Evaluating the Impact of Personalized Value Alignment in Human-Robot Interaction: Insights into Trust and Team Performance Outcomes

SHREYAS BHAT, JOSEPH B. LYONS, CONG SHI, X. JESSIE YANG

INDUSTRIAL & OPERATIONS ENGINEERING
UNIVERSITY OF MICHIGAN

UNIVERSITY OF MIAMI
MIAMI HERBERT BUSINESS SCHOOL

AFRL
THE AIR FORCE RESEARCH LABORATORY

Introduction

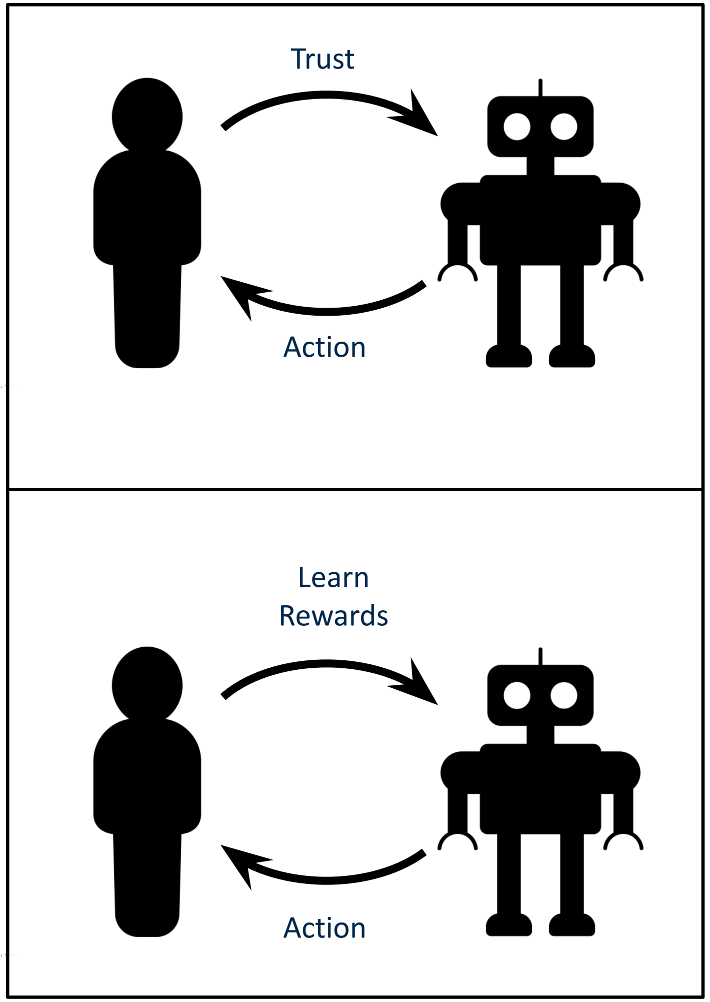
Trust is a key factor to facilitate effective collaboration [1]

Trust has been used to drive the decision-making of robots in human-robot teams [2, 3]

- However, most prior research fixes a reward function for the human and the robot that is known to the team

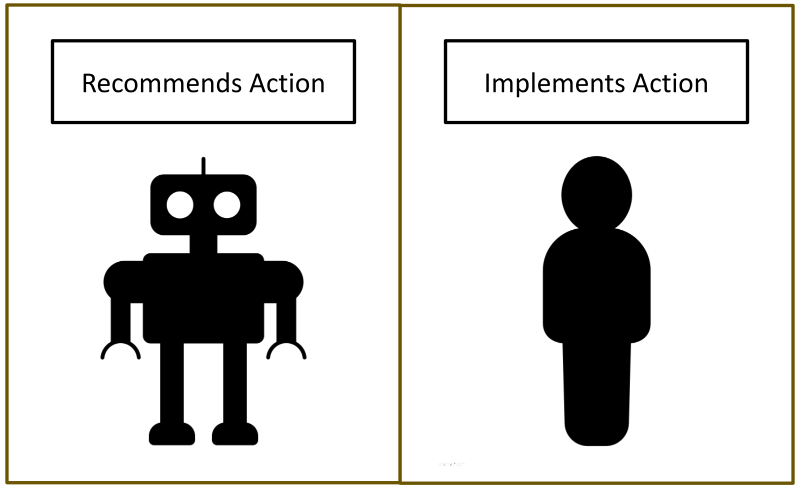
Value alignment [4, 5] refers to the field of study trying to match the “values” of the robot to that of the human

