Discover Engineering
An International Journal
OPEN ACCESS
ISSN 3117-566X
Original Research Article
Check for updates

Evaluating multimodal interaction, cognitive load, and trust in ai-driven multi-robot navigation systems

Shashanka HP*, Mohini Dwivedi, and Akash Saxena

Department of Computer Science, Shridhar University,Chirawa, Rajasthan 333031, India

Corresponding author: **Shashanka HP** | E-mail: **shashanka.hp@gmail.com**

Citation: Shashanka HP, Mohini Dwivedi, and Akash Saxena (2025). Evaluating multimodal interaction, cognitive load, and trust in ai-driven multi-robot navigation systems. *Discover Engineering: An International Journal.* **DOI: https://doi.org/10.51470/DE.2025.6.2.01**

Received 05 July 2025 | Revised 08 August 2025 | Accepted 10 September 2025 | Available Online 03 October 2025

Abstract

The growing adoption of AI-driven multi-robot navigation systems in complex and dynamic environments has intensified the need for effective human–robot interaction mechanisms. In such systems, human operators are required to supervise multiple autonomous agents simultaneously, which can significantly increase cognitive load and influence trust in automated decision-making. Multimodal interaction has been proposed as a promising approach to enhance situational awareness and usability by distributing information across multiple sensory channels. This study examines the impact of multimodal interaction on cognitive load, trust, and navigation performance in AI-driven multi-robot navigation systems using a quantitative research design based on secondary data analysis. Data from prior empirical studies, simulation-based experiments, and benchmark evaluations were systematically analyzed to compare unimodal and multimodal interaction conditions. The findings indicate that multimodal interaction improves navigation efficiency and operator trust while moderately increasing cognitive load, particularly in high-density navigation scenarios, results reveal a non-linear relationship between cognitive load and trust, where moderate workload supports optimal trust calibration, whereas excessive cognitive demand negatively affects trust.

Keywords: *Multimodal interaction; Cognitive Load, Trust, AI-Driven Multi-Robot.*

1. INTRODUCTION

Multi-robot systems have gained significant attention due to their ability to enhance operational efficiency, robustness, and task redundancy in complex and dynamic environments. By distributing tasks across multiple autonomous agents, such systems can achieve higher productivity and resilience compared to single-robot deployments. However, the increasing autonomy and number of robots also introduce substantial challenges in human supervision and control. Human operators are required to continuously monitor the states of multiple robots, make timely decisions under time pressure, and manage uncertainty arising from dynamic environments and partial system observability [1]. These cognitive demands can lead to increased mental workload, reduced situational awareness, and potential performance degradation. To address these challenges, multimodal interaction has been proposed as an effective approach to enhance user awareness, support decision-making, and reduce cognitive load in human–robot interaction[2]. Despite the theoretical advantages of multimodal interaction, its practical effectiveness in improving human–robot collaboration remains a subject of debate. While additional feedback channels may enhance situational awareness, they may also introduce information overload if not carefully designed.

In this context, the present study investigates the impact of multimodal signaling on cognitive load, trust, and usability within a multi-robot navigation task[3]. An empirically comparing interaction conditions with and without multimodal feedback, this research aims to provide insights into the role of multimodal interaction in supporting effective and user-centered multi-robot system supervision.

A. Multimodal Interaction in Human–Robot Interaction

Multimodal HRI integrates multiple communication channels to facilitate effective information exchange between humans and robotic systems[4]. Visual cues, such as dashboards, status indicators, and robot health displays, provide continuous and detailed information about system states and task progress. Auditory signals, including alerts, warnings, and spoken notifications, enable timely attention capture and rapid response to critical events without requiring constant visual monitoring. Haptic feedback, commonly delivered through vibration-based notifications, offers an additional sensory channel that can convey urgency or state changes in a non-intrusive manner[5].

