*Both authors contributed equally to this research.

HRI '24, March 11–14, 2024, Boulder, CO, USA

ACM ISBN 979-8-4007-0322-5/24/03.
https://doi.org/10.1145/3610977.3637481

ACM Reference Format:
Liubove Orlov Savko, Zhiqin Qian, Gregory Gremillion, Catherine Neubauer, Jonroy Canady, and Vaibhav Unhelkar. 2024. RW4T Dataset: Data of Human-Robot Behavior and Cognitive States in Simulated Disaster Response Tasks. In *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction (HRI '24), March 11–14, 2024, Boulder, CO, USA.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3610977.3637481

1 STUDY OVERVIEW

Robots are becoming integral to various industries such as healthcare, manufacturing, and disaster response. However, their integration into industries, especially small and medium-sized enterprises, is hindered since robots need to operate at a slower speed or halt entirely near humans to ensure safety [22, 26]. For robots and humans to collaborate effectively, it is essential that robots can infer human internal states[1] and predict human actions accurately [10, 14, 25, 27]. Recognizing this need, multiple data-driven techniques have been developed to model human behavior, including imitation learning [15, 19, 23], batch learning from observations [11], and probabilistic models [12, 18, 30]. However, majority of these techniques are validated using benchmarks like OpenAI Gym [2] that do not model human-robot interaction, due to the lack of suitable HRI datasets.

To ensure applicability for HRI in the real world, these techniques must be tested on challenging and realistic data of human-robot behavior. However, the paucity of such data has limited its use in validation processes. Gathering this kind of data is inherently challenging as human behavior is influenced by a myriad of factors. Some of these are directly observable, like environmental conditions, while others, such as fatigue, workload, and trust, are latent and less tangible [21, 28, 29, 31]. **A review of the existing datasets, as shown in Table 1, underscores that most lack annotations related to these latent human states.** Although direct measurement of human internal states is challenging, they can be indirectly estimated through physiological metrics, such as heart rate and pupillometry, and through self-assessment instruments [3, 8, 9, 16].

To address this need, in this paper, we provide RW4T dataset: a benchmark dataset of human-robot behavior with annotation of human internal states. To compile this dataset, we followed a two-phase approach. *1) Simulation Development:* Using the Unity simulation engine, we developed Rescue World for Teams (RW4T).

[1]We use the term human internal states as an umbrella term to refer to cognitive constructs such as workload and intent, which are latent but impact human behavior.

Dataset	Behavior	Internal States	Simulator
RW4T (ours)	✓	multiple	✓
MoGaze [13]	✓	intent	-
MHHRI [5]	-	engagement	-
UE-HRI [1]	-	engagement	-
Robosuite [32]	-	-	✓
MineRL [6]	✓	-	✓

Table 1: An overview of relevant datasets. Behavior refers to task-oriented MDP (state-action)-trajectories.

This configurable testbed is designed to mimic disaster response scenarios that require human-robot collaboration. *2) Data Collection:* We invited participants to engage with RW4T, performing specific tasks while we monitor their behavior and cognitive states. The core task for participants was to deliver medical kits to designated points in a simulated disaster, with the assistance of a robot drone. Additionally, to induce varying levels of cognitive workload, participants were given secondary tasks to manage concurrently. Through this process, we curated a dataset of 100 trajectories from 20 participants, each undertaking five 2.5-minute HRI tasks.

The RW4T dataset is multimodal, encompassing raw continuous task features derived from the Unity simulation; processed state-action sequences; physiological metrics derived from the Zephyr BioHarness and Tobii Pro Eyetracker; self-reported measures of workload, trust, and engagement; and demographic data such as age, gender, and expertise in video games. Notably, expertise in video games offers insight into a participant's comfort level with virtual environments. Such expertise can substantially influence performance in the simulated tasks. Thus, it is an important variable in the analysis of behavioral data collected from RW4T.