- This is done by learning reward functions for the human through their behavior
- The effect of the degree of value alignment on trust is an unexplored area of research

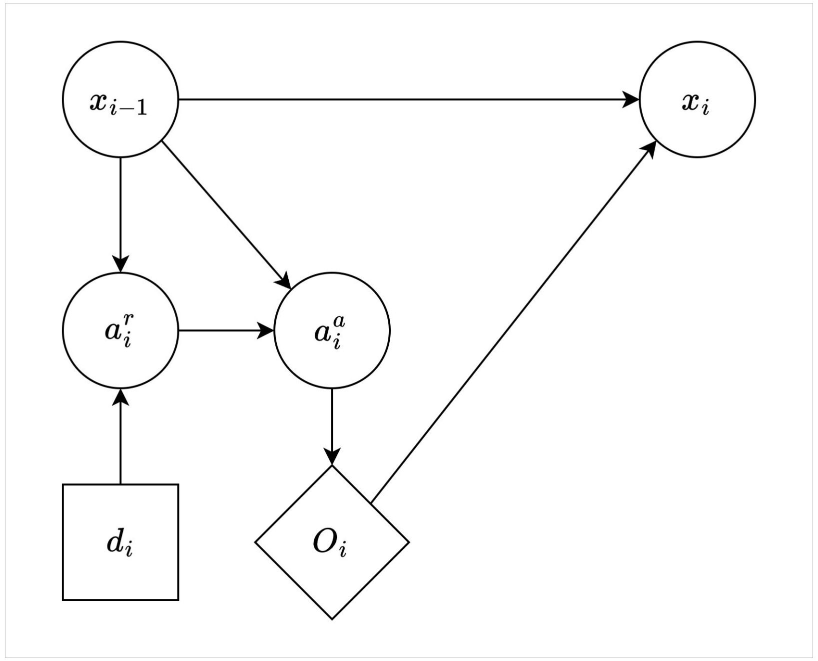


Problem Formulation

- We focus on tasks in which the robot acts an action recommender to the human
  - The human's chosen action is then implemented



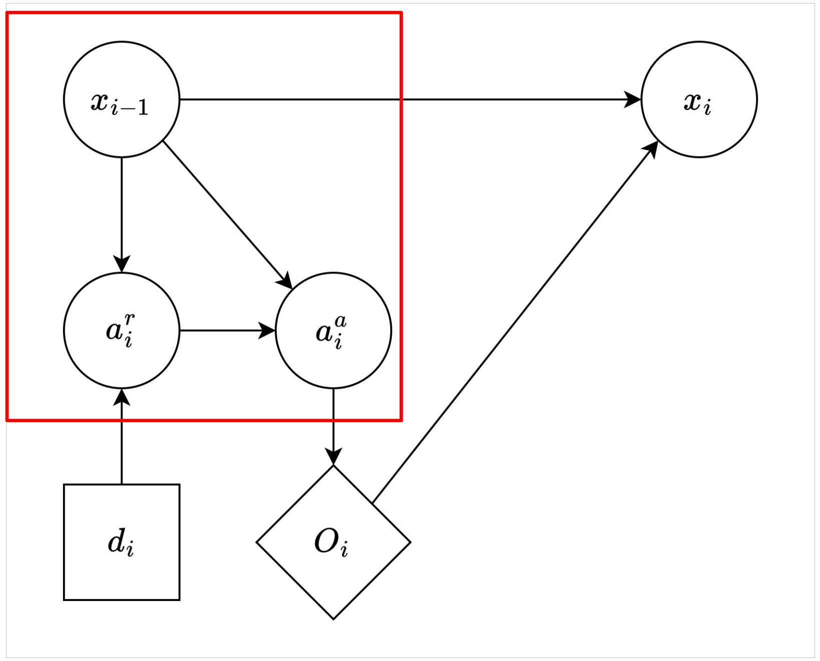
- We model the interaction as a Trust-aware Markov Decision Process
  - **States** – Trust of the human on the robot
  - **Actions** – Actions that can be recommended by the robot
  - **Transition function** – Trust dynamics model
  - **Reward function** – Rewards associated with outcomes observed by choosing actions
  - ***Human behavior model*** *– Probability of the human choosing a certain action given the recommended action and the state*