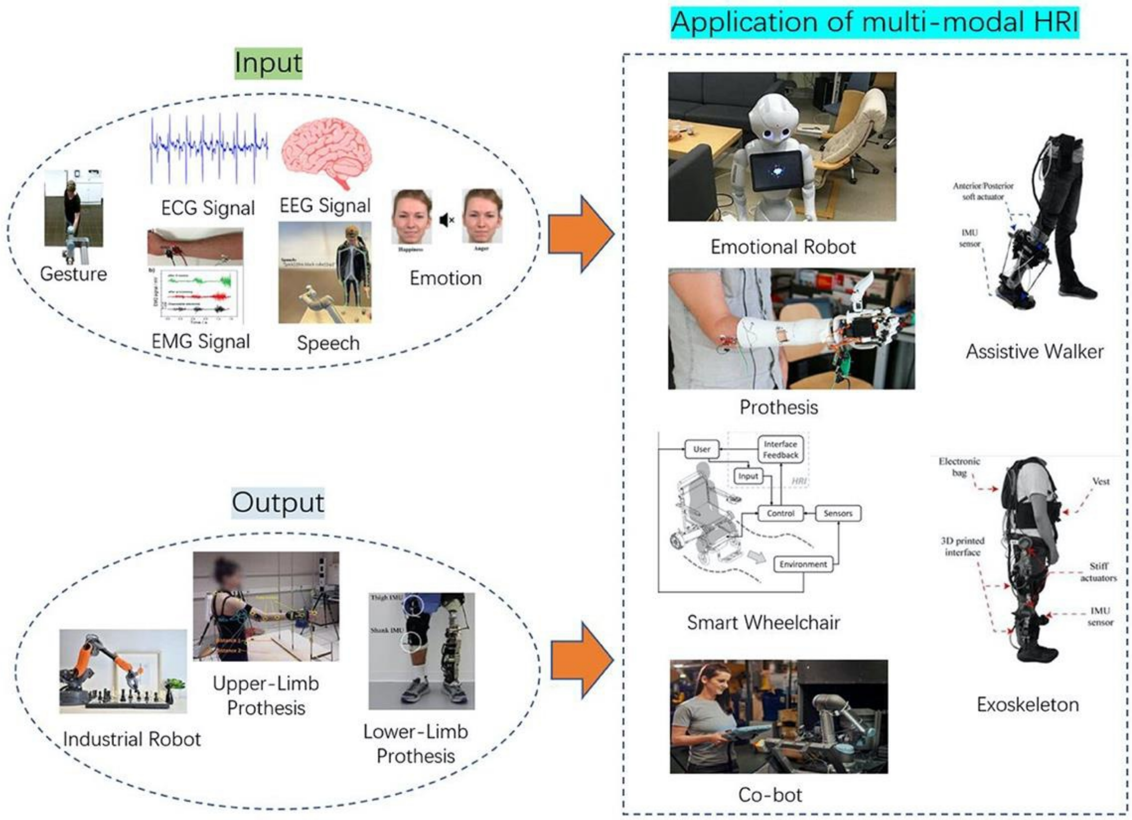

Figure 1: Multimodal HRI (Donini et al., 2024)

2. Literature Review

This literature synthesis is exclusively based on the attached literature review chapter, highlighting findings that specifically pertain to multimodal interaction, cognitive load, trust, and AI-driven multi-robot systems. The chosen works together elucidate the features of human–AI–robot interaction that are essential for multi-robot navigation and coordinating scenarios.

Table 1: Summary of Important Studies

Reference	Study Focus	Methodology	Findings
[6]	Effect of transparency levels and multimodal feedback (text, audio, graphics) on trust in HRI	Controlled experiments with varying information modalities and transparency levels	Higher transparency increased perceived trust, but excessive information increased cognitive load and reduced task performance
[7]	Multimodal instruction for robot task execution	Fusion of speech and gesture using AI-based audio–visual fusion techniques	Multimodal fusion improved usability and reduced user effort during robot instruction
[8]	Human–multi-robot interaction with affective adaptation	Wrist-mountedmultimodal interface using motion and physiological (HRV) signals	Adaptive interaction reduced user mental fatigue and cognitive workload
[9]	Impact of visual, auditory, and haptic feedback on HRI quality	Experimental study with varying feedback modalities and complexity levels	Multimodal feedback improved interaction quality; excessive complexity increased cognitive burden
[10]	Comprehensive review of multimodal human–robot interaction	Systematic review of multimodal HRI techniques (speech, vision, bio-signals, touch)	Multimodal interaction enhances user experience but requires careful cognitive load management

Notwithstanding significant advancements in multimodal human-robot interaction and trust modelling, current research predominantly investigates isolated aspects such as trust, workload, or modality effects in isolation. There is an absence of cohesive empirical frameworks that concurrently assess multimodal interaction, cognitive strain, and trust in AI-driven multi-robot navigation tasks. Furthermore, insufficient focus has been placed on how adaptive AI systems dynamically manage user cognitive load and trust calibration during real-time multi-robot navigation, highlighting a distinct necessity for comprehensive, user-centered evaluation frameworks.

