Enabling Effective Human-Robot Collaboration via Trust-Driven Decision-Making

SHREYAS BHAT

Agenda

- Introduction
- **Phase 1** – Trust-Aware Markov Decision Process
- **Phase 2** – Effects of Real-time Personalization of Reward Weights
- **Phase 3** – Effects of Fine-grained Reward Learning and State Space Exploration
- Future Research Directions

From Automation to Autonomy – From Tools to Teammates

Automation

- Rule based systems
- Task execution
- Human-directed
- Repetitive work

→

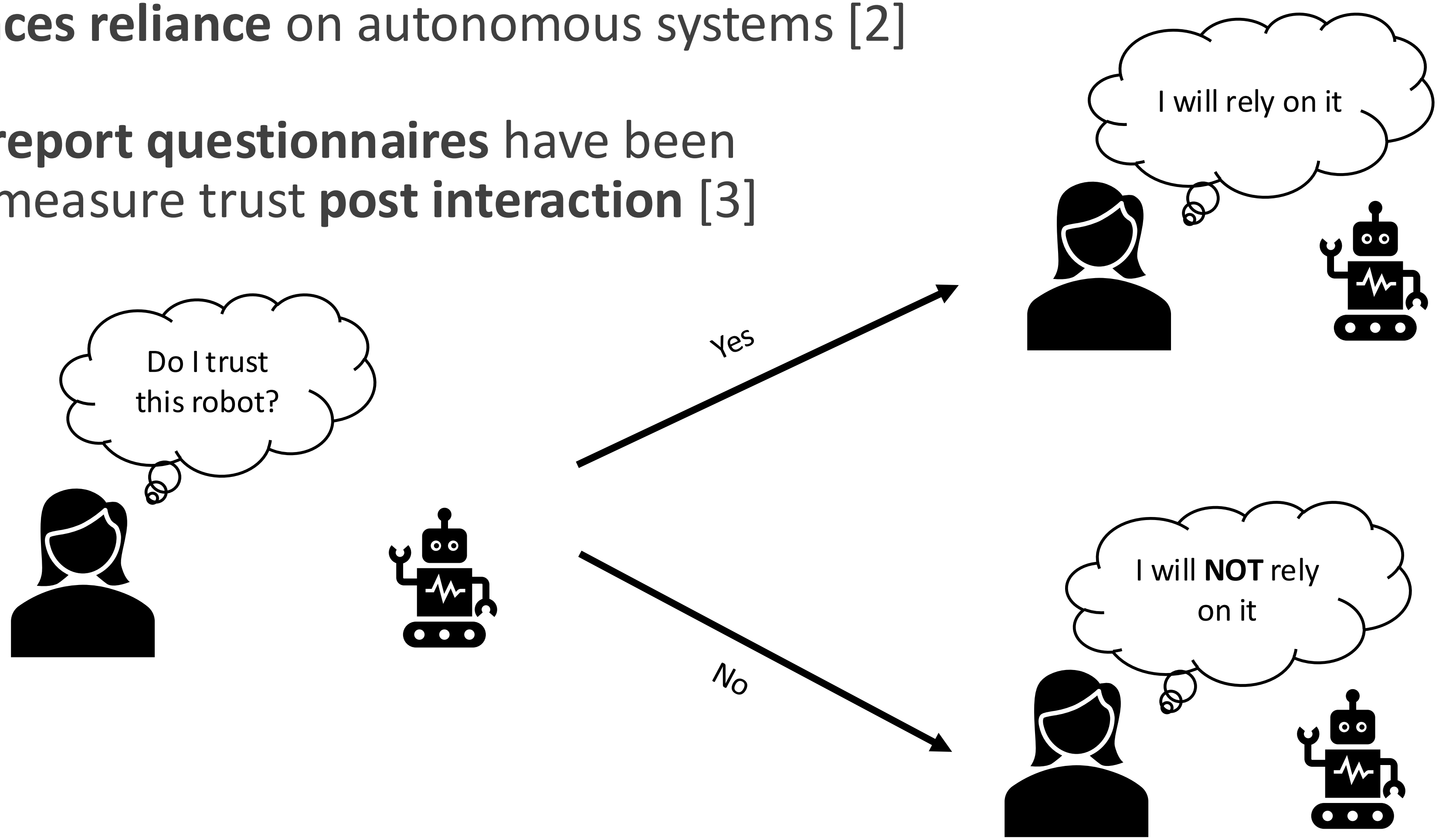
Autonomy

- Context-Aware
- Goal-Oriented
- Collaborative
- Adaptive

1. ENDSLEY (2017)

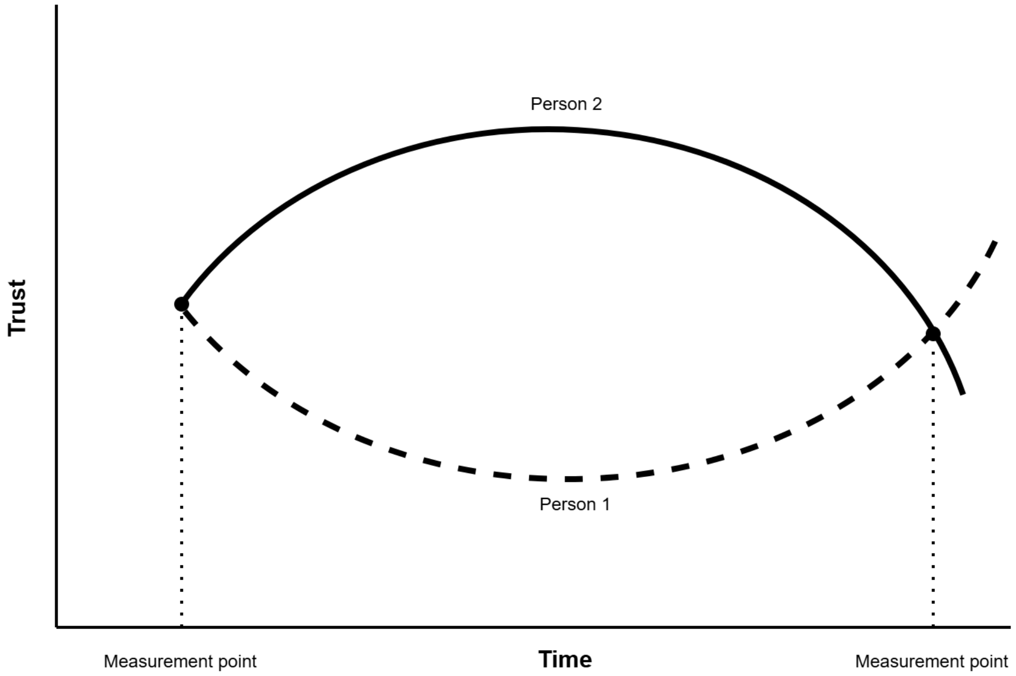
Why is Trust Important in HRI?

- **Trust influences reliance** on autonomous systems [2]
- Several **self-report questionnaires** have been designed to measure trust **post interaction** [3]



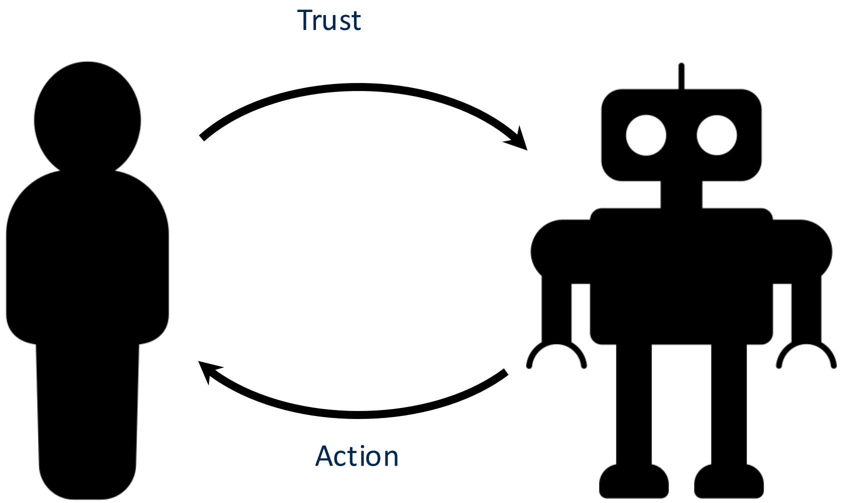
Why is Trust Important in HRI?

- Trust influences reliance on autonomous systems [2]
- Several self-report questionnaires have been designed to measure trust post interaction [3]
- **Trust is dynamic** and varies during interaction [4]
- Thus, mathematical models of trust were designed, capable of **estimating moment-to-moment trust** [5]



Why is Trust Important in HRI?

- Trust influences reliance on autonomous systems [2]
- Several self-report questionnaires have been designed to measure trust post interaction [3]
- Trust is dynamic and varies during interaction [4]
- Thus, mathematical models of trust were designed, capable of estimating moment-to-moment trust [5]
- **Open question** – How do we incorporate these trust models into the decision-making system of autonomous agents?



Agenda

- Introduction
- **Phase 1** – Trust-Aware Markov Decision Process
- Phase 2 – Effects of Real-time Personalization of Reward Weights
- Phase 3 – Effects of Fine-grained Reward Learning and State Space Exploration
- Future Research Directions

Trust-Aware Markov Decision Process (MDP)

- We mainly focus on scenarios where the **robot acts as an action recommender** to the human

Item	Description
States	**Trust**, Contextual Information
Actions	Actions recommended by the robot and implemented by the human
Transition Function	**Dynamic Trust Model**, Contextual Information Updates
Reward Function	Rewards obtained for choosing actions in specific states
Human Behavior Model	Probability of the **human choosing an action** given the recommendation

Table 1 - Components of the Trust-Aware MDP

- However, this framework can be **readily extended to other HRI scenarios**

Human-Robot Team Task

- Intelligence, Surveillance, and Reconnaissance (ISR) Mission scenario
- A human-robot team performs a sequential search for potential threats in an abandoned town
- At each search site, there are two actions
  - **USE** the armored robot – **costs time, safer**
  - **NOT USE** the armored robot – **may cost health, quicker**

Their objective is to minimize the loss of time and health

Trust Aware MDP – States and Transition Function

- **Trust** is modeled as a **Beta Distribution** with parameters α and β

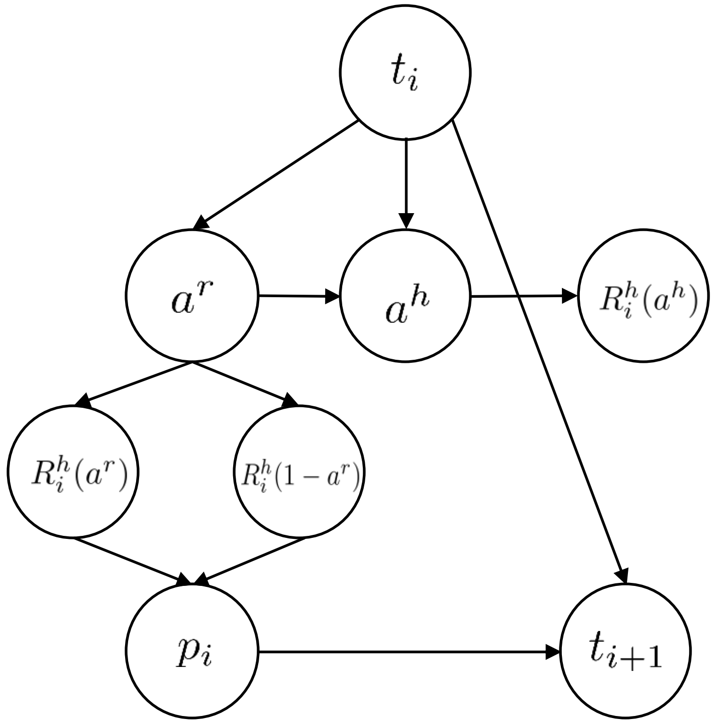
$$t_i \sim \mathrm{Beta}(\alpha_i, \beta_i)$$

- The **dynamic trust model** [3] is the transition function

$$\alpha_{i+1} = \alpha_i + v^s p_i,$$
$$\beta_{i+1} = \beta_i + v^f (1 - p_i).$$