A key feature of RW4T dataset is the availability of state-action sequences **with associated annotations of human internal states**, distinguishing it from other available datasets (cf. Table 1). The RW4T simulator and data together can be used in various facets of human modeling for HRI research. These include training, validation, and testing of predictive models of human behavior; investigations of how humans lean on robotic autonomy, especially in high-workload scenarios; determining the most suitable features for predicting human behavior; and delving into how human internal states shape the dynamics of human-robot collaboration.

2 METHODS

We now detail our approach to compile the RW4T dataset, which involved two phases: simulation creation and data collection.

2.1 Simulation Development

We designed a configurable simulation testbed called Rescue World for Teams (RW4T). RW4T simulates urban search-and-rescue tasks, which need to be collaboratively completed by a human-robot team. RW4T is implemented in Unity Engine to achieve photorealism and includes an API to facilitate data collection and task design. Further, RW4T includes secondary tasks and question prompts to easily modulate and measure human internal states corresponding to workload, intent, engagement, and trust. Finally, the testbed has been integrated with physiological sensors and lab streaming layer (LSL) to enable measurement of psychophysiological quantities that provide alternate mechanisms to estimate the ground truth values of human internal states [3, 9, 16].

Figure 1: First-person view of RW4T shows medical kits distribution and remaining time in the upper corners. The current score, with the maximum possible in parentheses, is in the middle. Participants control the robot drone via the robot panel, but the 'Auto' autonomy feature was disabled for this data collection. Above the robot panel is the minimap, which shows the full map (see Figure 2) when clicked. The lower right corner displays the F9-F12 windows, the interface for adjusting human cognitive workload in the secondary task.

2.1.1 Primary Task. As shown in Figs. 1-2, RW4T simulates dynamic and uncertain urban environments impacted by disasters. Some areas in the environment are unsafe for humans to enter and are marked as hazardous. Other areas are in need of rescue operations (i.e., delivery of medical supplies). A human first-responder is tasked with completing the rescue operations, with the help of a robot. The human-robot team's goal is to complete all rescue operations as soon as possible, while minimizing human exposure to the hazardous area. The team receives positive rewards for completing each rescue operations and negative rewards when the first responder enters the hazardous zone.

To successfully complete the task, the human-robot team needs to intelligently allocate tasks among team members. To enable study of HRI in a variety of scenarios, the task and environment can be reconfigured through a configuration API. For instance, one can change the environmental map, number and location of hazards, number and rescue locations, task rewards, and horizon. Further, the robot can be operated as autonomous (where it operates according to a predefined policy) or semi-autonomous (where it executes a goal-conditioned policy given a user's goal command).

In our data collection, the robot is configured as semi-autonomous. To emulate the unpredictable nature of robots, the drone was programmed with a 30% failure rate, potentially impacting participants' trust in its capabilities. The speed of the robot drone and the participant on the simulation testbed can also be configured. There is a trade-off between speeds and the time limit for completing a task: if the robot drone is faster than the participant, the participant will be prone to rely solely on the robot drone. Similar effect might occur if the time limit is large enough. Mathematically, the task can be viewed as a multi-agent Markov decision process (MDP), with fixed time horizon [17, 20].

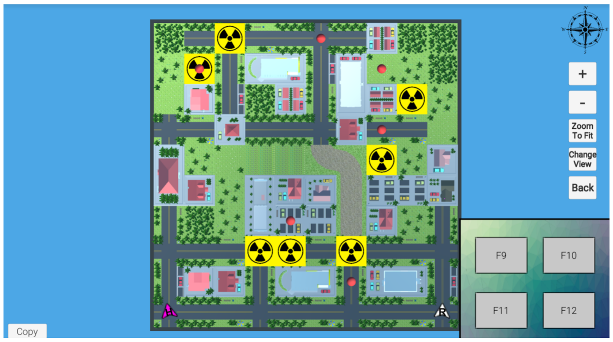

Figure 2: Configuration of RW4T map. Red dots represent locations to drop off medical kits, while hazard signs represent unsafe areas. The H and R arrows at the lower corners are the current positions of the human and robot, respectively.