Human Behavior Model

- Gives the probability of the human choosing an action, given
  - The current state
  - Recommended action

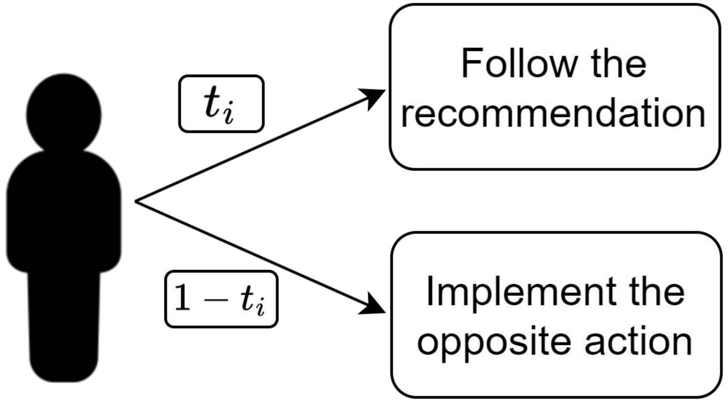
- We combine aspects of two popular models
  - Reverse psychology model [6]
  - Bounded rationality model [5]



$$P(a^h|a^r, x)$$

Human Behavior Model

Reverse Psychology Model [6]

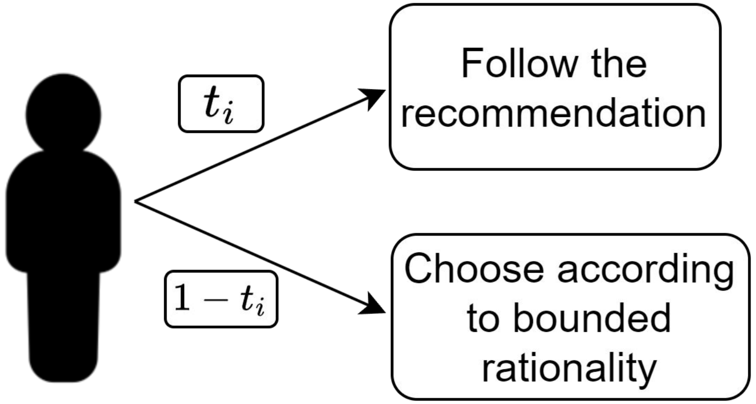


Bounded Rationality Model [5]

$$P(a^h) \propto \exp(E[\kappa R(a^h)])$$

Human Behavior Model

Bounded Rationality Disuse Model



Bounded Rationality Model [5]

$$P(a^h) \propto \exp(E[R(a^h)])$$

Reward Learning

- We assume that the reward function is a weighted sum of task-specific features

$$R(a) = \sum_{i=1}^{D} w_i \phi_i$$

- We maintain belief distributions $b_k(\mathbf{w})$ on the reward weights w_i and update them using Bayes' rule on the human behavior model [7]

$$b_{k+1}(\mathbf{w}) \propto b_k(\mathbf{w}) P(a^h | a^r, t, \mathbf{w})$$

- Note that we need an initial distribution on the reward weights $b_0(\mathbf{w})$ to start the process. The two human subject studies presented here differ in this initial distribution

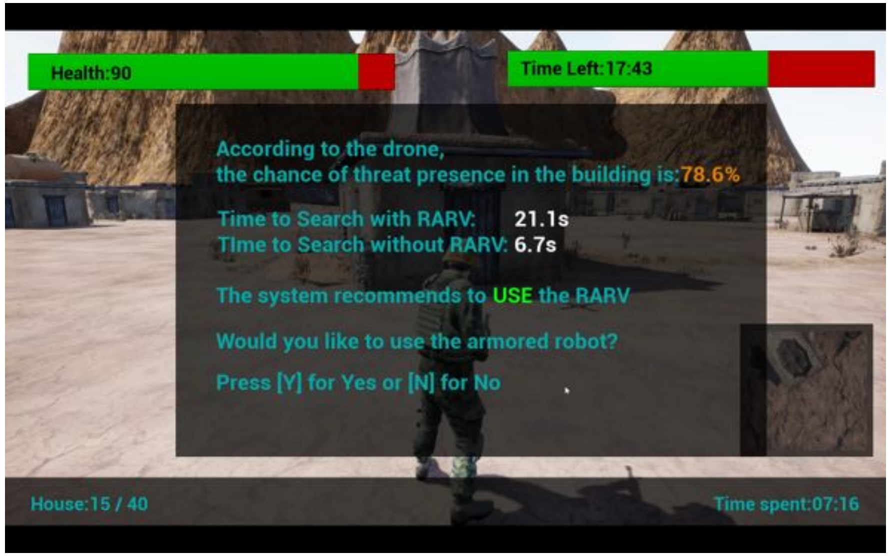
7. Ramachandran and Amir (2007)