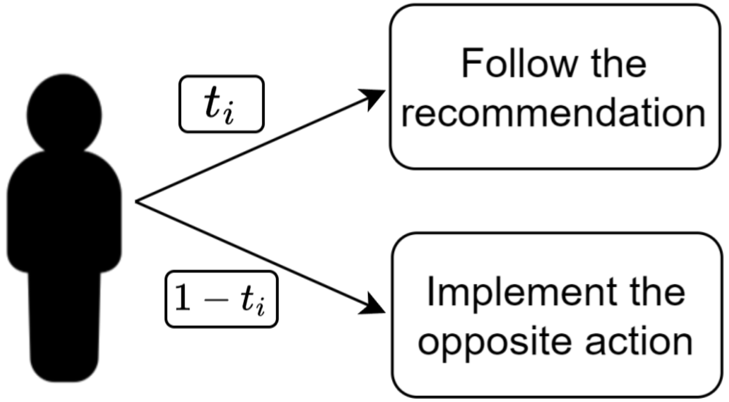
- We propose a **reward-based performance metric** for the sequential decision-making paradigm

$$p_i = \begin{cases} 1 & \text{if } \; R_i^h(a^r) \geq R_i^h(1-a^r), \\ 0 & \text{otherwise.} \end{cases}$$



Trust Aware MDP – Human Behavior Model

- We use the **reverse psychology model** of human behavior [6]



- Mathematically,

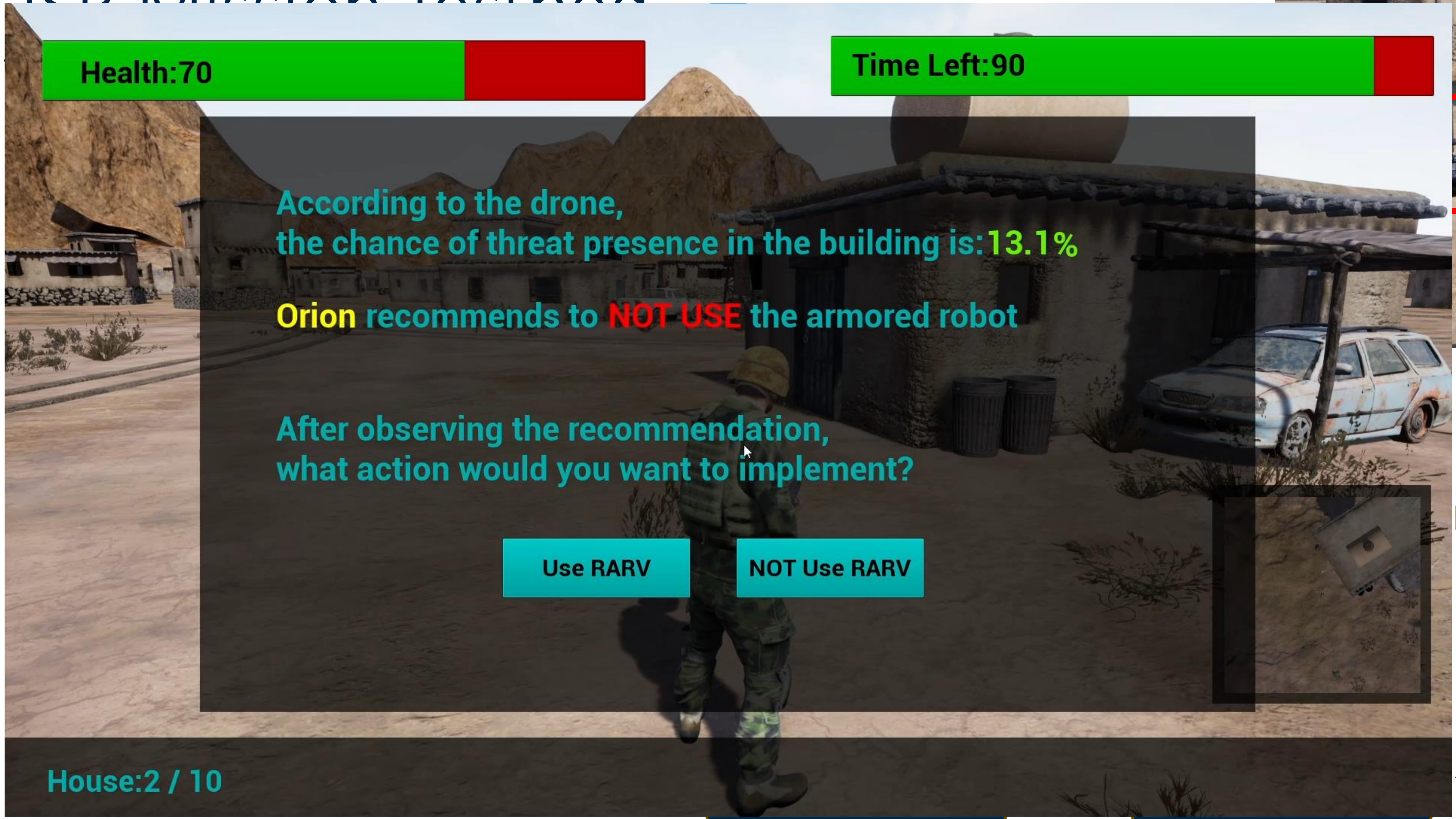
$$P(a_i^h = a_i^r) = t_i,$$
$$P(a_i^h = 1 - a_i^r) = 1 - t_i.$$

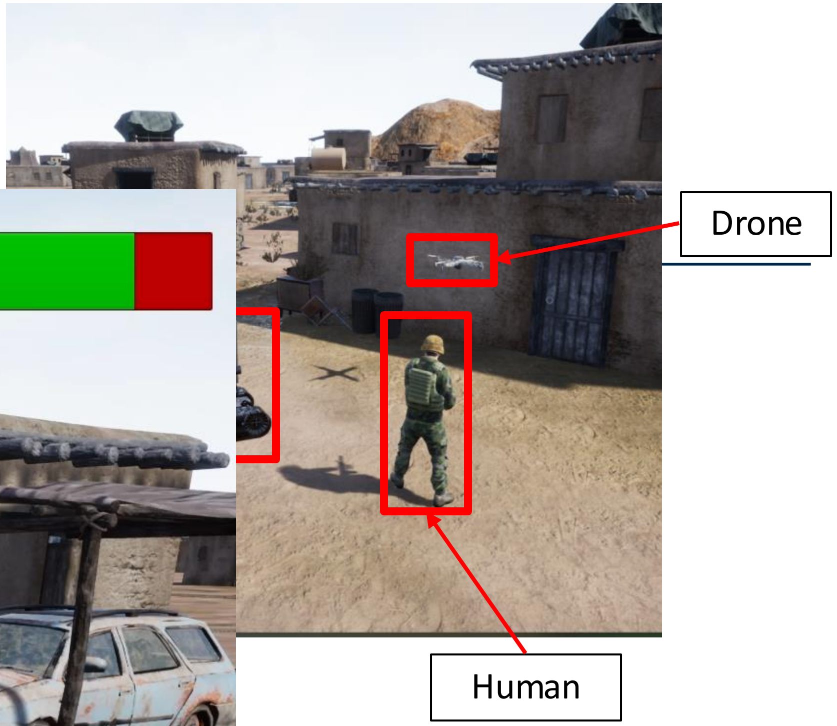
Trust Aware MDP - Reward Function

$$R_i(a_i^h, D_i) = \overbrace{-w_h h(a_i^h, D_i) - w_c c(a_i^h)}^{\text{Task Reward}} + \underbrace{\lambda_i \cdot \mathbb{1}(A)}_{\text{Trust Reward}}$$

- Guo et al. [6] observed that the robot tends to fall into the **reverse psychology loop** in the **absence of a trust-gaining reward term**
- So, we added a **trust-gaining reward term** to the robot's reward function
- The **performance metric only uses the task reward** to judge the robot's recommendation

ISR Mission Testbed





Agent recommends action

Human selects action

Human-Subjects Experiment

- 46 students from the University of Michigan participated
  - 21 Female, Age 22.8 ± 3.6 years

- Measures:
  - Big 5 Personality Traits [7]
  - Perfect Automation Schema [8]
  - Propensity to Trust [9]
  - Trust after each site
  - Post-experiment Trust [10]
  - Workload [11]

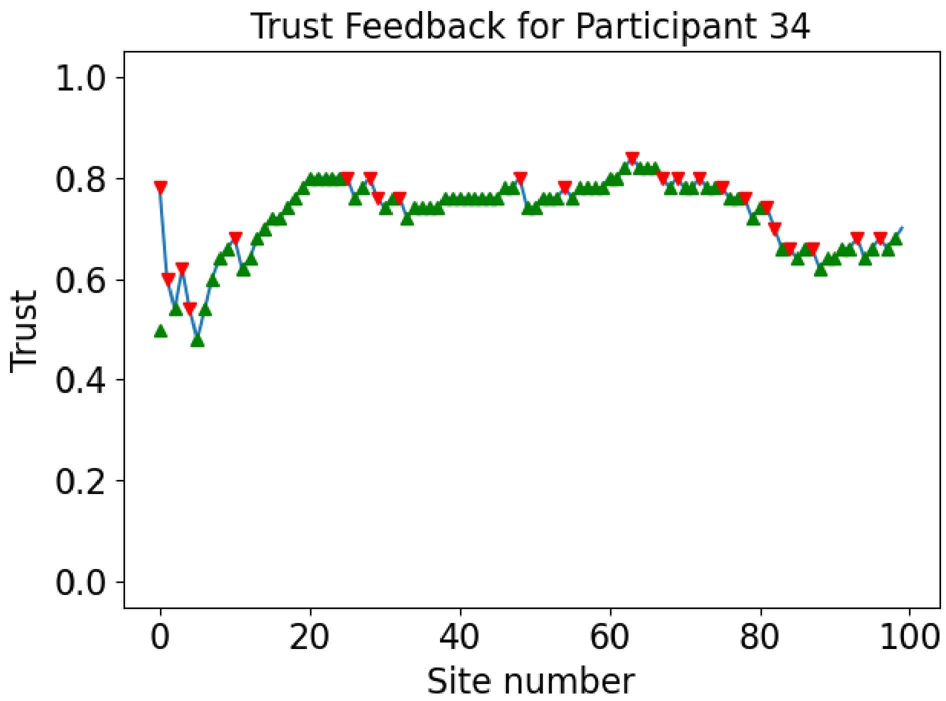
- Participants searched through 100 sites sequentially

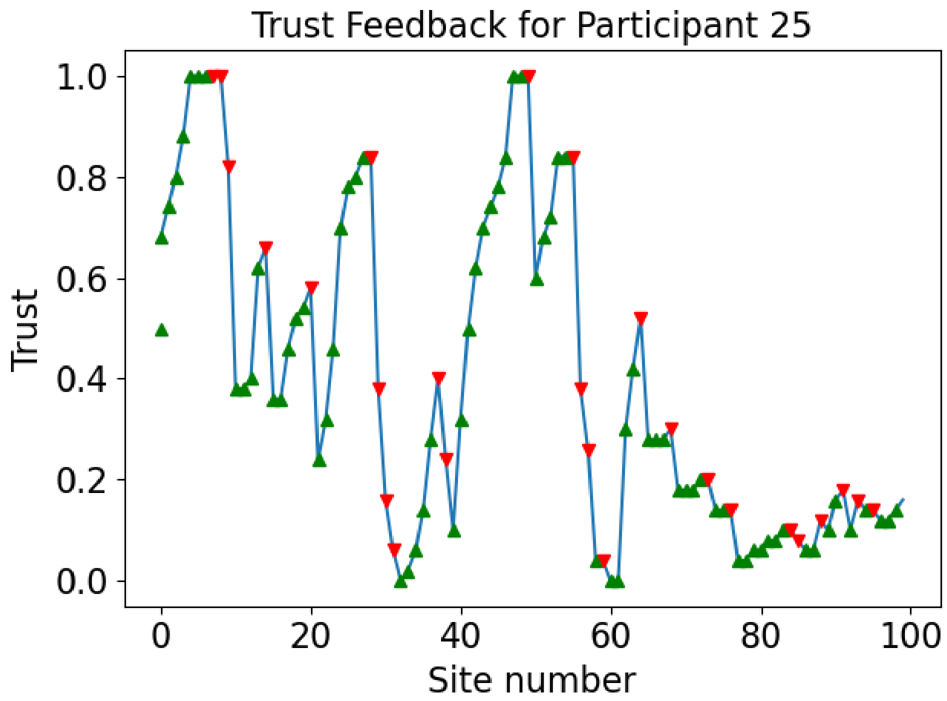
7. DONELLAN (2006) 8. LYONS (2019) 9. MERRITT (2013) 10. MUIR (1996) 11. HART (1988)

