POLITECNICO MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE E DELL'INFORMAZIONE

EXECUTIVE SUMMARY OF THE THESIS

Tino: First-Person Embodiment and Collaboration with a Non-Anthropomorphic Robot

LAUREA MAGISTRALE IN COMPUTER SCIENCE AND ENGINEERING

Author: ALESSANDRO RICCI

Advisor: PROF. ANDREA BONARINI

Co-advisor: FEDERICO ESPOSITI

Academic year: 2025-2026

---

1. Introduction

This thesis uses *Tino*, a non-anthropomorphic robot, as a platform to study first-person teleoperation in Virtual Reality (VR) and reduced-channel collaboration. One participant controls Tino from a first-person VR viewpoint, while a second participant shares the physical space with the robot and must coordinate through nonverbal cues such as motion, timing, proximity and simple signaling behaviors [2, 7]. The goal is to assess whether (i) the VR operator can experience embodiment toward a morphologically different body and (ii) the in-person participant can attribute communicative intent to the robot and cooperate with it under constrained communication.

The work follows the Physical Metaverse direction and builds on prior efforts to support embodiment with non-standard bodies through sensory translation and remapping [5, 9, 12]. Two experimental iterations were conducted and compared to evaluate how changes in control mapping and task constraints affect embodiment, communication legibility and collaboration outcomes.

The evaluation combines two perspectives: the VR operator reports embodiment-related outcomes (e.g., perceived control quality and ownership-like responses), while the in-person participant reports interaction outcomes (e.g., enjoyment and interpretability of the robot's indications). This separation allows assessing embodiment and collaboration as two linked but distinct layers of the same system.

The study is intentionally designed to reduce confounds from rich communication: participants are not instructed on an optimal strategy and the in-person participant is not told that the robot is controlled by another human. This supports the observation of spontaneous coordination and attribution under minimal, repeatable signals.

Contributions. This work contributes (i) a complete mixed physical–virtual platform for first-person VR teleoperation of a non-anthropomorphic robot, integrated with interactive scene devices and a unified data-logging pipeline and (ii) an iterative experimental evaluation showing how control coherence and task interdependence affect embodiment-related responses and reduced-channel collaboration.

Why Tino. Tino is intentionally non-humanoid: its morphology and locomotion differ from the human body and therefore do not provide the usual cues that support identification with a humanoid avatar. This makes it a suitable testbed to study whether embodiment can arise from control coherence and stable feedback rather than from visual similarity. In parallel, it allows evaluating whether a co-located human can treat a morphologically different robot as an intentional collaborator when communication is limited to simple, observable behaviors.

2. Background and research motivation

The thesis is positioned at the intersection of embodiment in VR, non-anthropomorphic robotics and collaborative Human–Robot Interaction (HRI) in hybrid physical–virtual environments. A key motivation is to understand whether phenomena often associated with humanoid avatars and rich communication channels can still emerge when morphology and communication bandwidth are constrained [6, 8, 10, 11].

The work focuses on two main aspects. The first concerns *embodiment*: whether a participant controlling Tino in first-person VR can experience the robot as the locus of action despite its non-human morphology. The second concerns *empathy and social attribution*: whether an in-person participant can interpret the robot as an intentional partner and show collaborative behavior even without verbal dialogue. Empathy is treated as a multi-component phenomenon expressed through perception, attribution and behavior, consistent with multidimensional approaches [4, 8].

A further motivation comes from the idea that embodiment can be facilitated by coherent sensorimotor contingencies and remapping, even when the controlled body differs from the human body schema [1, 5]. This supports the hypothesis that anthropomorphic resemblance is not a strict requirement; instead, interaction depends on the stability and readability of the action–feedback loop and on coherent multisensory cues [3].

3. Methodology and thesis approach

The adopted methodology combines system design with iterative experimental validation. Instead of relying on a purely simulated setting, the thesis builds a mixed physical–virtual setup and tests it with real participants in a collaborative task. This choice aligns with research that uses VR and mixed-reality systems as experimental media for embodied and social HRI [6, 10].

The approach includes three main steps. First, a VR teleoperation and interaction system is implemented around Tino, integrating robot control, interactive scene devices and logging. Second, a game-like collaborative environment is designed to elicit coordination under reduced-channel communication. Third, two experimental iterations are conducted and compared to assess how changes in control mappings and task constraints influence user experience and collaboration.