Human-Subject Studies

- We designed a reconnaissance mission scenario for our human-subject studies

- Human-robot team searches through a town for potential threats (armed gunmen)

- The robot recommends whether
  - the human should breach the site directly
  - or they should deploy an armored robot for protection



Human-Subject Studies - Conditions

- We design three interaction strategies for the robot

Non-learner:	Assumes that the human shares the robot’s reward function
Non-adaptive-learner:	Learns personalized reward functions for each human. It only uses these for performance estimation and behavior prediction. It still optimizes its original reward function
Adaptive-learner:	Learns personalized reward functions for each human and adopts it as its own

Human-Subject Studies - Details

STUDY 1 – INFORMED PRIOR

- The robot starts its learning algorithm from an informed prior on the reward weights
- 30 participants
- Under this condition, there was not much room for adaptation, leading to minimal differences in the three interaction strategies

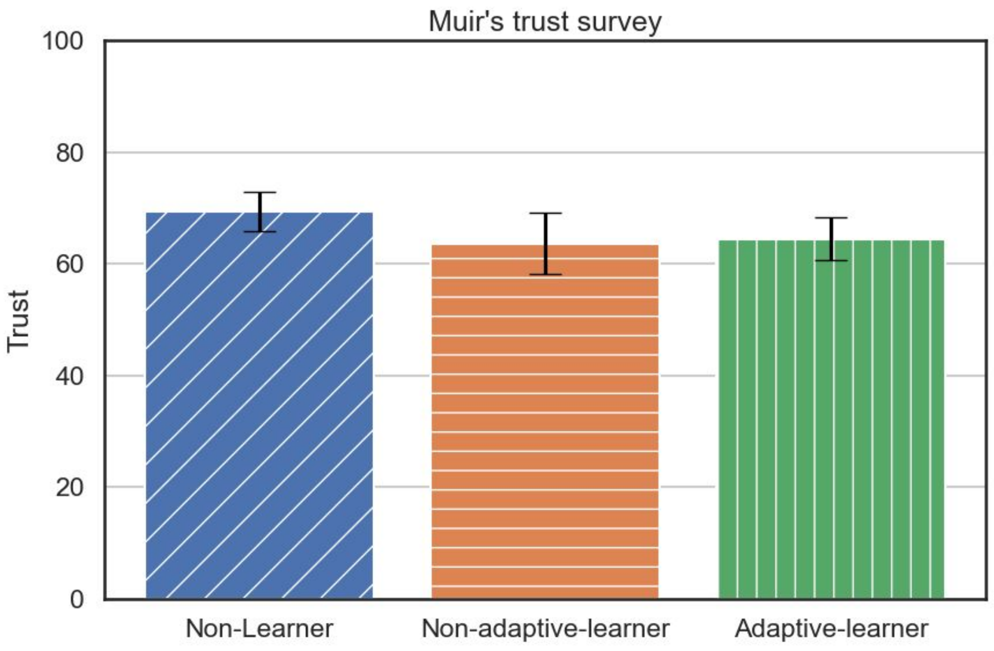
STUDY 2 – UNIFORM PRIOR

- The robot starts its learning algorithm from a uniform prior on the reward weights
- 24 participants
- Under this condition, there was room for adaptation, leading to better performance of the adaptive-learner strategy

In both studies, the non-adaptive-learner and the non-learner strategies use the mean of this prior as the weights for the robot's reward function