Results – Reward based performance metric

- $p_i = 1$

- $p_i = 0$



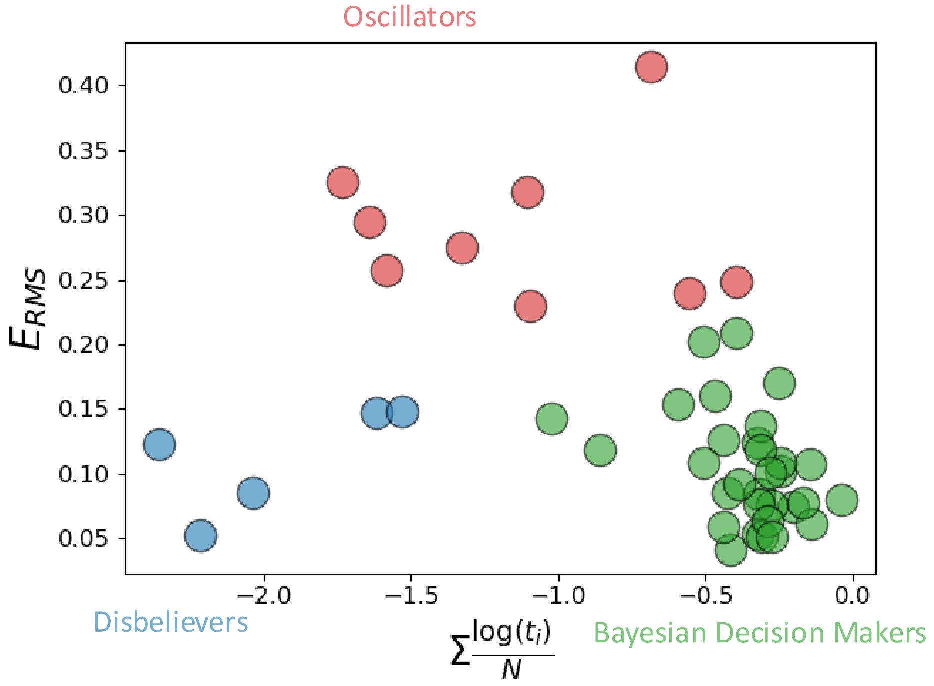


Results – Clustering Analysis

- K-means clustering analysis

- **Features**:
  - RMSE between feedback and predicted trust
  - Average log trust

- Elbow heuristic and silhouette scores indicate **3 significant clusters**



Results – Associations with Personal Characteristics

- Disbelievers **less extroverted** than Oscillators
- Disbelievers have **lower expectations** from autonomy

TABLE I
MEAN AND STANDARD DEVIATION (SD) OF PERSONAL CHARACTERISTICS BETWEEN THE THREE DIFFERENT TRUST DYNAMICS (BDM = BAYESIAN DECISION MAKER)

Personal Characteristic	BDM	Disbeliever	Oscillator
Extraversion (/20) *	9.5 (3.3)	5.8 (2.8)	11.3 (2.9)
Agreeableness (/20) *	13.5 (2.5)	10.4 (5.0)	14.1 (1.8)
Conscientiousness (/20)	13.1 (2.7)	12.4 (3.0)	12.1 (4.5)
Neuroticism (/20)	7.9 (2.7)	6.8 (3.6)	10.2 (4.7)
Intellect/Imagination (/20) †	11.7 (2.0)	9.8 (1.8)	12.2 (1.8)
High Expectations (/28) **	12.7 (3.9)	6.4 (2.8)	12.4 (4.2)
All or None Thinking (/21)	6.6 (2.9)	6.4 (3.4)	7.1 (3.1)
Trust Propensity (/30) †	20.2 (4.4)	17.2 (4.1)	22.8 (3.2)

$** - p < 0.01$, $* - p < 0.05$, $\dagger - p < 0.1$

Phase 1 – Key Takeaways and Limitations

KEY TAKEAWAYS

- Demonstrated the efficacy of the **trust-aware MDP framework**
- Found **3 types of trust dynamics** exhibited by people
- Showed the effectiveness of the **reward-based performance metric** to capture the **internal trust dynamics** of humans

LIMITATIONS

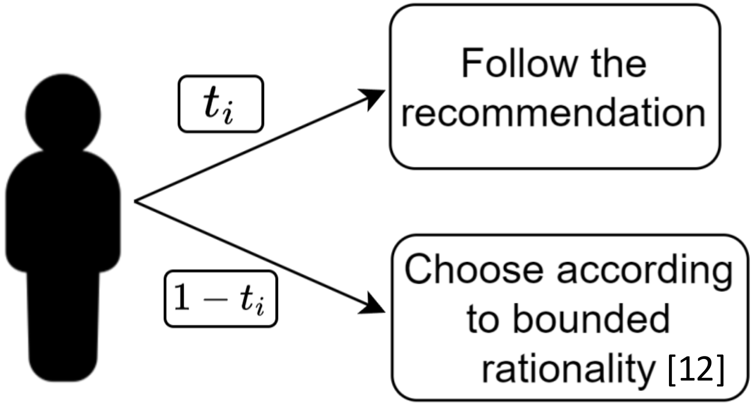
- Assumed that the human and robot **share a common reward function**
- Used the Reverse Psychology Model
  - Only applicable to **binary action** scenarios
  - Does **not consider the preferences** of the human
  - Requires a **trust-gaining reward term**

Agenda

- Introduction
- Phase 1 – Trust-Driven Markov Decision Process
- **Phase 2** – Effects of real-time personalization of reward weights
- Phase 3 – Effects of fine-grained reward learning and state space exploration
- Future Research Directions

Bounded Rationality Disuse Model of Human Behavior

- We introduce the **Bounded Rationality Disuse Model** of Human Behavior



- Mathematically,

$$P(a_i^h = a | a_i^r = a, t_i, w^h) = t_i + (1 - t_i) q_a(w^h),$$

$$P(a_i^h = 1 - a | a_i^r = a, t_i, w^h) = (1 - t_i)(1 - q_a(w^h)).$$

$$q_a(w^h) = \frac{\exp(\kappa E[R_i^h(a)])}{\sum_{a' \in \{0,1\}} \exp(\kappa E[R_i^h(a')])}.$$

Benefits Over Reverse Psychology Model

BOUNDED RATIONALITY DISUSE

- **Uses the underlying preferences** of the human
- Extensible to **multi-action scenarios**
- Robot **only prefers high-trust states**, thus **removing** the requirement for a **trust-gain reward term**

REVERSE PSYCHOLOGY

- **Does not use the underlying preferences** of the human
- Restricted to **binary action scenarios**
- Robot prefers both high- and **low-trust** states, **requiring the trust-gain reward term**

Reward Function

Phase 1

$$R_i(a_i^h, D_i) = \overbrace{-w_h h(a_i^h, D_i) - w_c c(a_i^h)}^{\text{Task Reward}} + \underbrace{\lambda_i \cdot \mathbb{1}(A)}_{\text{Trust Reward}}$$

Phase 2

$$R^o(a_i^h, D_i) = -w^o h(a_i^h, D_i) - (1 - w^o) c(a_i^h) \qquad \text{where, } o \in \{h, r\}$$

- We **remove the trust-gain reward term** from Phase 1
- We separate the reward functions for the two agents – Human (h) and Robot (r)
- We assume that the reward is a **convex combination** of health-loss cost and time-loss cost

Bayesian Inverse Reinforcement Learning

- **Core Idea:**
  - Maintain a **belief distribution on the reward weights** of the human $b_i(w^h)$
  - **Update it** after observing the interaction **using Bayes' rule** on the Behavior Model

$$b_{i+1}(w) \propto \begin{cases} P(a_i^h = a_i^r | a_i^r, t_i, w) b_i(w), & \text{if } a_i^h = a_i^r, \\ P(a_i^h = 1 - a_i^r | a_i^r, t_i, w) b_i(w), & \text{otherwise.} \end{cases}$$

- The algorithm needs an **initial distribution** $b_0(w)$ to get started
- We present results from **two human-subjects' studies that differ in this initial distribution**

Human-Subject Studies - Conditions

- We design three interaction strategies for the robot

Non-Learner

- Assumes that the human and the robot **share the same reward weights**
- **No reward learning** is performed
- **Similar to Phase 1**

Non-Adaptive Learner

- **Learns personalized reward weights** for each human it interacts with
- Only uses these personalized weights for
  - Behavior Prediction
  - Performance Estimation
- **Solves trust-aware MDP for fixed reward weights** (similar to non-learner)

Adaptive Learner

- Learns **personalized reward weights** for each human it interacts with
- Uses these personalized weights for
  - Behavior Prediction
  - Performance Estimation
  - Solving trust-aware MDP
- In essence, **adopts these learned reward weights** as its own

Human-Subjects Experiments

STUDY 1 – INFORMED PRIOR

- The robot starts its learning algorithm from an **informed prior** on the reward weights
- 30 participants

STUDY 2 – UNIFORM PRIOR

- The robot starts its learning algorithm from a **uniform prior** on the reward weights
- 24 participants

NOTE

The **non-adaptive learner** and the **non-learner** strategies use the **mean of the corresponding prior** as the weights for the robot’s reward function

Human-Subjects Experiments – Details

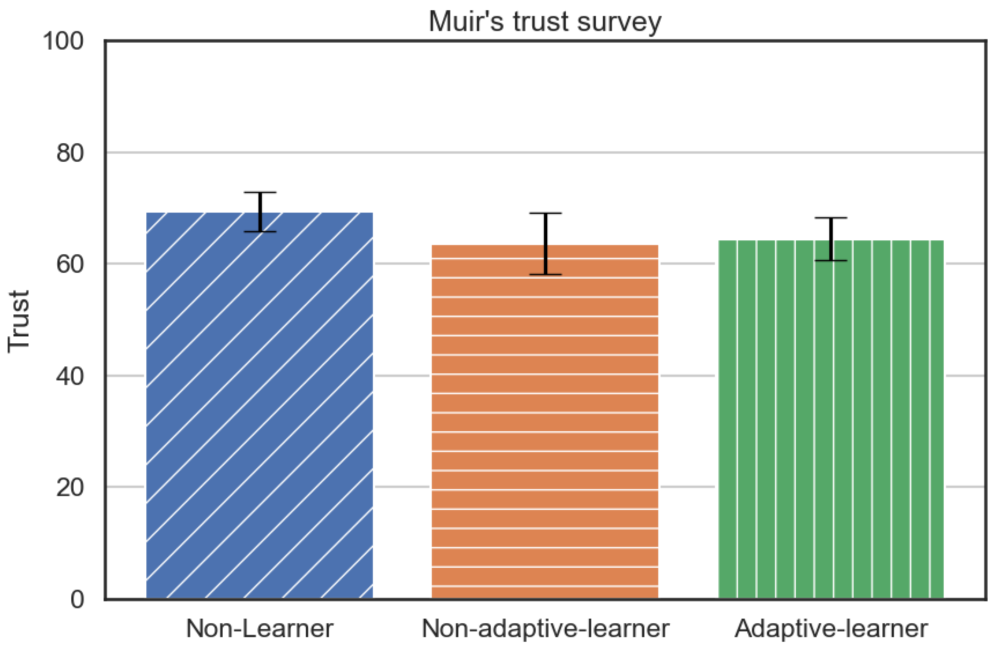
- Within-subjects design
  - Each participant completed **3 missions**
  - Each mission used 1 of the 3 interaction strategies
  - Counterbalanced ordering