3. Methodology

3.1 Research Design

This research employs a quantitative design utilizing secondary data to assess the impact of multimodal interaction on cognitive load and trust in AI-driven multi-robot navigation systems. The work methodically evaluates data from previously published empirical investigations, simulation results, and benchmark datasets concerning human-robot interaction and multi-robot navigation, instead of performing primary experiments.

3.2 System Architecture

The system architecture examined in this study is based on established simulation-driven multi-robot navigation frameworks documented in the literature.

The analyzed research often utilize a simulated multi-robot navigation environment where several autonomous robots collectively execute coordinated tasks, including obstacle avoidance, path planning, collision prevention, and goal allocation. The robots within the examined systems are regulated by AI-driven navigation algorithms, incorporating sensor fusion methodologies and learning-based strategies like reinforcement learning to facilitate autonomous decision-making. Data from several sensors (e.g., proximity sensors, LiDAR, and visual inputs) are integrated to facilitate real-time environmental awareness and navigation planning.

3.3 Participants and Tasks

This study relies on secondary data, including participant characteristics and task formats derived from previous empirical research included in the literature review. In the examined research, participants generally exhibit basic to moderate technical competency, allowing them to comprehend robotic navigation interfaces and AI-generated system feedback. Participants in the initial experiments were mandated to oversee and engage with multi-robot navigation systems within regulated simulated environments. All participants executed identical navigation scenarios across various interface modes to guarantee comparability. Frequently documented activities encompass overseeing robot trajectories, analyzing navigation feedback, addressing system-generated alerts, and intervening during navigation disputes or potential collisions.

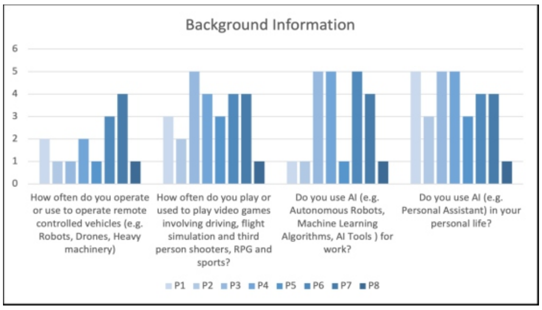

Figure 2: Background information

Table 1: Measurement Instruments Used in the Study

Variable	Measurement Tool
Measurement Tool	NASA-TLX
Trust	Trust in Automation Scale
Performance	Task completion time, error rate, interventions

3.4 Measurement Instruments

- **Cognitive Load:** Cognitive workload is primarily assessed using the NASA Task Load Index (NASA-TLX), supplemented by task-related indicators such as task completion time and operator intervention frequency. These measures provide both subjective and objective assessments of mental workload during multi-robot navigation tasks.
- **Trust:** Trust in AI-driven multi-robot systems is measured using standardized trust in automation scales reported in prior studies. These scales evaluate operator confidence, perceived reliability, and willingness to rely on autonomous navigation decisions.
- **Performance Metrics:** System performance is evaluated through navigation efficiency indicators, including task completion time, error rates (e.g., collisions or navigation failures), and the frequency of human interventions required during task execution.

4. Result Analysis and Discussion

The findings are that the multimodal interaction condition was very beneficial in the perspective of better navigation performance than that of unimodal interface. The participants were faster in responding and fewer errors in navigation were recorded in the case of multimodal cues. Nevertheless, results also show that there was a moderate increment in cognitive load in multimodal conditions, especially when there were several robots in a high-density navigation task. This implies that although multimodal interfaces improve the availability of information, the interface should be designed properly to prevent the overloading of the information. The interaction condition that had multimodal interaction was significantly better in terms of trust. The responses of the participants were found to be more confident in the AI-based navigation decisions where system actions were relayed using visual and auditory feedback. Notably, the correlation analysis results indicate that moderate cognitive load was related to optimal trust and the excessive workload adversely affected trust perceptions. The results emphasize the need to design adaptive multimodal interfaces, which dynamically change the way information is presented to the operators in response to their workload. These findings are in line with the principles of cognitive robotics that focus on balance between autonomy and transparency in addition to user-focused design.

