A Puppeteering Interface for Teleoperating High Degree-of-Freedom Social Robots

Roel van de Laar
Vrije Universiteit Amsterdam
Amsterdam, Netherlands
r.vande.laar@student.vu.nl

Kim Baraka
Vrije Universiteit Amsterdam
Amsterdam, Netherlands
k.baraka@vu.nl

Abstract

Teleoperating high degree-of-freedom (DoF) robots such as humanoids in interactive settings remains challenging due to high operator workload, and limited situational awareness. The majority of existing interfaces often relies on graphical dashboards, limiting natural, embodied control. This study explores an alternative paradigm: "Robot-as-Interface," where one humanoid robot (the puppet) is physically manipulated to control another (the performer) through direct joint-to-joint mapping. Following a co-design session with expert users, we developed an improved interface featuring joint locking, head orientation control, blockage detection, and a pausing toggle. A between-subjects user study (N=26) compared this expert-informed system against a baseline. Results show significantly improved system usability (SUS) and a reduction in perceived workload. Observations further revealed the importance of operator pacing, spatial positioning, and clear system feedback. Overall, results indicate that expert-informed enhancements can improve usability and operator experience in puppet–performer teleoperation, provided that hardware limits and user training are carefully addressed.

CCS Concepts

• **Human-centered computing** → **Usability testing**; **User interface management systems**; • **Computer systems organization** → **Robotic control**.

Keywords

Teleoperation interfaces, Wizard-of-Oz, Shadowing, Embodied interfaces, Puppeteering.

1 Introduction

Social robots often require context-dependent behaviors that exceed current autonomous capabilities [18]. In such cases, a hidden human operator can teleoperate the robot, commonly referred to as a

Wizard-of-Oz (WoZ) setup, to give the impression of advanced intelligence or interaction skills [11]. However, effectively teleoperating social robots remains challenging: operators often face difficulties in maintaining situational awareness, managing the complexity of control, and handling high cognitive load, particularly when interfaces lack a user-centered design [9, 10].

Most existing approaches, such as GUIs, separate the operator from natural, human-like control modalities. This study explores an alternative paradigm in which one social robot, termed the *puppet*, is physically manipulated to control an identical robot, the *performer*. By leveraging the human-like embodiment and built-in interaction modalities of social robots, we explore the idea that physically guiding the puppet—with a direct, one-to-one joint mapping to the performer—can facilitate more intuitive and natural control.

This research addresses two questions: (1) How practical is it to use a physical replica of a teleoperated robot as the interface? (2) How can such an interface be made more usable through expert-informed design improvements? We conducted a co-design study to explore interface improvements and evaluated these in a comparative user study focusing on usability and workload.

2 Related Work

Wizard-of-Oz (WoZ) enables naturalistic interaction studies before full autonomy is feasible [11]. However, WoZ interfaces remain labour-intensive, inconsistent across studies, and highly dependent on operator skill. Riek's review [11] found that most WoZ systems rely on ad-hoc GUIs, and later work highlights persistent issues such as high wizard workload and extensive preparation demands [1, 12].

Prior research has examined various methods of teleoperation, including GUI panels [10, 14], immersive VR/AR [14, 15, 19], wearables [3, 8], and proxies [4, 5]. Overall, these studies converge on the need to balance intuitive embodied control with multimodal feedback and support for high-dimensional complexity. Yet, to our knowledge, no work has investigated one humanoid directly teleoperating another, a gap our puppet–performer study addresses.

Puppeteering replaces GUI-based teleoperation with embodied, proxy-based control. In social HRI, Troughton et al. [20] showed that expert improvisers can generate rich behaviours when puppeteering Pepper, motivating efforts to formalize reusable strategies. Schrage-Prent et al. [13] demonstrated that incremental, feedback-rich puppeteering can be more usable than graphical interfaces, highlighting the value of accessible, embodied programming.