- Each mission contained **40 sequential searches**

- Team started with **100 health** and **100 time** points
  - Time cost – **5 points**
  - Health cost – **5 points**

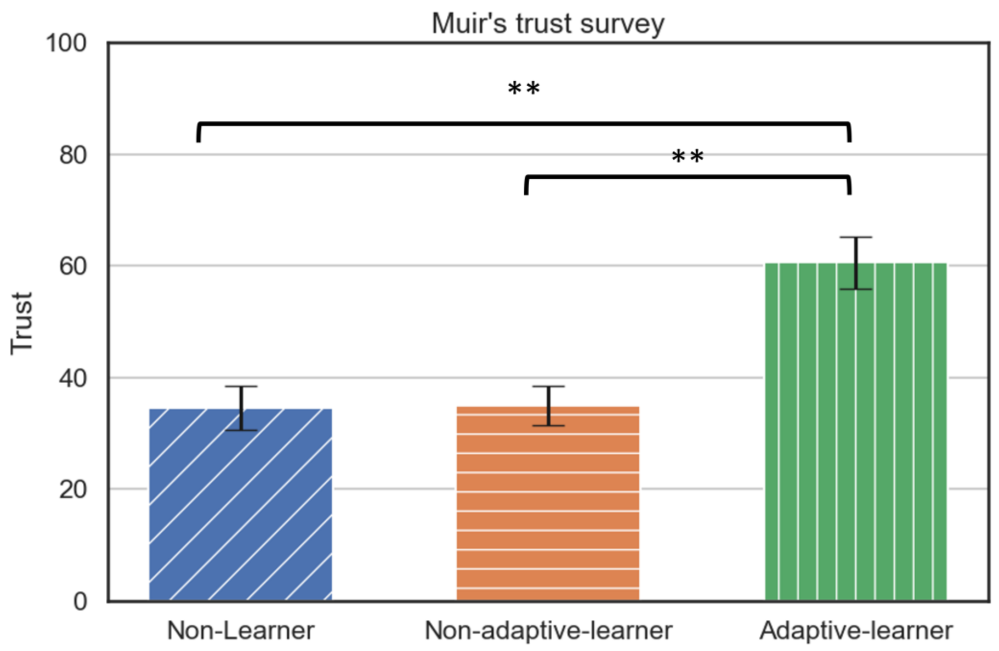
Results – Subjective Trust

STUDY 1 – INFORMED PRIOR



- **No significant difference** between the three strategies

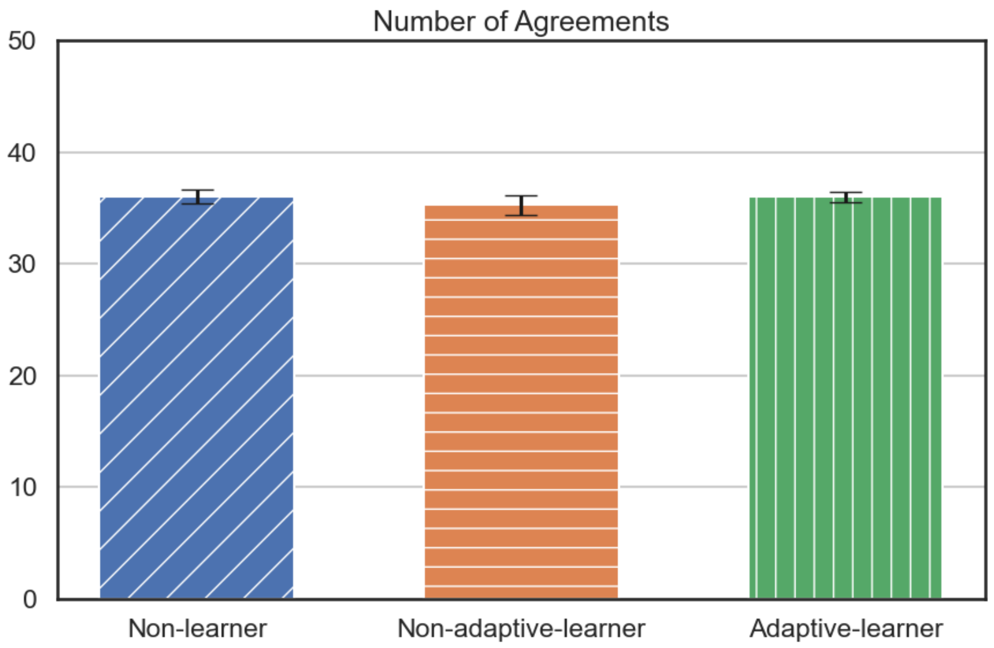
STUDY 2 – UNIFORM PRIOR



- **Adaptive strategy** dominates trust

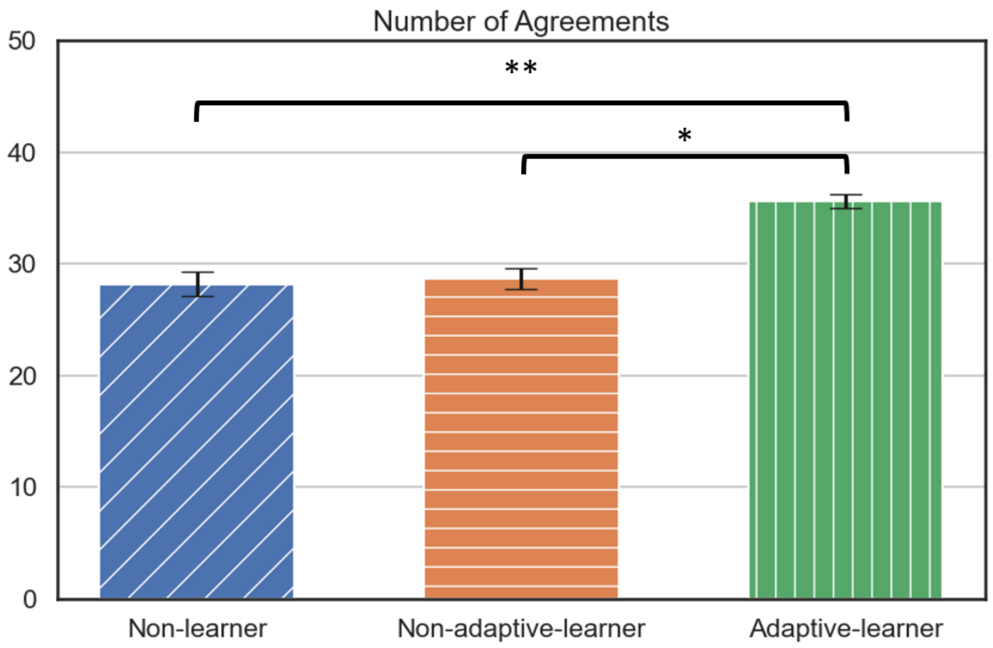
Results – Behavioral Trust

STUDY 1 – INFORMED PRIOR



- **No significant difference** between the three strategies

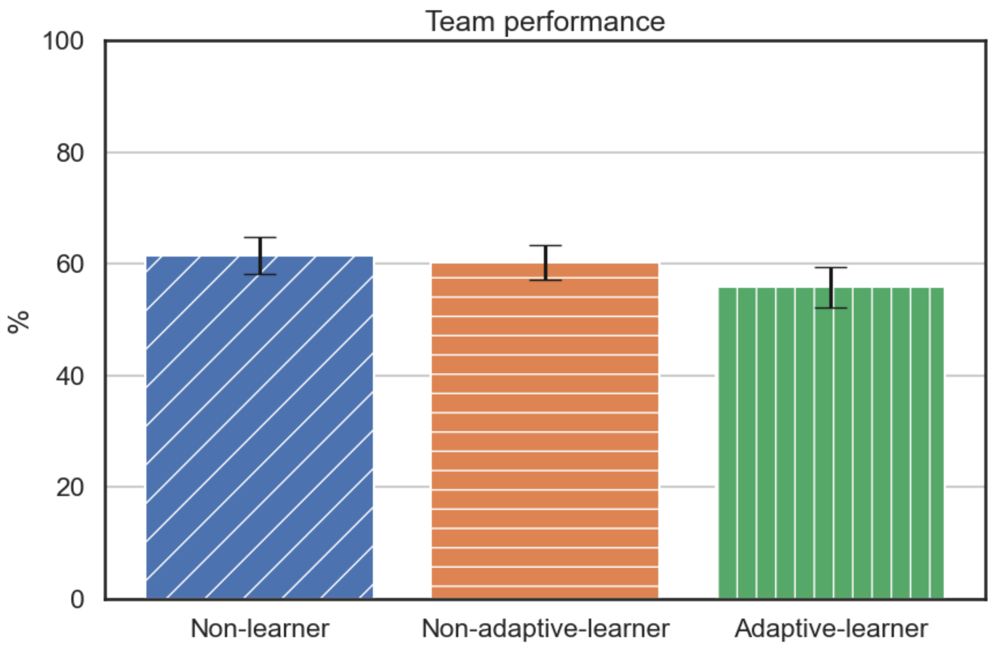
STUDY 2 – UNIFORM PRIOR



- Adaptive strategy is **most agreeable**

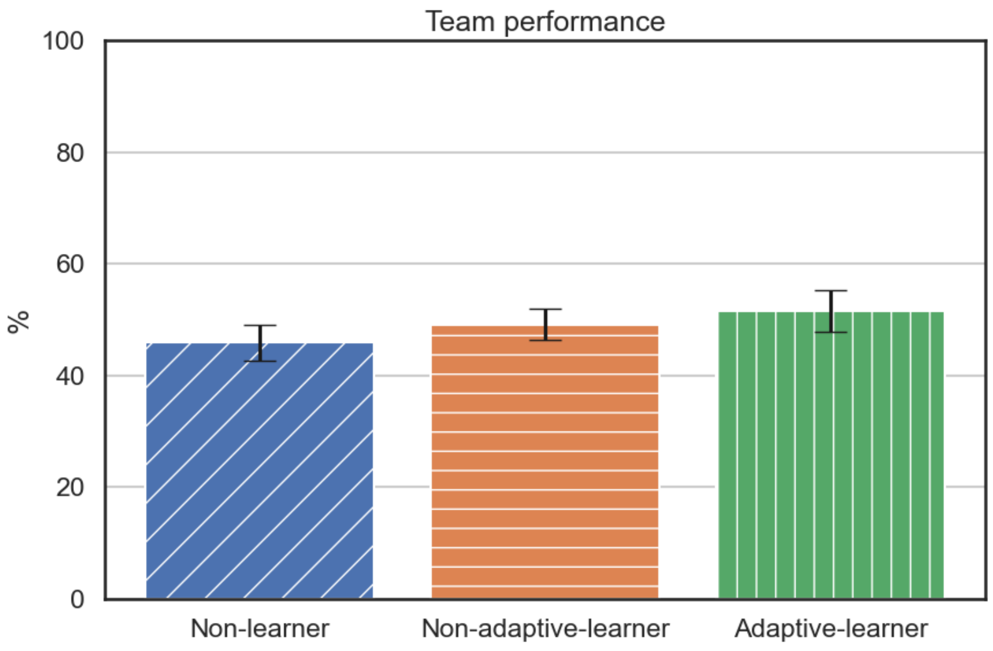
Results – Team Performance

STUDY 1 – INFORMED PRIOR



- **No significant difference** between the three strategies

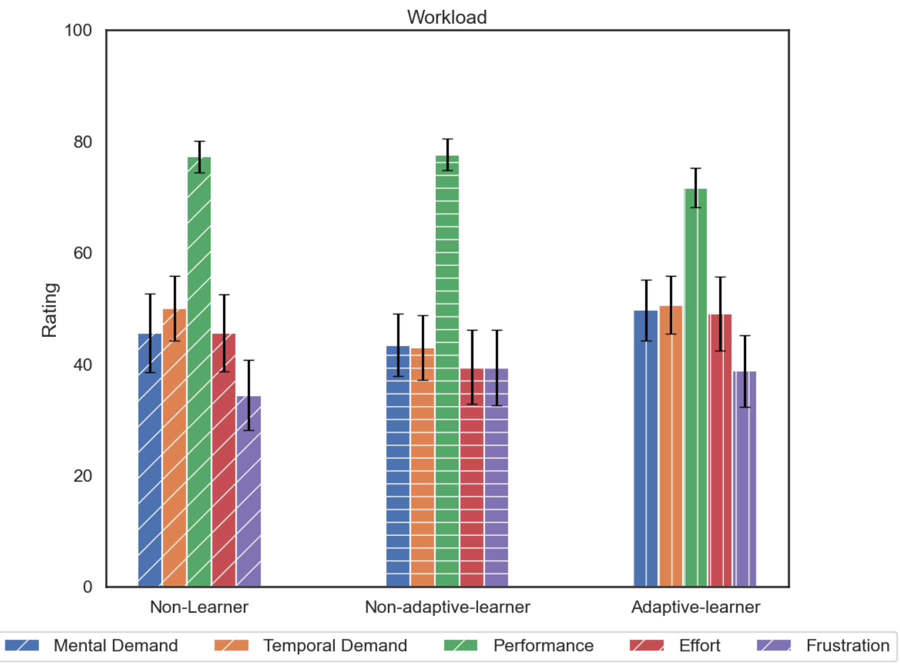
STUDY 2 – UNIFORM PRIOR



- **No significant difference** between the three strategies

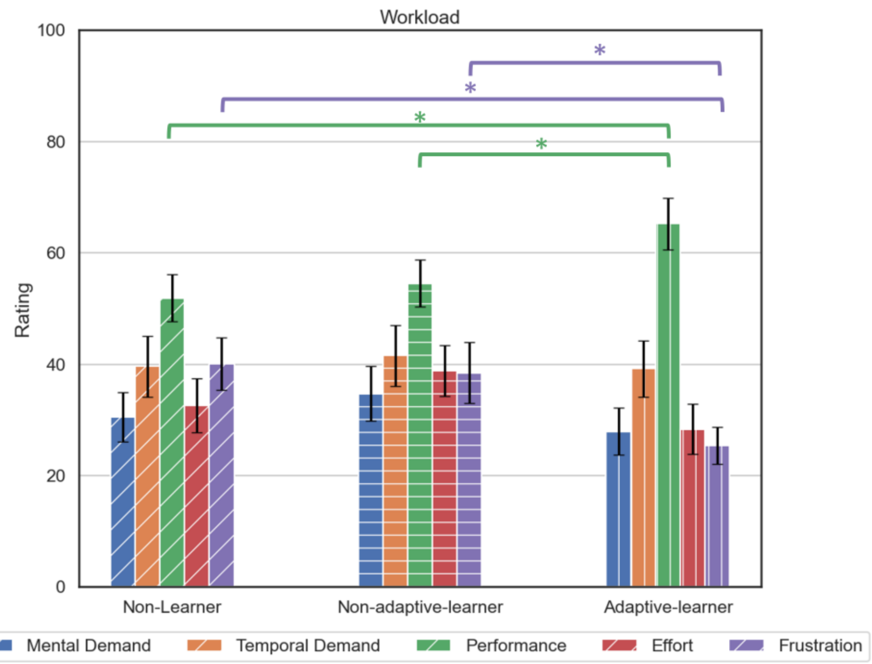
Results – Workload

STUDY 1 – INFORMED PRIOR



- **No significant difference** between the three strategies

STUDY 2 – UNIFORM PRIOR



- **Adaptive strategy** associated with **lowest frustration** and **highest perceived performance**

Phase 2 – Key Takeaways and Limitations

KEY TAKEAWAYS

- Proposed the **Bounded Rationality Disuse Model** of human behavior
- Proposed a **framework for personalized reward learning** using Bayesian IRL
- Personalized reward alignment works better when starting with a uniform prior on reward weights

LIMITATIONS

- **No context dependence** in the reward function
- **Limited exploration** of the health and time contexts
- **Limited variance** in the threat levels presented to the participants

Agenda

- Introduction
- Phase 1 – Trust-Driven Markov Decision Process
- Phase 2 – Effects of real-time personalization of reward weights
- **Phase 3** – Effects of fine-grained reward learning and state space exploration
- Future Research Directions

Reward Function

Phase 2

$$R^o(a_i^h, D_i) = -w^o h(a_i^h, D_i) - (1 - w^o)c(a_i^h) \qquad \text{where, } o \in \{h, r\}$$

Phase 3

Health

$$R(H_i, C_i, a_i^h, D_i) = -w(H_i, C_i)h(a_i^h, D_i) - (1 - w(H_i, C_i))c(a_i^h)$$

Time

- We add the **current health** and **current time** in the state space of the trust-aware MDP
- We **explicitly vary the reward weights** based on the current health and time
- We **do not separate reward functions** for the two agents

The Critical Chance of Threat Presence - d^*