Reduced-channel interaction is enforced by design. The two participants are physically separated and are not allowed to coordinate through direct conversation. Coordination must therefore emerge from what can be perceived in the environment: the agents' movements, relative positioning, timing, and simplesignaling behaviors.

The evaluation relies on post-session questionnaires for both roles. The VR operator focuses on embodiment-related items (e.g., perceived control quality, ownership-like responses and perceived constraints), while the in-person participant focuses on collaboration-related items (e.g., enjoyment, interpretability of indications and perceived communicative intent). In the results, Likert items are reported on a 1–5 scale and categorical items as percentages.

4. System architecture and implementation

4.1. Hardware and interactive setup

The system is built around Tino, a non-anthropomorphic robot with an effortful dragging locomotion and an upper-body module used for expressive motion and signaling. The platform extends prior work on Tino-based interaction experiments [2, 7]. During this thesis, the platform underwent multiple hardware iterations to improve stability and usability during live sessions.

The physical experience takes place in a room equipped with interactive elements mirrored in Unity, including platforms, buttons and doors. Each physical element has a corresponding virtual representation and controller logic so that state

(a) Physical robot Tino used in the experimental setup.

(b) Virtual avatar of Tino used for first-person VR control.

Figure 1: Tino in the physical setup and its corresponding virtual avatar used in the VR environment.

changes remain consistent between the physical and virtual spaces.

Figure 1 shows the physical robot Tino and the corresponding virtual avatar used for first-person control in the VR environment.

4.2. Unity environment and VR control

Unity is used to implement the virtual environment, the Tino avatar and the interaction logic. The VR operator controls Tino from a first-person perspective through a VR setup based on Meta Quest. The avatar includes the camera viewpoint, controller tracking logic, a locomotion mapping and feedback layers related to Tino's state and terrain conditions.

A relevant design aspect is that the virtual environment is not only a visual replica of the physical room. It is also designed to support legibility from Tino's perspective through coherent visual and audio cues, so that both participants can rely on stable action–feedback patterns.

4.3. Communication and data logging

The system integrates Unity, the physical robot and multiple scene devices through a lightweight network layer that keeps the virtual and physical states synchronized in real time. The communication design prioritizes robustness during live sessions (device discovery, state updates and acknowledgements where needed) and supports reproducible experimental runs.

The main information flows are: (i) robot state and pose updates to keep the VR viewpoint consistent with the physical platform, (ii) control commands from the VR station to drive Tino's locomotion and signaling motions and (iii) interaction events from scene devices (e.g., platforms, doors and buttons) to update the shared task state in both the physical room and the virtual environment. Logged data include questionnaire responses and session events to enable post-hoc comparison across iterations.

A dedicated logging pipeline records key events and signals during the experiments. This supports both subjective questionnaire analysis and the interpretation of behavioral patterns observed during the collaborative tasks.

5. Experimental design

5.1. Collaborative task and workflow

The experiments are structured as an escape-room-like collaborative activity distributed across multiple rooms. Progression requires coordination between the human participant and Tino. The tasks are designed to create situations in which simple actions such as waiting, moving together or choosing the correct button become indicators of communicative interpretation and collaboration.

At the beginning of a session, the participant in the physical room explores the environment while the VR operator starts controlling Tino from its initial position. In the first room, the two participants must coordinate to reach their respective platforms to open the next door. In the second room, the human may either press buttons immediately or wait for Tino's indication of the target button. These situations help assess whether the human treats the robot as a necessary teammate or as a secondary object.

The first room is mainly aimed at establishing a shared objective and forces the participants to align spatially, since progression requires both to reach their respective target positions. The second room is the key interaction point: the in-person participant faces multiple buttons and can either act immediately or wait for the robot to approach and provide an indication. This design exposes a typical reduced-channel ambiguity: acting too early leads to trial-and-error, while waiting enables interpreting the robot as a partner whose behavior carries information.

Participants were given only a high-level explanation of the game and the available controls, without being told which action triggered which effect. This choice was intended to reduce instruction bias and encourage coordination strategies to emerge from interaction.