Key telepresence and leader–follower systems illustrate the design trade-offs of puppeteering. TELESAR VI [17] achieves high-fidelity full-body telexistence but requires complex hardware,
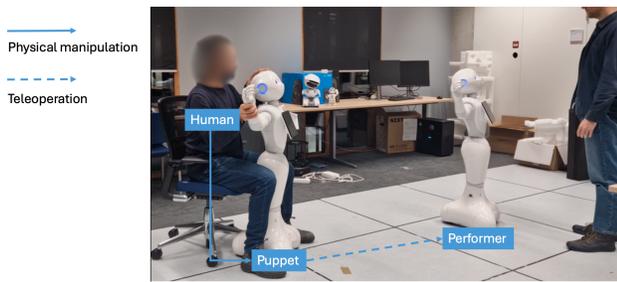
Figure 1: Visualization of the puppeteering setup showing the interaction between the human operator and the performer.

whereas low-cost replicas such as *ROBOPuppet* [4] enable intuitive mapping for novices through physical miniatures. Together, these works show that puppeteering offers intuitive control but faces recurring challenges in workload, mapping ambiguity, and hardware constraints. Our work extends this combined trajectory through a robot-as-interface paradigm in which one Pepper robot directly teleoperates another. By integrating co-design-informed improvements that go beyond simple shadowing of join values, we empirically evaluate how embodied puppeteering can deliver more intuitive, effective, and user-centered teleoperation, specifically in interactive scenarios with highly articulated social robots.

3 Methodology

3.1 System Design

3.1.1 System Overview. The basic version of our system uses two SoftBank Robotics Pepper robots integrated via the Social Interaction Cloud (SIC) framework [16][1][2]. The "Puppet" is physically manipulated by the operator, and its joint positions, including base displacement, are streamed in real-time to the "Performer" through a controllable sampling rate (figure 1).

3.1.2 Co-Design Session. To identify usability challenges, we conducted a participatory co-design session with three expert users: a professional puppeteer, circus artist and juggler, a Cultural Data & AI researcher with dancing experience, and a MSc AI graduate. Participants explored the baseline system through a semi-structured setup and identified several issues: (a) motion misalignment caused by dropped or desynchronized commands, (b) the inability to recover or reposition the robots once they became obstructed, and (c) the lack of clear orientation or gaze cues, which made social gestures difficult to judge and perform naturally.

3.1.3 Implemented Features. Based on the co-design outcomes, we implemented the following visible interface features:

Physical Head Handle: To address issue (c), a handle was attached to the back of the puppet's head (figure 2a) to facilitate intuitive gaze direction and orientation control.

Joint Locking: Operators can lock specific joint chains (left arm, right arm, or head) by pressing the robot's side bumpers or head sensor (figure 2b). This allows the operator to stabilize one limb while manipulating another, reducing cognitive load.

Blockage Feedback: To address limited situational awareness, the system compares the puppet's intended position with the performer's actual position. If a discrepancy persists (indicating an obstacle), the puppet's chest LEDs turn red and a spoken notification is triggered (figure 2c).

Pause/Resume Toggle: To address issue (b), a toggle to pause the puppeteering link is mapped to the back bumper (figure 2d). It allows the operator to reposition the puppet without moving the performer, facilitating recovery from awkward physical positions.

3.2 Evaluation Study

We conducted a between-subjects user study (N = 26) comparing the improved puppeteering system against a baseline model that streamed only real-time joint positions and base movement from puppet to performer, without any additional control or feedback mechanisms.

3.2.1 Ethical Considerations. The experimental study was conducted in accordance with institutional ethical guidelines, according to an online self-check. All participants were informed about the study procedures, their right to withdraw at any time without penalty, and the voluntary nature of their participation. Written informed consent was obtained prior to the start of the experiment. Each participant received €10 as compensation for their participation in the form of a gift card.

3.2.2 Experimental Setup and Tasks. Participants performed three tasks (figure 3), focusing on different aspects of robot control that were aimed at testing the newly added features. The participants using the new model were made familiar with the new features but not explicitly encouraged to use them. They were told to use them if they believed it would be beneficial to complete the task. The participants were given a maximum time of 5 minutes to complete each task. Not succeeding to complete a task in the given time limit or upon giving up themselves, the task would be marked as a failure.

The participants completed the following tasks:

Navigation: Navigating the robot around an obstacle (Figure 3a), requiring base movement and potential repositioning (testing the Pausing feature).

Guided Tour: Navigating the performer to two points (Figure 3b), pointing and directing gaze at a cup at each location (testing Head Handle and coordination).

Careful Delivery: Carrying a small cube along a pad to a drop-off point (Figure 3c), requiring stable arms during movement (testing Joint Locking).

3.2.3 Measures. The two systems were compared using the System Usability Scale (SUS) [2], NASA-TLX [6], and three qualitative questions (requested additions, most helpful feature, most unnecessary feature).

(a) Prototype for a physical head handle.

(b) Operator pressing Pepper's bumper to lock an arm.