- We see that at a certain chance of threat presence, the two actions result in the same expected reward

- At a chance below d^*, NOT USING the armored robot is better on average

- At a chance above d^*, USING the armored robot is better on average

$$d^*(H,C) = \frac{(1-w(H,C))c}{w(H,C)h}$$

$$\Updownarrow$$

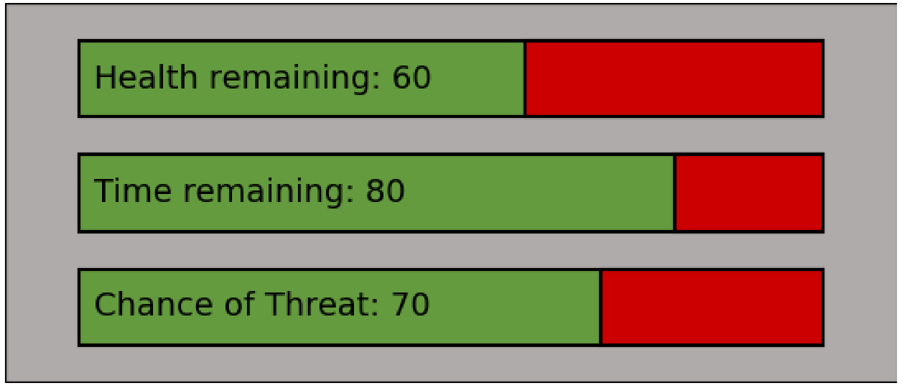
$$w(H,C) = \frac{c}{c + hd^*(H,C)}$$

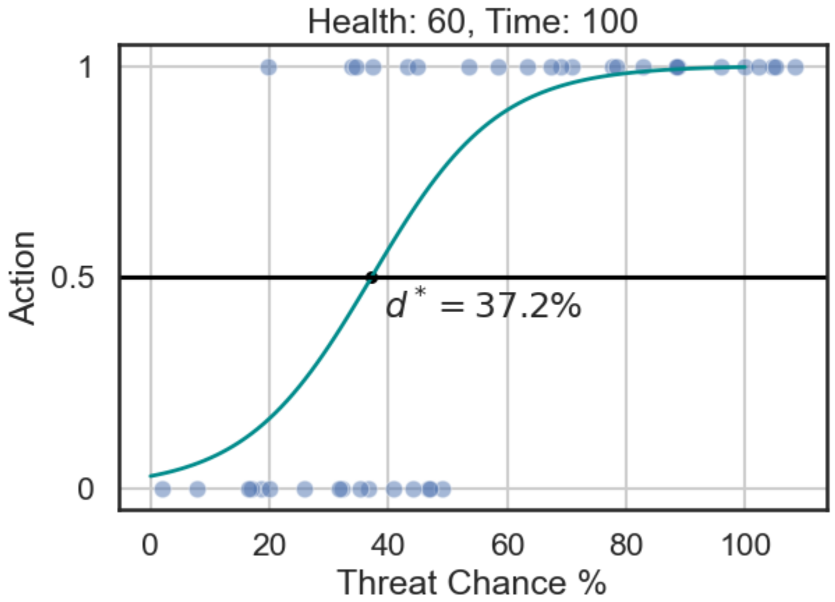
Time Loss Cost

Health Loss Cost

Learning State Dependence of Rewards

- For a set of states $\{H_i, C_i\}_{i \in N}$ get responses from participants about their choice of action for a range of threat levels $d_k \in [0, 100\%]$
- Train logistic regressions for each i
  - The threat level d^* is the threat level at which the classifier gives an equal probability for both actions for the state H_i, C_i
- Data collected via Amazon Mechanical Turk
  - 396 queries (6 health * 6 time * 11 threat levels)
  - 124 workers
  - 4092 responses





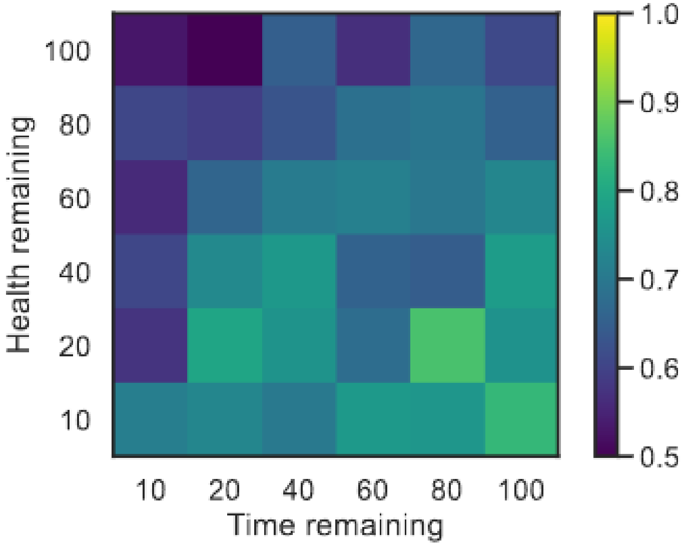
State Dependent Reward Function

- Raw data of learned reward weights is then smoothed by fitting a logistic regression model

- Forward selection using the Akaike Information Criterion (AIC) for selecting features for the final model

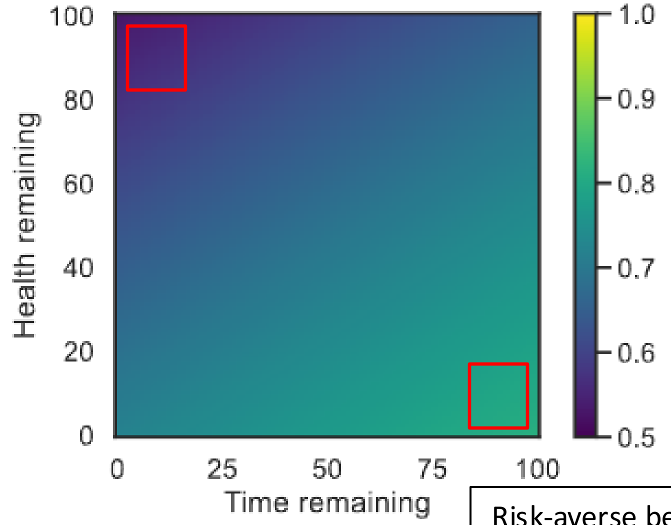
$$w(H, C) = \frac{1}{1 + \exp(0.26H - 0.17C - 0.79)}$$

Learnt Reward weights



Risk taking behavior

Smoothed Reward Weights



Risk-averse behavior

Human-Subject Studies – Interaction Strategies

- We design two interaction strategies for the robot

Constant

- Uses **constant reward weights** associated with the costs of losing health and losing time
- Similar to the **non-learner strategy** from Phase 2, **with the informed prior**
- Chosen as a **baseline** since the non-learner performed as well as the adaptive-learner in the informed prior case in phase 2

State Dependent

- Uses the learned **state-dependent reward function** for the reward weights
- Changes **risk appetite** depending on the **current context** of interaction
- **Goal** is to check if people can identify the **changing risk appetite and prefer it**

Human-Subject Studies – Vulnerability to the Human

- We also consider two conditions of vulnerability for the human

Low Vulnerability

- The human starts with a pool of **100 health points and 100 time points**
- Operationalized as **a high level of armor** and **high available mission time**

High Vulnerability

- The human starts with a pool of **40 health points and 40 time points**
- Operationalized as a **low level of armor** and **low available mission time**