In practice, Tino's "communication" is limited to what can be inferred from its behavior: where it moves, when it stops, how it orients its body and how it performs simple signaling motions. The task structure is used to make these cues meaningful. For example, in the second room the intended strategy is to delay button presses until the robot reaches the interaction area, so that the in-person participant can interpret the robot's indications rather than proceeding by trial-and-error.

5.2. From first experiment to final experiment

We compared two experimental iterations. The final experiment introduced three main changes with respect to the first one. First, the VR participant was placed in a separate room, reducing the possibility of direct verbal communication. Second, the initial locomotion control was replaced by a movement mechanism that constrained motion more clearly and produced more repeatable control actions. Third, the logic of the second room was modified so that the correct button could open the door only when Tino was in the room and close enough to the buttons, increasing interdependence between participants.

These changes made the final experiment a stronger test of reduced-channel coordination and allowed a clearer comparison of embodiment and collaboration outcomes across iterations.

These changes targeted two issues observed in the first iteration: (i) limited controllability and learnability of the locomotion mapping, which affected the VR experience and (ii) weak task interdependence, which allowed the in-person participant to progress by trial-and-error without relying on the robot. The final design strengthens interdependence and makes the robot's actions more consequential for task success.

6. Results and discussion

6.1. Participants and completion rates

In the first experiment, 10 participants took part in total. Nine completed the VR phase and eight completed the physical collaboration phase, with one further reduction in the effective physical sample for part of the analysis due to a session issue. In the final experiment, 20 participants took part: 14 completed both sessions, while 6 completed only one (3 only VR and 3 only collaborative). Therefore, 17 participants were included in the VR-phase analysis and 17 in the collaborative-phase analysis; for some physical-phase analyses, the effective sample was further reduced to 16 because one participant did not complete the relevant part.

Unless otherwise stated, questionnaire items are reported as percentages; Likert items use a 1–5 scale.

Tables 1 and 2 summarize the comparative indicators discussed in this section.

6.2. Embodiment outcomes

The comparison suggests an overall improvement in the VR experience after the introduction of the revised control solution (Table 2) and an increase in ownership-related responses (Table 1). The improvement is consistent with increased appreciation of the movement system and with a reduction in participants describing the controls as non-intuitive.

Table 1: Binary and categorical questionnaire outcomes (percentages).

Indicator	First experiment	Final experiment	Change
Body felt like own body (body ownership)	0%	41%	↑
Movement appreciation	11%	59%	↑
Movements reported as not intuitive	67%	53%	↓
Perceived freedom of movement	78%	35%	↓
Participants reporting Tino was trying to communicate	43%	56%	↑
Waiting behavior in the second room	29%	50%	↑
Participants who touched the robot	71%	50%	↓

Table 2: Likert-scale outcomes (1–5 scale).

Indicator	First	Final
Average VR rating	3.33	3.82
Physical interaction enjoyment	4.29	4.31
Ease of understanding Tino's button indication	1.43	2.38

A relevant outcome is the reduction in perceived freedom of movement. In this context, the decrease is interpreted as evidence that participants more often recognized the intended environmental constraints (e.g., muddy unsafe areas) and adapted their behavior accordingly, rather than treating the environment as freely traversable.

Importantly, the observed improvements are attributed to a more coherent control loop and clearer constraints rather than to human-like appearance. In practical terms, embodiment increased when actions became easier to predict and when feedback remained consistent across repeated attempts.

Taken together, these results support the thesis claim that embodiment in this scenario depends primarily on the stability and readability of the sensorimotor loop rather than on anthropomorphic resemblance.

6.3. Interaction and communication outcomes

In the physical interaction phase, acceptance remained high across experiments, with nearly unchanged enjoyment scores (Table 2). This suggests that interaction with Tino remained engaging despite the robot's limitations and the reduced-channel communication setting.

The final experiment shows improvements in signal readability and coordination. Participants reported a higher ease of understanding Tino's indication of the correct button and a larger share interpreted the robot as trying to communicate (Tables 1 and 2). Waiting behavior in the second room also increased, consistent with the revised game logic making Tino's presence more relevant for progression.

The collaboration results suggest that interpretability is improved when the environment prevents premature actions (e.g., pressing buttons before the robot arrives) and when the task explicitly requires the robot's presence. This supports the design claim that legibility and interdependence are key drivers for reduced-channel cooperation.