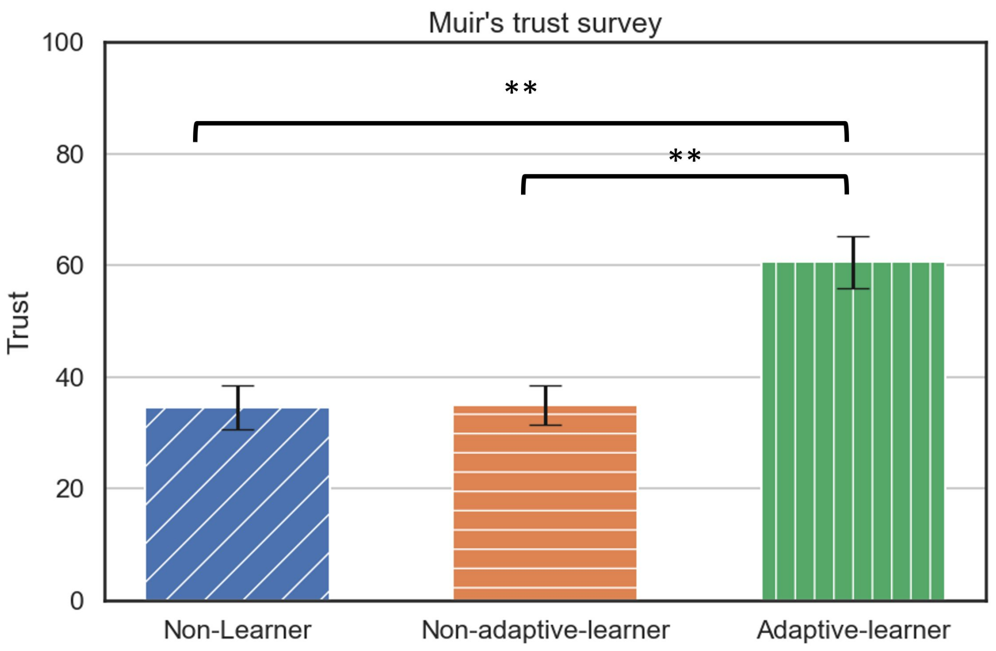
Results – Subjective Trust

STUDY 1 – INFORMED PRIOR



- No significant difference between the three strategies

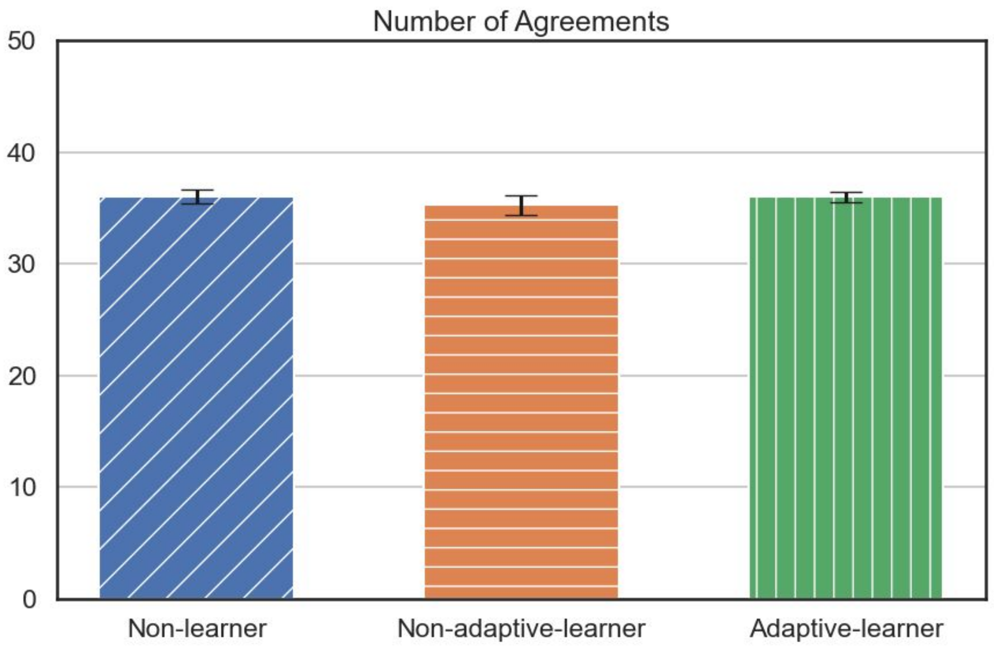
STUDY 2 – UNIFORM PRIOR



- Adaptive strategy rated the highest in trust

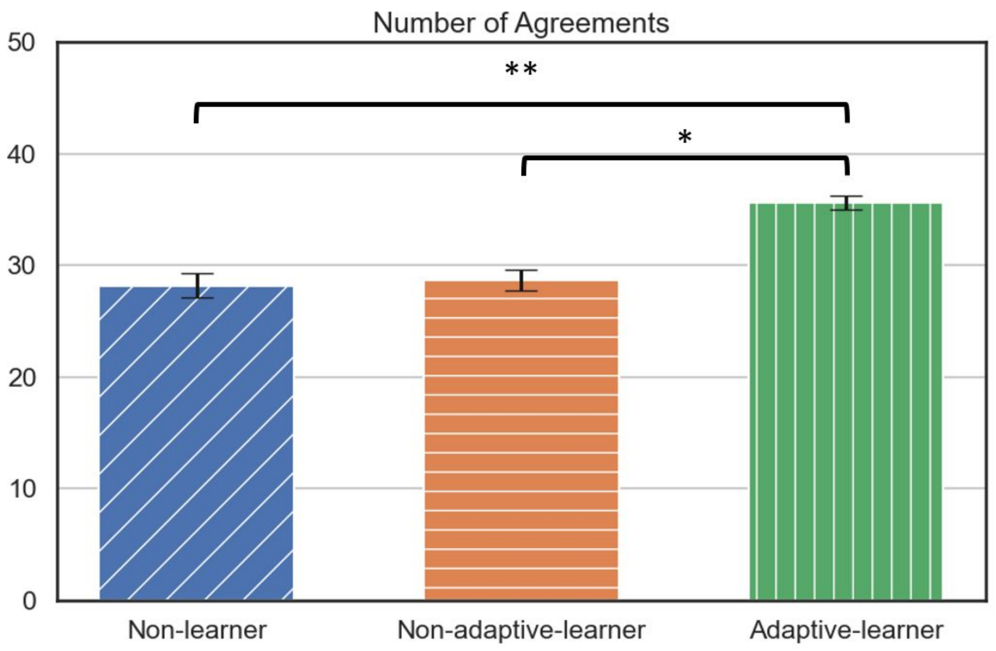
Results – Behavioral Trust

STUDY 1 – INFORMED PRIOR



- No significant difference between the three strategies

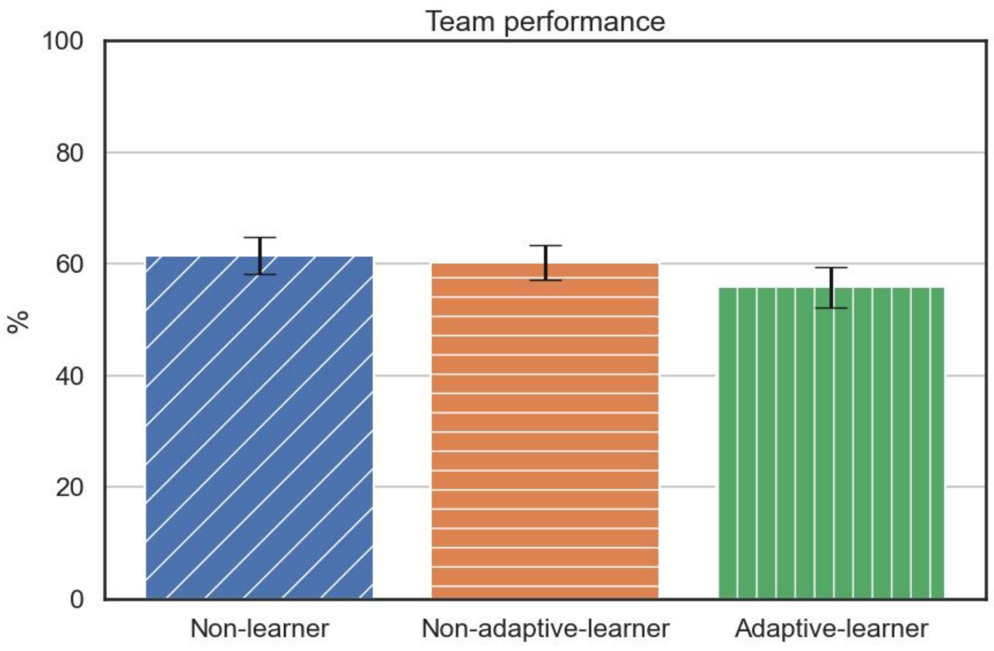
STUDY 2 – UNIFORM PRIOR



- Adaptive strategy had the most agreements with the participants

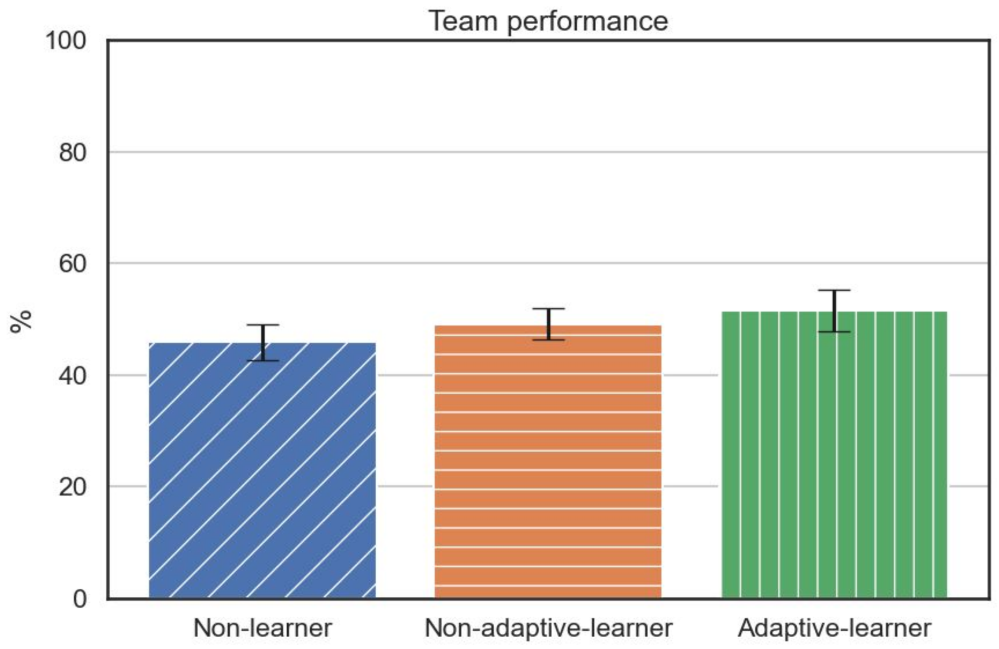
Results – Team Performance

STUDY 1 – INFORMED PRIOR



- No significant difference between the three strategies

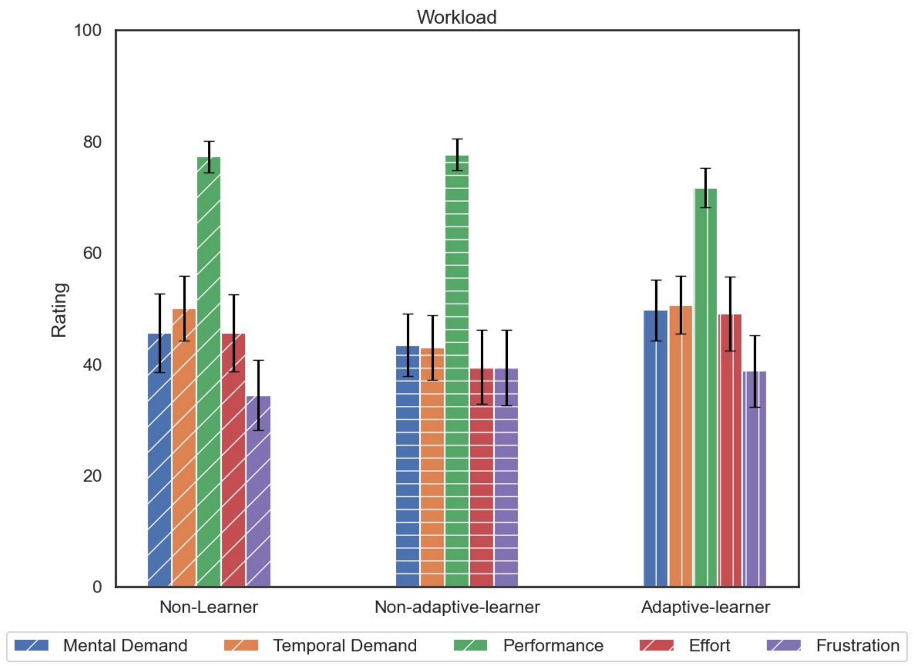
STUDY 2 – UNIFORM PRIOR