2.1.2 Secondary Task. The simulator also includes an (optional) secondary task interface to allow researchers to modulate human cognitive states corresponding to workload. The secondary task design is informed by the system monitoring task from OpenMATB [4]. The human needs to monitor a panel of 4 gray buttons and press the corresponding key when a button changes color to yellow. The frequency of color change depends on the desired difficulty of the secondary task, which can be specified via the configuration API. The participant has limited time (`default=3` seconds) to press the corresponding key (`F9, F10, F11, F12`) to restore it to base color. If the secondary task is fulfilled successfully, the participant's reward increases by preset number of points (`default=10`), else they receive a penalty. Similar to the primary task, secondary task parameters (difficulty, reward, penalty) too can be configured.

2.1.3 Human Internal States. The simulator can readily capture environmental and behavioral data corresponding to the rescue task using the Unity engine, which can be further extracted via its Python API. Additionally, to estimate human internal states, RW4T includes two mechanisms: user prompts and integration with physiological sensors. For instance, to measure workload, the simulator can prompt users to report their workloads at a fixed frequency [7, 24]. We also provide integration with LSL and Python scripts to collect, synchronize, and process physiological measurements using Tobii eye tracker and Zephyr BioHarness [3, 9, 16].

2.2 Data Collection

Using the simulator, we collect data of human-robot behavior and human internal states through a human subject study, which was approved by Rice University's Institutional Review Board.

2.2.1 Participants. The dataset includes behavior of 20 participants, mostly campus graduate students and researchers. Their ages ranged from 19 to 54 years, median age = 27 years, 9 female.

2.2.2 Experimental Procedure. After giving informed consent, participants completed a demographic survey (Table 2), wore the BioHarness, and completed the calibration process of the eye tracker. To establish baseline physiological measurements, participants were requested to wait in front of the monitor for 3 minutes. They were then introduced to the RW4T simulator and experimental tasks through a tutorial and training trial. Participants completed five test trials, delivering 6 medical kits in each (Figure 2), with trials divided into low and high intensity periods for secondary tasks (c.f. Table 3). The high intensity periods involved a secondary task every five seconds, and no secondary tasks in low intensity periods. Participants filled out questionnaires during and post-trial.

ID	Demographic question
1	Age
2	Gender
3	What is your prior experience with video games
4	How often do you play video games?
5	On average, how much time do you spend each time you play a video game (in minutes)?
6	Open-ended comments (optional)
ID	**Post task questions**
1	How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)?
2	How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred?
3	How successful do you think you were in accomplishing the goals of the task?
4	How hard did you have to work (mentally) to accomplish your level of performance?
5	How stressed did you feel during the task?
ID	**In-task questions**
1	Rate your workload from 1 (low) to 5 (high)
2	Rate trust in robot from 1 (low) to 5 (high)
3	Rate level of engagement from 1 (low) to 5 (high)

Table 2: Experimental Questionnaires

Scenario	Duration	Secondary Task Intensity	
	(min)	Period 1	Period 2
Tutorial	≈20	-	-
Training trial	3	Low	Low
Test trial #1	2.5	Low	Low
Test trial #2	2.5	Low	High
Test trial #3	2.5	High	Low
Test trial #4	2.5	High	High
Test trial #5	2.5	Low	Low

Table 3: Overview of Experimental Procedure

2.2.3 Measurements. Task states, team actions, and physiological measurements were recorded throughout the whole experiment session. In-task questionnaires, listed in Table 2 were administered during each trial by pausing the task every 75 sec. The customized NASA-TLX questionnaire of Table 2 was administered upon completion of each trial [7]. Multimodal data arriving from different sources (Unity, physiological sensors, and questionnaires) was synchronized using the Lab Streaming Layer (LSL).

2.2.4 Compensation. Participants were compensated with $12 for their participation. We further motivated participants by awarding additional $12 to the best teaming performance with the robot.[2]

3 RW4T DATASET

The RW4T data consists of 100 trajectories of human-robot behavior and human internal states. As detailed in Sec. 4, the data is housed in a public GitHub repository. In this section, we summarize the structure, contents, and limitations of the dataset.