In both conditions, participants lost

- **10 points of time** on using the RARV
- **10 points of health** on encountering threat without the RARV

Active Threat Selection

- What we want to convey?
  - The **constant strategy may be too conservative**
  - The **state dependent strategy changes risk appetite**
- So, we need to be smart about how we set the threats and threat levels
- Threats set
  - **Randomly**, with 50% chance
  - **Actively**, with 50% chance

Random Threat Selection (50%)

$$D \sim \text{Bernoulli}(0.6)$$

$$d \begin{cases} \approx 0.9, & \text{if } D = 1, \\ \approx 0.1, & \text{otherwise.} \end{cases}$$

Active Threat Selection (50%)

$$d_1^*(H, C) = \frac{(1 - w(H, C))c}{w(H, C)h}$$

$$d_2^* = \frac{(1 - w)c}{wh}$$

$$d_2^* < d < d_1^*$$

$$D \sim \text{Bernoulli}(d)$$

Human-Subject Studies – Data Collection

- **2 x 2** mixed factorial design study
  - **Robot strategy** – within-subjects variable
  - **Vulnerability** – between-subjects variable

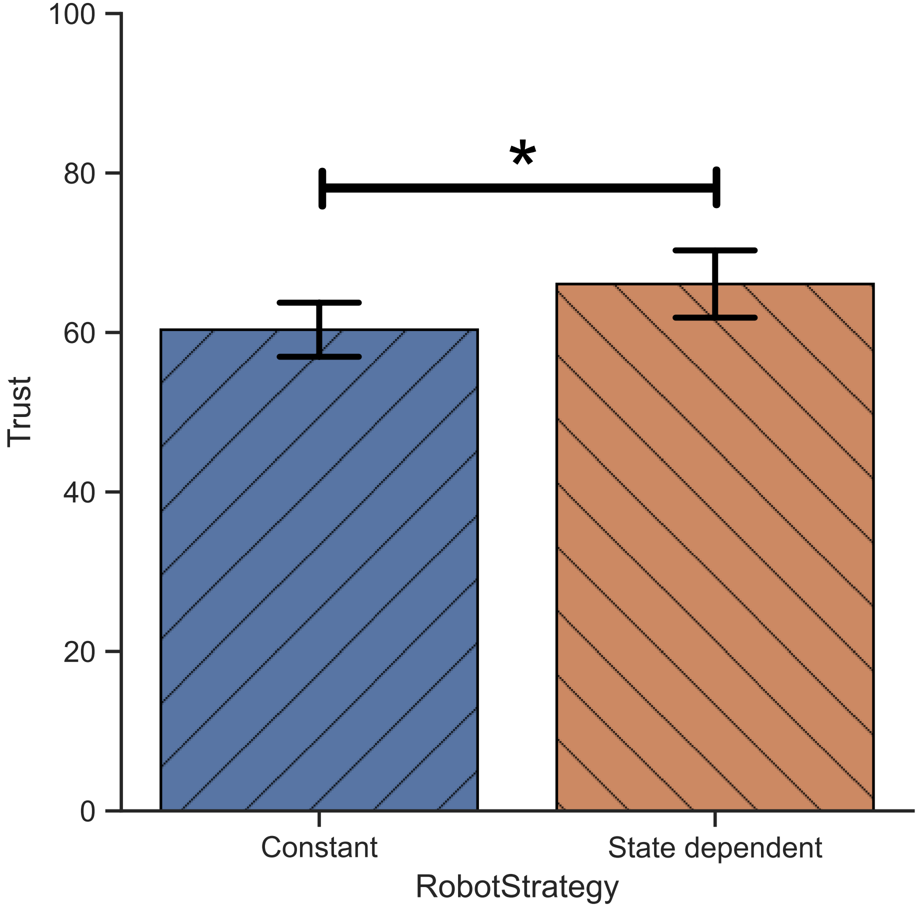
- 40 participants
  - Each participant did **2 missions**
  - Each mission had **10 sequential searches**

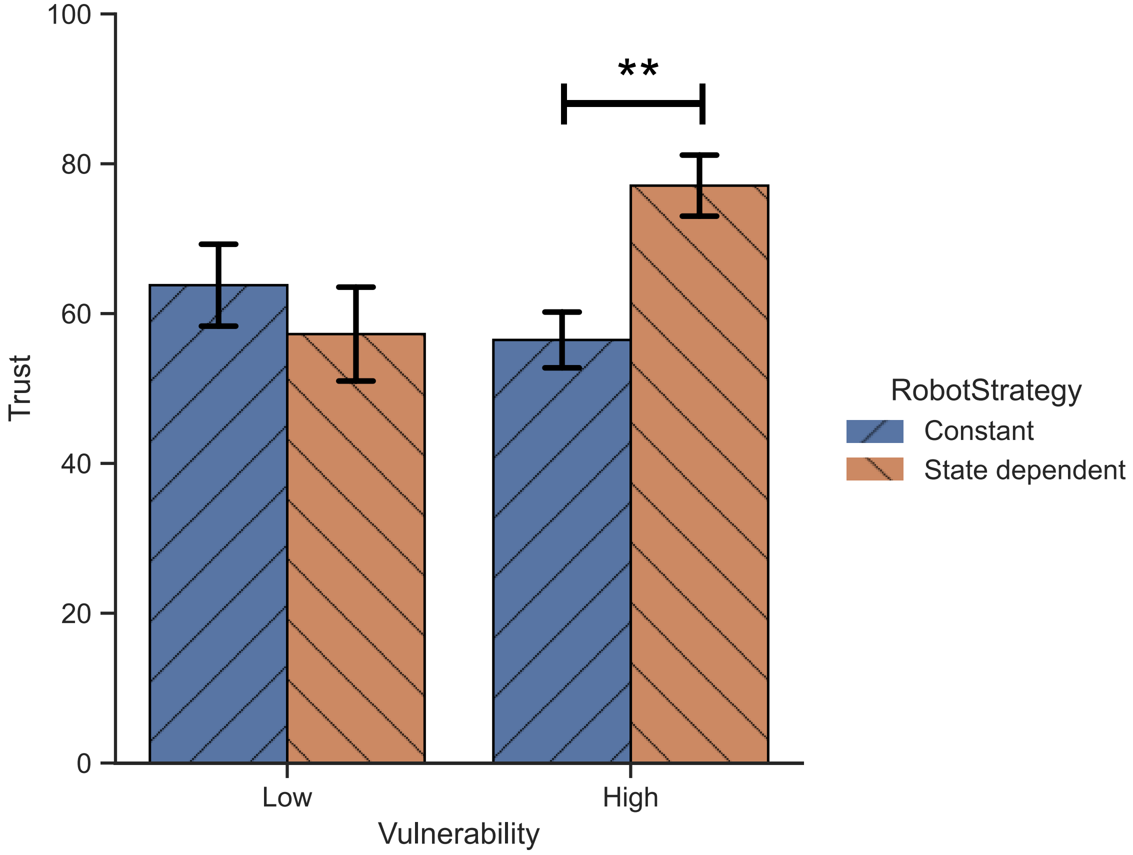
- Removed data from **7 participants** who did not complete the mission
  - 6 from high vulnerability condition

- Final dataset – 33 participants
  - **19 participants** – low vulnerability
  - **14 participants** – high vulnerability

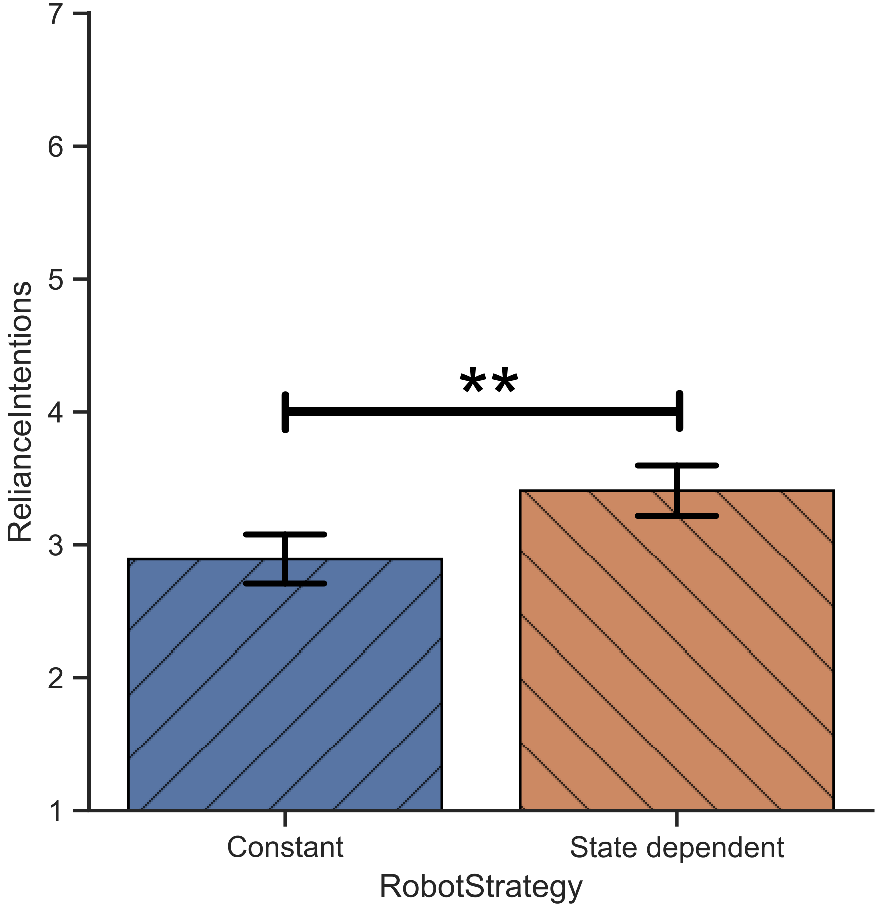
Participants lost
- **10 points of time** on using the RARV
- **10 points of health** on encountering threat without the RARV