- No significant difference between the three strategies

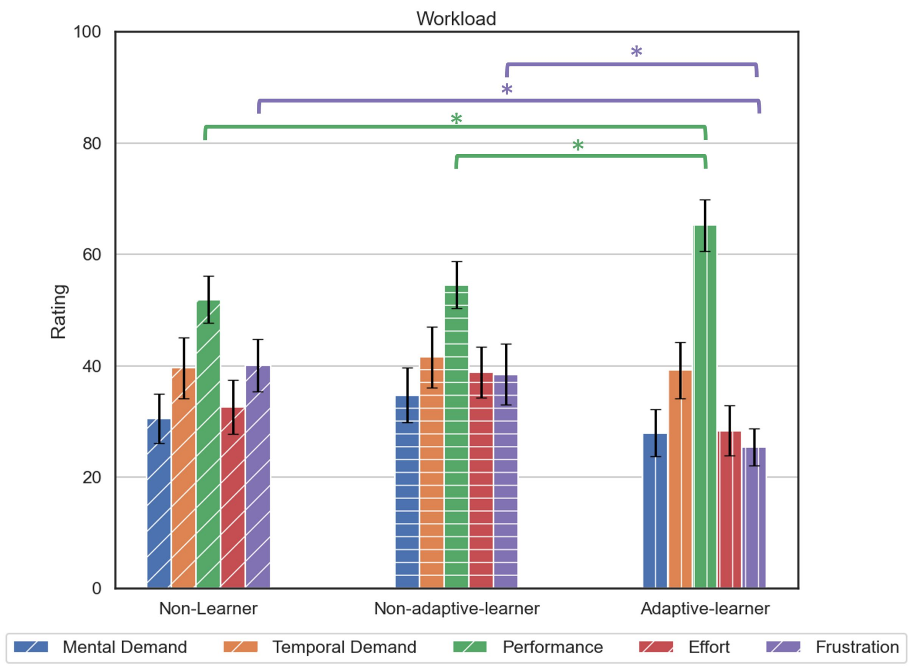
Results – Workload

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

Conclusion

- We present three interaction strategies for a robot in a human-robot team

- We identified conditions where personalized value alignment is beneficial for the team

- Under such conditions, our adaptive-learner strategy results in
  - High trust (subjective and behavioral)
  - Low workload
  - High perceived performance
  - Low frustration

Thank You. Questions?

SHREYAS BHAT

(shreyasb@umich.edu)

Research funded by:

References

1. P. A. Hancock, T. T. Kessler, A. D. Kaplan, J. C. Brill, and J. L. Szalma, “Evolving Trust in Robots: Specification Through Sequential and Comparative Meta-Analyses,” Human Factors, vol. 63, no. 7, pp. 1196–1229, 2020.