Data derived from the Unity Simulator. The dataset includes two versions of behavioral measurements: `raw` and `trajectories`. The `raw` data is provided as seven `csv` files per participant, corresponding to two training trials and five test trials. Each file contains time series data collected at 60 Hz of the task-relevant features (such as coordinates of the participant and robot, status of medical kits) and participants' actions (such as keyboard inputs and mouse clicks). Please refer to the associated data description (stored in the parent GitHub repository) for a complete list of raw features. The `raw` data provides uncompressed information about the task and, thus, can be used to replicate participants' trials. To facilitate algorithmic HRI research, we also provide `processed` version of this data in the form of MDP (state, action)-trajectories. This format is typically used in human modeling and imitation learning research to learn behavioral policies. We include both discrete and continuous representation of the task state, which consists of participant's position, robot's state, and the status of medical kits. We manually append the state-action trajectories with intent annotations, by observing the rescue locations the participant has visited. To arrive at the discretized version of the data, we model the environment as a 10×10 grid. We also provide the script to convert the `raw` data to its `processed` form. Using this script, the data can be processed at an alternate discretization as needed in applications.

Data derived from the Physiological Sensors. Physiological indicators can provide indirect measurements of human internal states such as workload and trust [9, 16]. While these measurements are indirect (i.e., they do not pinpoint the exact value of human internal states without additional post-processing), they can be collected at high temporal resolution without interfering with the HRI task. The dataset includes eye tracking data collected at 60 Hz using the Tobii Pro Nano sensor. The data is stored as one `json` file per participant. Each file contains time-stamped values of gaze location, pupil diameter for each eye, together with their validity. Further, the dataset includes additional physiological data collected at 1 Hz using the Zephyr BioHarness. This data is stored as one `csv` file per participant. Each file contains time-stamped values of physiological features such as heart rate, respiration rate, and posture.

Data derived from Questionnaires. Questionnaires provide direct measurements of human internal states through self-reports. However, they are subjective in nature, and querying the user during an ongoing task may be intrusive and hinder task performance [3]. The dataset includes responses to the NASA-TLX questionnaire, administered after each trial, and is stored in five `csv` files named as 'Post-task #.csv', where # is the trial number. Each row corresponds to a single participant. Responses to in-task prompts are provided as `txt` files, one corresponding to each participant. Lastly, the dataset includes demographic responses, stored in one `csv` file, with each row corresponding to one participant.

Data Organization. The simulator data, physiological data and responses to in-task questionnaires are grouped in the same folder by the participant's identifier, which is of the form 'user###', where each # represents a digit, inside the `raw` directory. An additional directory, named `trajectories`, contains files with MDP (state-action) trajectories. Responses to the post-task (NASA-TLX) and demographic questionnaires are located in the outer directory. Each row corresponds to a participant, which can be matched with the rest of the collected data through completion timestamps.

4 USAGE NOTES

Repositories and Documentation. The RW4T dataset, the eponymous simulator, and the scripts to collect and process data can be found at https://github.com/unhelkarlab/rw4t-dataset. The repository also offers documentation and scripts for loading the data specifically for research on human behavior modeling.

Implications. The RW4T dataset can serve as a useful benchmark for training and validating techniques for predictive modeling of human behavior in HRI. The trajectories in the dataset are diverse, making the human modeling task both realistic and challenging. It contains multimodal measures of human internal states and demographic data, both of which are intrinsically tied to human decision-making processes. Incorporating this data allows for generation of robust and personalized models of behavior, enhancing the efficacy and cohesion of human-robot collaboration. Furthermore, the multimodal data can allow researchers to analyze human-robot teaming at different levels of abstraction. For instance, besides predictive models, the dataset can be used to investigate how humans lean on robotic autonomy and how robot failures affect this process.

Contributing to the Dataset. Researchers using this dataset should also recognize its limitations. It was curated within a controlled experimental setting, restricted to search and rescue scenarios, with data corresponding to 20 individuals recruited from Rice University, with fixed robot failure rate to lessen the variability in this already highly variable dataset. As such, it is expected to represent only a subset of HRI scenarios observed in the real world. To facilitate mitigating some of these limitations, we also release the configurable RW4T simulator and data collection scripts. These artifacts can help researchers collect additional data of human-robot behavior, in the context of simulated disaster response. We hope that researchers as needed use the simulator to not only collect more data but also share the new data with the community, thereby making the dataset more comprehensive for its applications in HRI research.