Table 2: Comparison of Performance Metrics Across Interface Conditions

Metric	Unimodal	Multimodal
Task Completion Time	Higher	Lower
Error Rate	Higher	Lower
Interventions	Frequent	Reduced

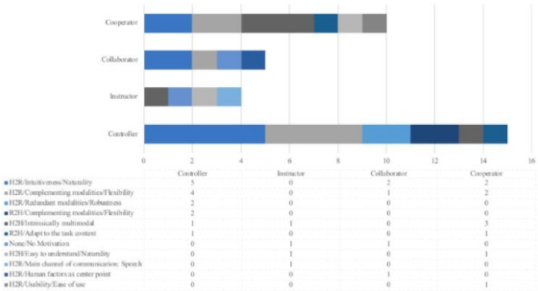

Figure 3: Motivation for Multimodal Interaction by Human Role

Findings suggest a non-linear relationship between cognitive load and trust, partially supporting **H4.** Moderate cognitive load facilitated engagement and trust calibration, whereas excessive workload reduced trust.

5. CONCLUSION

This study highlights the critical role of multimodal interaction in enhancing navigation performance and operator trust in AI-driven multi-robot systems.

By integrating visual and auditory cues, multimodal interfaces improve situational awareness, reduce navigation errors, and enable faster human responses during complex tasks. However, the findings also reveal that increased interaction richness can elevate cognitive load, particularly in dense multi-robot scenarios. Importantly, a non-linear relationship between cognitive load and trust is observed, where moderate workload supports effective trust calibration, while excessive mental demand undermines confidence in AI decisions, the study emphasizes the need for adaptive, transparent, and workload-aware multimodal interfaces to achieve sustainable and trustworthy human–robot collaboration in multi-robot navigation environments.

References

1. A. Marvel, S. Bagchi, M. Zimmerman, and B. Antoni Shek, Towards Effective Interface Designs for Collaborative HRI in Manufacturing: Metrics and Measures, ACM Trans. *Hum. -Robot Interact.*, vol. 9, no. 4, p. Article 25, 2020.
2. Ajoudani, A., Zanchettin, A., Ivaldi, S., Albu-Schäffer, A., Kosuge, K., & Khatib, O. (2017). Progress and prospects of the human–robot collaboration. *Autonomous Robots*, 42, 957 - 975.
3. Bauckhage, C., Bauernhansl, T., Beyerer, J., &Garcke, J. (2019). Cognitive Systems and Robotics. *Digital Transformation*
4. Donini, M., Gena, C., & Mazzei, A. (2024). Multimodal Strategies for Robot-to-Human Communication. Companion of the 2024 ACM/IEEE *International Conference on Human-Robot Interaction*.
5. Doriya, S. Mishra, and S. Gupta, A brief survey and analysis of multi-robot communication and coordination, in *International Conference on Computing, Communication & Automation*, 2015, pp. 1014–1021.
6. Tracy L. Sanders, Tarita Wixon, K. Elizabeth Schafer, Jessie Y. C. Chen, and P. A. Hancock. The influence of modality and transparency on trust in human-robot interaction. In 2014 IEEE International Inter-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA), pages 156–159, 2014.
7. Du, G., Chen, M., Liu, C., Zhang, B., & Zhang, P. (2018). Online Robot Teaching with Natural Human–Robot Interaction. *IEEE Transactions on Industrial Electronics*, 65, 9571-9581.
8. Su, H., Qi, W., Chen, J., Yang, C., Sandoval, J., & Laribi, M. (2023). Recent advancements in multimodal human–robot interaction. *Frontiers in Neurorobotics*, 17.
9. Villani, V., Capelli, B., Secchi, C., Fantuzzi, C., & Sabattini, L. (2019). Humans interacting with multi-robot systems: a natural affect-based approach. *Autonomous Robots*, 44, 601 - 616.
10. Noa Markfeld, Samuel Olatunji, Dana Gutman, Shay Givati, Vardit Sarne Fleischmann, and Yael Edan. Feedback modalities for a table setting robot assistant for elder care. 03 2021.