The reduction in the rate of participants touching the robot suggests that Tino's appearance may still discourage direct physical contact. However, observed collaborative behaviors (waiting, encouraging and following indications) support the idea that empathy and social attribution can emerge through interaction dynamics and task meaning.

6.4. Key findings

Overall, the comparison suggests three main trends. First, the revised control solution improved the VR experience and increased ownership-related responses, indicating stronger embodiment. Second, task constraints that made Tino's presence necessary for progression improved coordination timing (e.g., more waiting behavior) and increased the perception of communicative intent. Third, despite remaining limitations, participants reported consistently high engagement in the physical interaction phase, supporting the feasibility of reduced-channel collaboration with a non-anthropomorphic robot.

6.5. Limitations

The study has two main limitations. First, reduced-channel communication makes some signals inherently ambiguous, especially when the in-person participant acts before observing the robot's behavior. Second, despite improvements, locomotion remains effortful and can still constrain learnability for part of the participants. These limitations motivate the future work directions aimed at increasing cue clarity and control robustness.

7. Conclusions and future developments

This thesis demonstrates a feasible direction for HRI in Physical Metaverse scenarios where the robot is non-anthropomorphic and communication is constrained. The main contribution is the implementation of a complete mixed physical–virtual experimental platform and the evidence that embodiment and social attribution can be supported without relying on human-like morphology.

The comparative results indicate that improving control mappings and interaction constraints strengthens embodiment-related outcomes and increases the readability of the robot's communicative behavior. This leads to a broader design conclusion: *interaction legibility*, more than human likeness, is the key variable for enabling meaningful collaboration with a non-anthropomorphic embodied robot.

Future developments should focus on increasing system robustness and improving signal clarity for both participants. Relevant directions include improving the stability of the embodiment input pipeline, extending multi-modal cues (especially audio cues for Tino's state), refining the readability of the physical set and validating additional tracking components for more consistent scene understanding.

Bibliography

[1] Anna Berti and Francesca Frassinetti. When far becomes near: Remapping of space by tool use. *Journal of Cognitive Neuroscience*, 12(3):415–420, 2000.

[2] Davide Cardillo. Conveying illusion of life and empathy through a non-anthropomorphic robot and non-verbal communication: The tino experience. Master's thesis, Politecnico di Milano, Milan, Italy, 2024.

[3] Yi-Chuan Chen and Charles Spence. Assessing the role of the "unity assumption" on multisensory integration: A review. *Frontiers in Psychology*, 8:445, 2017.

[4] Mark H. Davis. Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, 44(1):113–126, 1983.

[5] Federico Espositi and Andrea Bonarini. Towards a framework for embodying any-body through sensory translation and proprioceptive remapping: A pilot study. In *Proceedings of xCoAx 2023: 11th Conference on Computation, Communication, Aesthetics & X*, 2023.

[6] Yohei Kobayashi, Kazuki Murao, Kenji Hattori, Kenji Mase, Masaki Okada, and Tetsunari Inamura. Sigverse: A cloud-based vr platform for research on social and embodied human–robot interaction. *arXiv preprint arXiv:2005.00825*, 2020.

[7] Giovanni Sassone Massari. From teleoperation to embodiment: Designing nonverbal communication strategies for the non-anthropomorphic robot tino. Master's thesis, Politecnico di Milano, Milan, Italy, 2025.

[8] Ádám Miklósi, Péter Korondi, Vicente Matellán, and Márta Gácsi. Ethorobotics: A new approach to human-robot relationship. *Frontiers in Psychology*, 8:958, 2017.

[9] Erica Panelli. Odile: an expressive robotic agent for emotional exchange and information sharing. Master's thesis, Politecnico di Milano, Milan, Italy, 2023.

[10] Oliver Roesler, Phuong H. Nguyen, and Takayuki Kanda. Mixed reality as communication medium for human–robot collaboration. *arXiv preprint arXiv:2206.15380*, 2022.

[11] Sofia Seinfeld, Tiare Feuchtner, Antonella Maselli, and Jörg Müller. User representations in human–computer interaction. *Human-Computer Interaction*, 36(5-6):400–438, 2021.

[12] Maurizio Vetere. Bridging realities: Advancing human-avatar interaction through sensory translation systems in the physical metaverse. Master's thesis, Politecnico di Milano, Milan, Italy, 2023.