ACKNOWLEDGMENTS

This research was supported by the Army Research Office through CA# W911NF-20-2-0214. The authors also thank COMEXUS and the Fulbright Program for supporting Liubove Orlov-Savko through the Fulbright-García Robles fellowship for Graduate Studies.

[2]To quantify performance, the human-robot team was awarded 25 points for distributing each first aid kit and 10 points for completing a secondary task. The team was penalized by 10 points for each unsuccessful attempt on a secondary task and for every second the human spent in a radioactive zone. If the team completed the task before the allotted time, the remaining time (in seconds) was added to their final score.

REFERENCES

[1] Atef Ben-Youssef, Chloé Clavel, Slim Essid, Miriam Bilac, Marine Chamoux, and Angelica Lim. 2017. UE-HRI: A new dataset for the study of user engagement in spontaneous human-robot interactions. In *ACM International Conference on Multimodal Interaction (ICMI)*. 464–472.

[2] Greg Brockman, Vicki Cheung, Ludwig Pettersson, Jonas Schneider, John Schulman, Jie Tang, and Wojciech Zaremba. 2016. OpenAI gym. *arXiv preprint arXiv:1606.01540* (2016).

[3] Stephen M Casner and Brian F Gore. 2010. Measuring and evaluating workload: A primer. *NASA Technical Memorandum* 216395, 2010.

[4] J Cegarra, B Valéry, Eugénie Avril, C Calmettes, and J Navarro. 2020. OpenMATB: A Multi-Attribute Task Battery promoting task customization, software extensibility and experiment replicability. *Behavior research methods* 52 (2020), 1980–1990.

[5] Oya Celiktutan, Efstratios Skordos, and Hatice Gunes. 2017. Multimodal human-human-robot interactions (MHHRI) dataset for studying personality and engagement. *IEEE Transactions on Affective Computing* 10, 4 (2017), 484–497.

[6] William H Guss, Brandon Houghton, Nicholay Topin, Phillip Wang, Cayden Codel, Manuela Veloso, and Ruslan Salakhutdinov. 2019. MineRL: A large-scale dataset of minecraft demonstrations. (2019), 2442–2448.

[7] 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.

[8] Jamison Heard and Julie A Adams. 2019. Multi-dimensional human workload assessment for supervisory human–machine teams. *Journal of Cognitive Engineering and Decision Making* 13, 3 (2019), 146–170.

[9] Jamison Heard, Caroline E. Harriott, and Julie A. Adams. 2018. A survey of workload assessment algorithms. *IEEE Transactions on Human-Machine Systems* 48, 5 (2018), 434–451.

[10] Laura M Hiatt, Cody Narber, Esube Bekele, Sangeet S Khemlani, and J Gregory Trafton. 2017. Human modeling for human–robot collaboration. *The International Journal of Robotics Research* 36, 5-7 (2017), 580–596.

[11] Daniel Jarrett, Ioana Bica, and Mihaela van der Schaar. 2020. Strictly batch imitation learning by energy-based distribution matching. *Advances in Neural Information Processing Systems* 33, 7354–7365.

[12] Ece Kamar, Ya'akov Gal, and Barbara J Grosz. 2009. Incorporating helpful behavior into collaborative planning. In *International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. Springer Verlag.

[13] Philipp Kratzer, Simon Bihlmaier, Niteesh Balachandra Midlagajni, Rohit Prakash, Marc Toussaint, and Jim Mainprice. 2020. MoGaze: A dataset of full-body motions that includes workspace geometry and eye-gaze. *IEEE Robotics and Automation Letters* 6, 2 (2020), 367–373.

[14] Przemyslaw A Lasota, Terrence Fong, Julie A Shah, et al. 2017. *A survey of methods for safe human-robot interaction*. Now Publishers.

[15] Yunzhu Li, Jiaming Song, and Stefano Ermon. 2017. Infogail: Interpretable imitation learning from visual demonstrations. *Advances in Neural Information Processing Systems (NeurIPS)* 30 (2017).