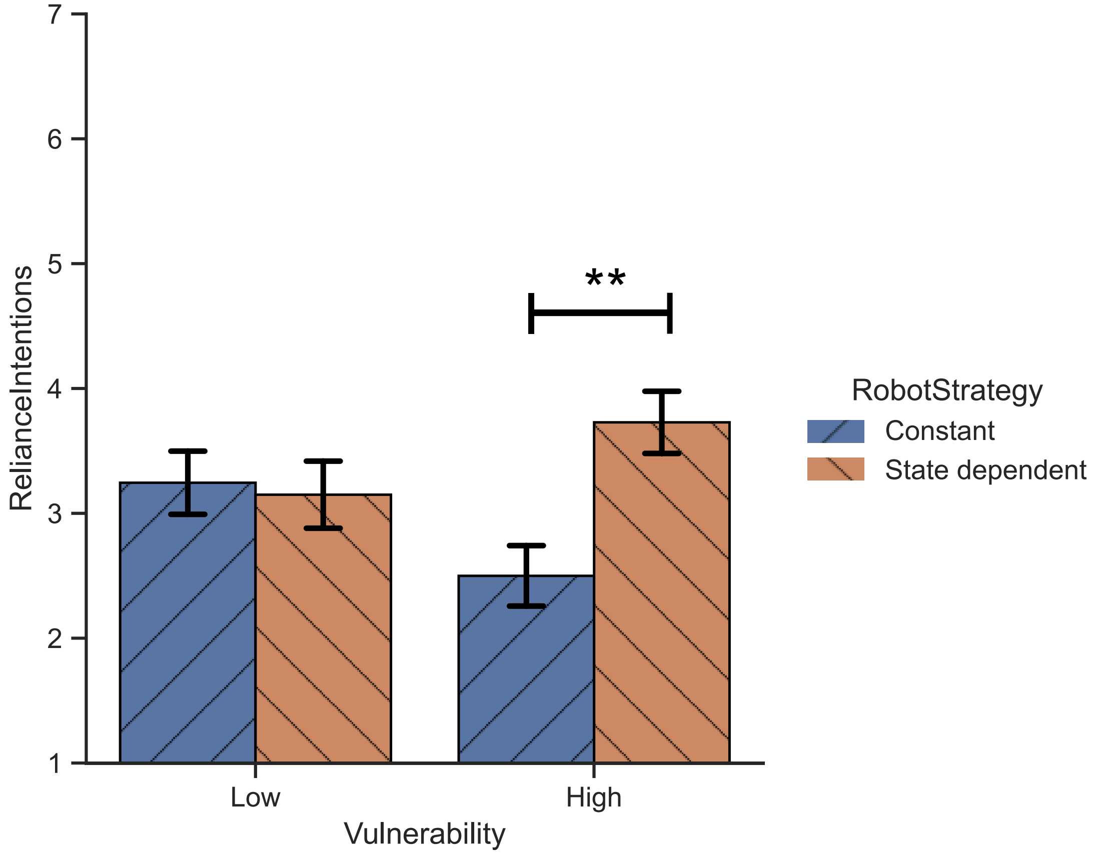
Results – Subjective Trust



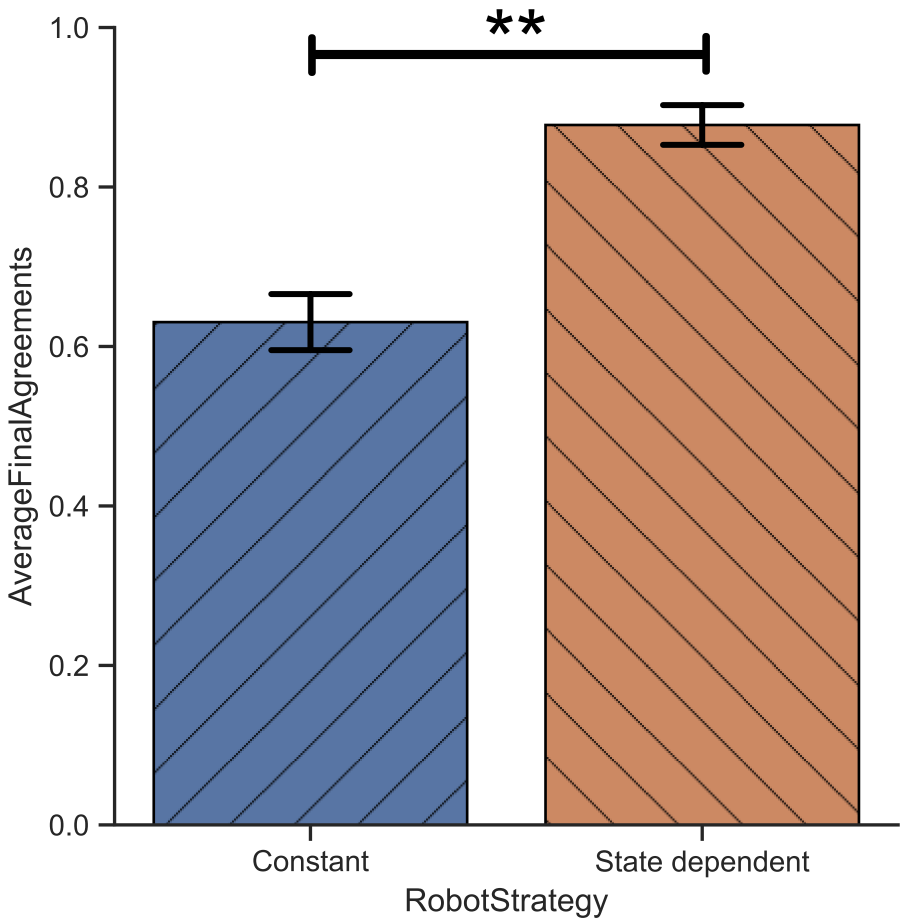


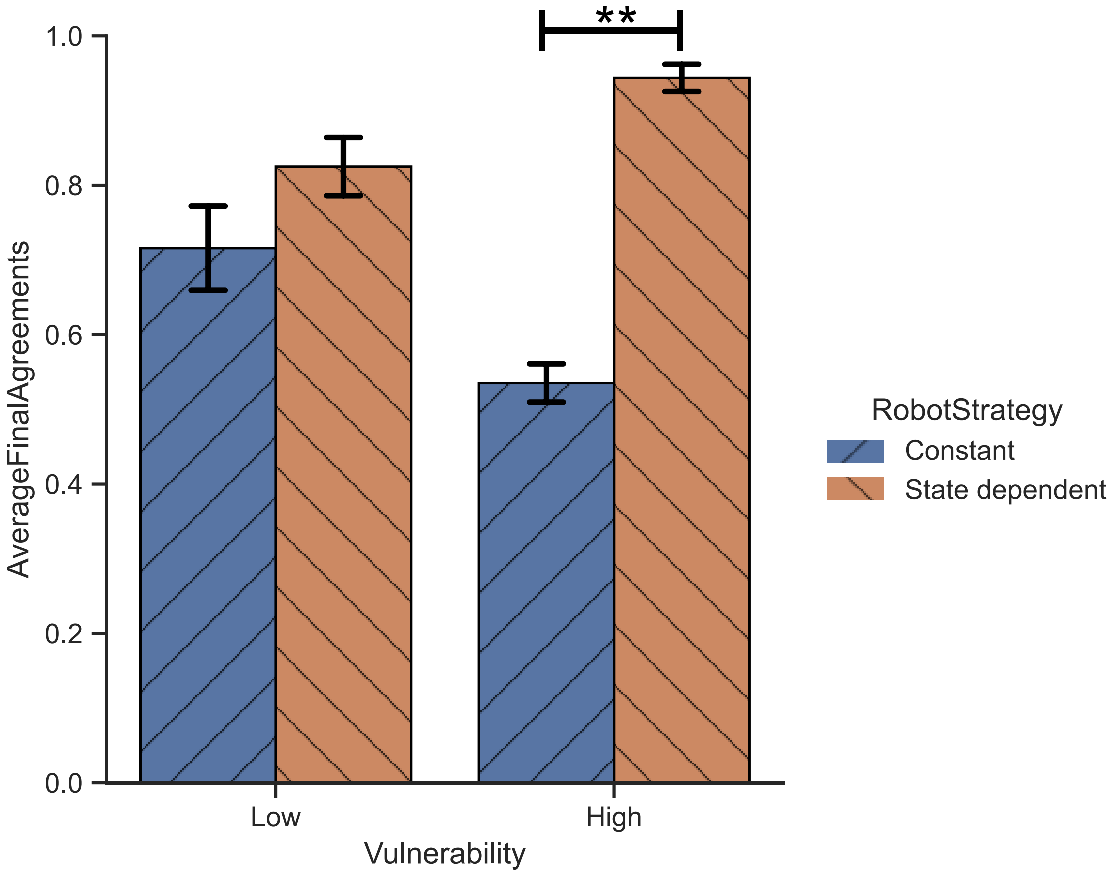
Results – Reliance Intentions



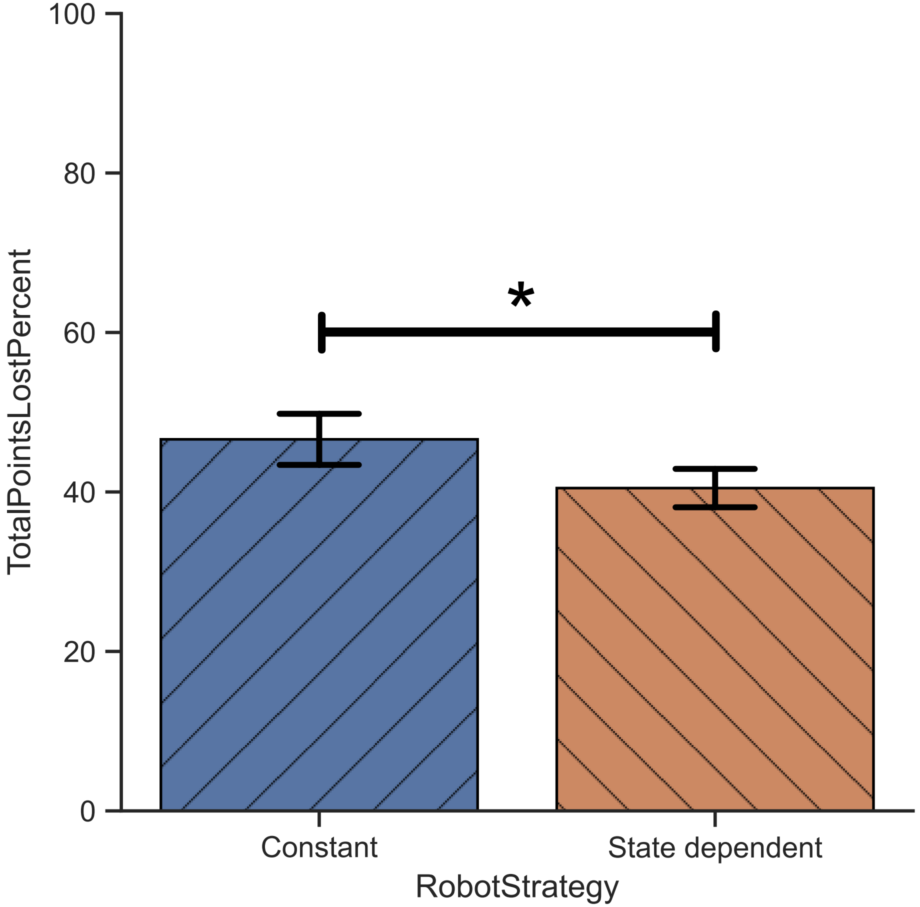


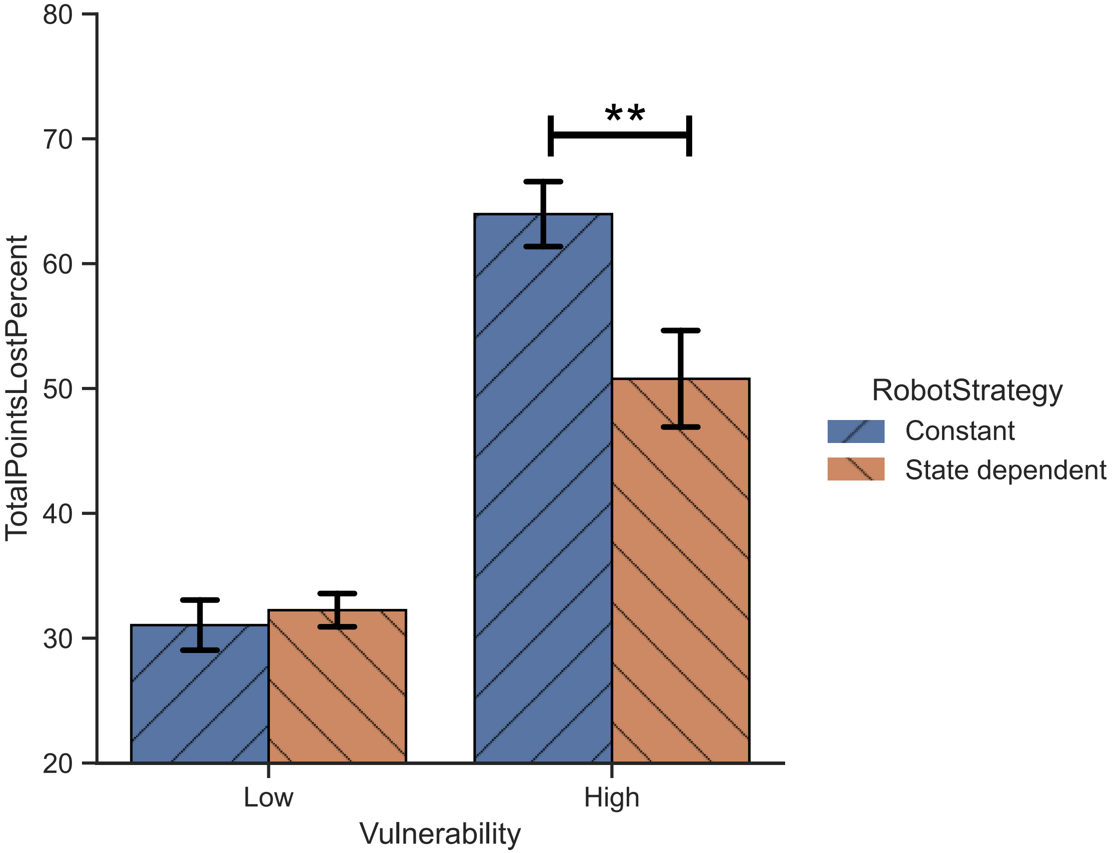
Results – Behavioral Trust





Results – Team Performance





Phase 3 – Key Takeaways and Limitations

KEY TAKEAWAYS

- Proposed a **framework for learning state-dependent reward functions** in an ISR mission context
- Conducted **empirical studies** to learn this reward function and to evaluate its **effects on trust and team performance**
- Fine-grained reward functions are **better for trust and team performance**, especially when the **stakes are high**