2. S. Bhat, J. B. Lyons, C. Shi and X. J. Yang, "Clustering Trust Dynamics in a Human-Robot Sequential Decision-Making Task," in IEEE Robotics and Automation Letters, vol. 7, no. 4, pp. 8815-8822, Oct. 2022, doi: 10.1109/LRA.2022.3188902.
3. Min Chen, Stefanos Nikolaidis, Harold Soh, David Hsu, and Siddhartha Srinivasa. 2020. Trust-Aware Decision Making for Human-Robot Collaboration: Model Learning and Planning. J. Hum.-Robot Interact. 9, 2, Article 9 (June 2020), 23 pages. https://doi.org/10.1145/3359616
4. Luyao Yuan et al., In situ bidirectional human-robot value alignment. Sci. Robot.7,eabm4183(2022).DOI:10.1126/scirobotics.abm4183
5. Fisac, J.F. et al. (2020). Pragmatic-Pedagogic Value Alignment. In: Amato, N., Hager, G., Thomas, S., Torres-Torriti, M. (eds) Robotics Research. Springer Proceedings in Advanced Robotics, vol 10. Springer, Cham. https://doi.org/10.1007/978-3-030-28619-4_7
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. Ramachandran, D., & Amir, E. (2007, January). Bayesian Inverse Reinforcement Learning. In *IJCAI* (Vol. 7, pp. 2586-2591).
8. Bonnie Muir and Neville Moray. 1996. Trust in automation. Part II. Experimental studies of trust and human intervention in a process control simulation. Ergonomics 39 (04 1996), 429–60.
9. Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In Human Mental Workload, Peter A. Hancock and Najmedin Meshkati (Eds.). Advances in Psychology, Vol. 52. North-Holland, 139–183.