[16] Catherine Neubauer, Kristin E Schaefer, Ashley H Oiknine, Steven Thurman, Benjamin Files, Stephen Gordon, J Cortney Bradford, Derek Spangler, and Gregory Gremillion. 2020. *Multimodal Physiological and Behavioral Measures to Estimate Human States and Decisions for Improved Human Autonomy Teaming*. Technical Report. CCDC Army Research Laboratory, Aberdeen Proving Ground, United States.

[17] Frans A Oliehoek, Christopher Amato, et al. 2016. *A concise introduction to decentralized POMDPs*. Vol. 1. Springer.

[18] Liubove Orlov-Savko, Abhinav Jain, Gregory M Gremillion, Catherine E Neubauer, Jonroy D Canady, and Vaibhav Unhelkar. 2022. Factorial Agent Markov Model: Modeling Other Agents' Behavior in presence of Dynamic Latent Decision Factors. In *International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. 982–990.

[19] Takayuki Osa, Joni Pajarinen, Gerhard Neumann, J Andrew Bagnell, Pieter Abbeel, Jan Peters, et al. 2018. An Algorithmic Perspective on Imitation Learning. *Foundations and Trends® in Robotics* 7, 1-2 (2018), 1–179.

[20] Martin L Puterman. 2014. *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons.

[21] Peizhu Qian and Vaibhav V Unhelkar. 2022. Evaluating the Role of Interactivity on Improving Transparency in Autonomous Agents. In *International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*. IFAAMAS.

[22] Lindsay Sanneman, Christopher Fourie, Julie A Shah, et al. 2021. The state of industrial robotics: Emerging technologies, challenges, and key research directions. *Foundations and Trends® in Robotics* 8, 3 (2021), 225–306.

[23] Sangwon Seo and Vaibhav Unhelkar. 2022. Semi-Supervised Imitation Learning of Team Policies from Suboptimal Demonstrations. In *International Joint Conference on Artificial Intelligence (IJCAI)*.

[24] Andrew J Tattersall and Penelope S Foord. 1996. An experimental evaluation of instantaneous self-assessment as a measure of workload. *Ergonomics* 39, 5 (1996), 740–748.

[25] Andrea Thomaz, Guy Hoffman, and Maya Cakmak. 2016. Computational human-robot interaction. *Foundations and Trends in Robotics* 4, 2-3 (2016), 105–223.

[26] Peter Trautman and Andreas Krause. 2010. Unfreezing the robot: Navigation in dense, interacting crowds. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 797–803.

[27] Vaibhav V Unhelkar, Shen Li, and Julie A Shah. 2019. Semi-Supervised Learning of Decision-Making Models for Human-Robot Collaboration. In *Conference on Robot Learning (CoRL)*.

[28] Vaibhav V Unhelkar, Shen Li, and Julie A Shah. 2020. Decision-Making for Bidirectional Communication in Sequential Human-Robot Collaborative Tasks. In *ACM/IEEE International Conference on Human-Robot Interaction (HRI)*.

[29] Vaibhav V Unhelkar and Julie A Shah. 2018. Learning and Communicating the Latent States of Human-Machine Collaboration.. In *International Joint Conference on Artificial Intelligence (IJCAI)*. 5789–5790.

[30] Vaibhav V Unhelkar and Julie A Shah. 2019. Learning Models of Sequential Decision-Making with Partial Specification of Agent Behavior. In *AAAI Conference on Artificial Intelligence (AAAI)*.

[31] X. Jessie Yang, Vaibhav V. Unhelkar, Kevin Li, and Julie A. Shah. 2017. Evaluating Effects of User Experience and System Transparency on Trust in Automation. In *ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (Vienna, Austria). 408–416.

[32] Yuke Zhu, Josiah Wong, Ajay Mandlekar, Roberto Martín-Martín, Abhishek Joshi, Soroush Nasiriany, and Yifeng Zhu. 2020. robosuite: A modular simulation framework and benchmark for robot learning. *arXiv preprint arXiv:2009.12293* (2020).