LIMITATIONS

- Studies only involved **binary action choices**
- Studies involved dual-objective scenarios in which there is an **obvious bias towards one of the objectives**
- We only considered **dyadic human-robot interaction** scenarios

Agenda

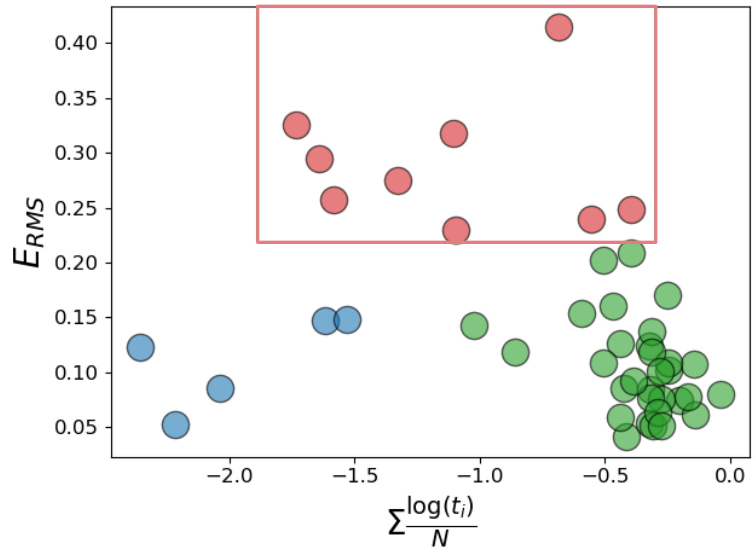
- Introduction
- Phase 1 – Trust-Driven Markov Decision Process
- Phase 2 – Effects of real-time personalization of reward weights
- Phase 3 – Effects of fine-grained reward learning and state space exploration
- Future Research Directions

More Personalized Trust Dynamics Models

- We found that some people exhibit the **Oscillator type trust dynamics**

- The **Beta distribution** trust dynamics model used in this work **struggles to model these dynamics**

- Possible future direction
  - Predict if a person is an oscillator based on their personal characteristics
  - Use a specialized trust dynamics model suited for oscillators when interacting with this person

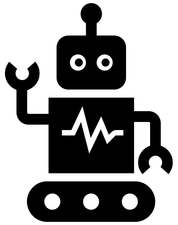


$$\alpha_i = \alpha_0 + \sum_{j=1}^{i} \gamma^{i-j} p_j v^s$$

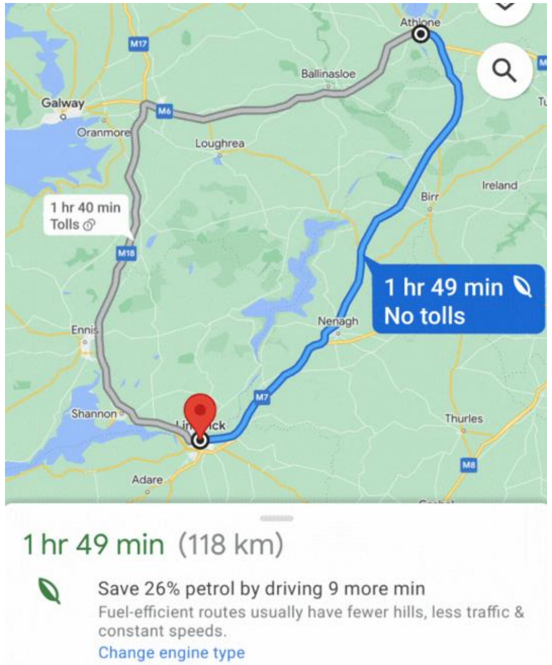
$$\beta_i = \beta_0 + \sum_{j=1}^{i} \gamma^{i-j} (1 - p_j) v^f$$

Exploring Other HRI Domains

- Other types of interactions that would benefit from a trust-driven approach
  - Human-supervisor, robot-worker
  - Human and robot doing separate tasks toward a common goal
  - Human assigning tasks to a robot



- In the ISR mission context, **most humans prefer saving the soldier's health** than saving mission time
  - More research needs to be done to see if our results translate to situations where **human preferences are more varied**
  - E.g. Time vs Quality of Work, Speed vs Eco-friendliness



Incorporating Multiple Actions

- We focused on a **binary action scenario** – USE or NOT USE the RARV
  - Easily define the intuitive **reward-based performance metric**
- In case **multiple actions are available**, humans may
  - Have a non-binary performance metric
  - Exhibit satisficing behavior [13]
  - Something else?
- Studying the **performance metric** could be an interesting direction for future research

Non-binary performance metric

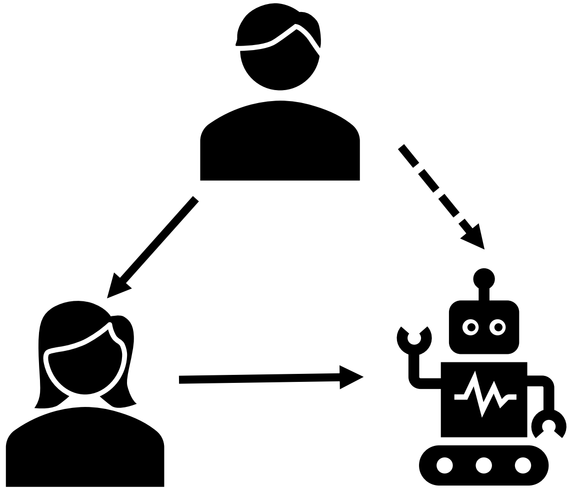
$$p \in (0, 1)$$

Satisficing performance metric

$$p = \begin{cases} 1, & \text{if } R(a^r) > R_t \\ 0, & \text{otherwise} \end{cases}$$

Multi-Human Multi-Robot Scenarios

- We focused on a **dyadic human-robot interaction** scenario
  - Extending this to a **multi-human multi-robot paradigm** brings in an entirely new set of challenges
- **Trust dynamics model**
  - Trust not only evolves through **direct experience** with a robot, but also propagates through **indirect experiences** through another human teammate [14, 15]
- **Task Assignment**
  - How to assign sub-tasks to each dyadic human-robot team?



Thank you!

Air Force Office of Scientific Research

Committee members:

- Dr. Jessie Yang
- Dr. Cong Shi
- Dr. Patricial Alves-Oliveira
- Dr. Brian Denton
- Dr. Joseph Lyons

Questions?

References

1. Endsley, M. R. (2016). From Here to Autonomy: Lessons Learned From Human–Automation Research: Lessons Learned From Human–Automation Research. Human Factors: The Journal of the Human Factors and Ergonomics Society, 59(1), 5-27. https://doi.org/10.1177/0018720816681350 (Original work published 2017)
2. Lee, J. D., & See, K. A. (2004). Trust in Automation: Designing for Appropriate Reliance. Human Factors: The Journal of the Human Factors and Ergonomics Society, 46(1), 50-80. https://doi.org/10.1518/hfes.46.1.50_30392 (Original work published 2004)
3. S. C. Kohn, E. J. De Visser, E. Wiese, Y.-C. Lee, and T. H. Shaw, "Measurement of Trust in Automation: A Narrative Review and Reference Guide," Front. Psychol., vol. 12, p. 604977, Oct. 2021, doi: 10.3389/fpsyg.2021.604977.
4. Guo, Y., Yang, X.J. Modeling and Predicting Trust Dynamics in Human–Robot Teaming: A Bayesian Inference Approach. Int J of Soc Robotics 13, 1899–1909 (2021). https://doi.org/10.1007/s12369-020-00703-3
5. L. Rodriguez Rodriguez, C. E. Bustamante Orellana, E. K. Chiou, L. Huang, N. Cooke, and Y. Kang, "A review of mathematical models of human trust in automation," Front. Neuroergonomics, vol. 4, p. 1171403, Jun. 2023, doi: 10.3389/fnrgo.2023.1171403.
6. Y. Guo, C. Shi and X. J. Yang, "Reverse Psychology in Trust-Aware Human-Robot Interaction," in IEEE Robotics and Automation Letters, vol. 6, no. 3, pp. 4851-4858, July 2021, doi: 10.1109/LRA.2021.3067626
7. Donnellan, M. Brent, et al. "The mini-IPIP scales: tiny-yet-effective measures of the Big Five factors of personality." Psychological assessment 18.2 (2006): 192.
8. Lyons, J. B., & Guznov, S. Y. (2019). Individual differences in human–machine trust: A multi-study look at the perfect automation schema. Theoretical Issues in Ergonomics Science, 20(4), 440-458.
9. Merritt, S. M., Heimbaugh, H., LaChapell, J., & Lee, D. (2013). I trust it, but I don't know why: Effects of implicit attitudes toward automation on trust in an automated system. Human factors, 55(3), 520-534.
10. Muir, B. M., & Moray, N. (1996). Trust in automation. Part II. Experimental studies of trust and human intervention in a process control simulation. Ergonomics, 39(3), 429-460.
11. Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In Advances in psychology (Vol. 52, pp. 139-183). North-Holland.
12. Xuanming Su, (2008) Bounded Rationality in Newsvendor Models. Manufacturing & Service Operations Management 10(4):566-589. https://doi.org/10.1287/msom.1070.0200
13. Simon, H.A. 1955. A behavioral model of rational choice. Quarterly Journal of Economics 69: 99–118. Reprinted in Simon (1982), ch. 7.2.
14. Yaohui Guo, X. Jessie Yang, and Cong Shi. 2023. TIP: A Trust Inference and Propagation Model in Multi-Human Multi-Robot Teams. In Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction (HRI '23). Association for Computing Machinery, New York, NY, USA, 639–643. https://doi.org/10.1145/3568294.3580164
15. Hidalgo, D. T., Guo, Y., Lyons, J. B., Shi, C., & Yang, X. J. (2025). Trust Propagation in Multi-Operator Multi-Autonomy (MOMA) Teams: Friends Versus Strangers. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 0(0). https://doi.org/10.1177/10711813251370736
16. S. Bhat, J. B. Lyons, C. Shi, and X. J. Yang, "Clustering Trust Dynamics in a Human-Robot Sequential Decision-Making Task," IEEE Robot. Autom. Lett., vol. 7, no. 4, pp. 8815–8822, Oct. 2022, doi: 10.1109/LRA.2022.3188902.
17. S. Bhat, J. B. Lyons, C. Shi, and X. J. Yang, "Evaluating the Impact of Personalized Value Alignment in Human-Robot Interaction: Insights into Trust and Team Performance Outcomes," in Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction, Boulder CO USA: ACM, Mar. 2024, pp. 32–41. doi: 10.1145/3610977.3634921.
18. S. Bhat, J. B. Lyons, C. Shi, and X. J. Yang, "Value Alignment and Trust in Human-Robot Interaction: Insights from Simulation and User Study," in Discovering the Frontiers of Human-Robot Interaction, R. Vinjamuri, Ed., Cham: Springer Nature Switzerland, 2024, pp. 39–63. doi: 10.1007/978-3-031-66656-8_3.
19. S. Bhat, J. B. Lyons, C. Shi and X. J. Yang, "Effects of Learning State Dependence of Reward Weights on Trust and Team Performance in a Human-Robot Sequential Decision-Making Task," 2025 IEEE 5th International Conference on Human-Machine Systems (ICHMS), Abu Dhabi, United Arab Emirates, 2025, pp. 35-40, doi: 10.1109/ICHMS65439.2025.11154177.