(c) LEDs turn red after joint mismatch.

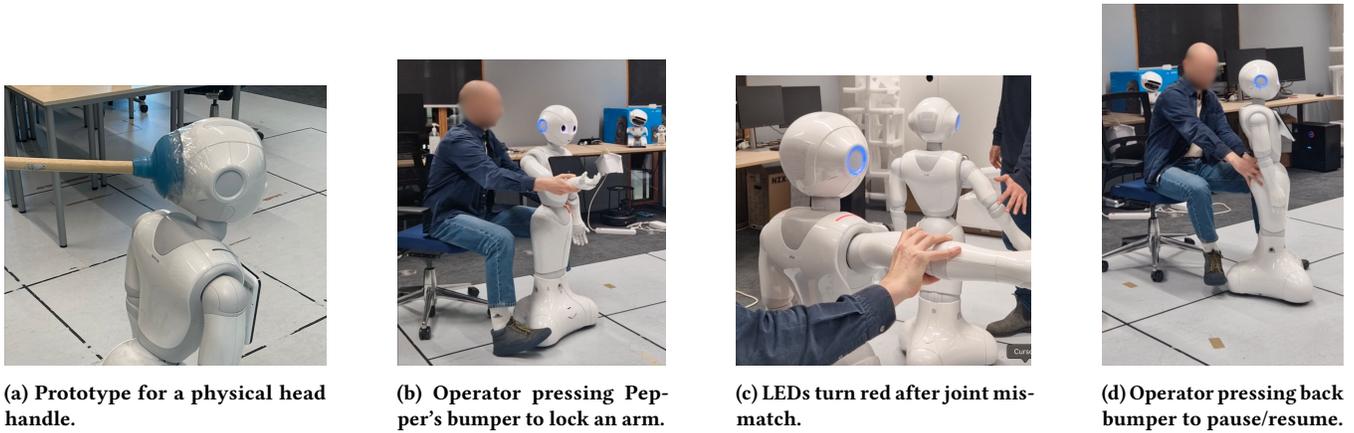**(d) Operator pressing back bumper to pause/resume.**

Figure 2: Overview of implemented features. (A) Physical head handle. (B) Joint locking via side bumpers. (C) Blockage feedback through LEDs and audio. (D) Pause/resume toggle on back bumper.

(a) Task 1: Navigation

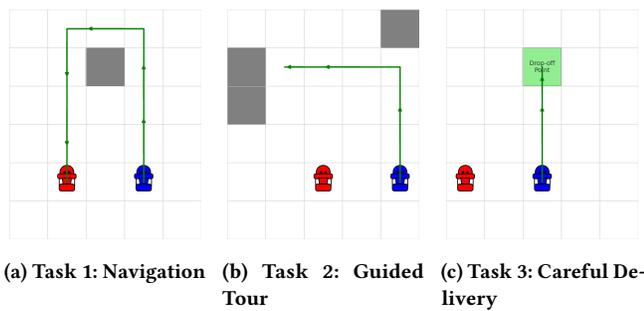**(b) Task 2: Guided Tour**

(c) Task 3: Careful Delivery

Figure 3: Top-down visualization of the three tasks. The Puppet robot is depicted in red and the Performer robot in blue. A single grey square indicates a chair, and double grey squares indicate a table.
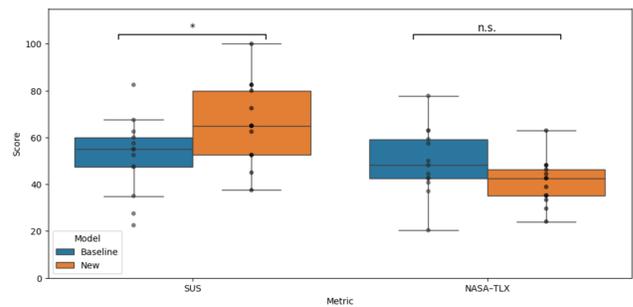

Figure 4: Box plots showing System Usability Scale (SUS; higher = better) and NASA–TLX workload scores (lower = less workload) for the baseline and new models. The asterisk (*) denotes a significant difference ($p < .05$) based on independent-samples t-tests; "n.s." = not significant.

4 Results

4.1 Participant Demographics

A total of 26 participants took part, evenly split between baseline and improved system conditions. Ages ranged from 21 to 68 years, with $n = 15$ identifying as male, $n = 9$ as female, and $n = 2$ as non-binary. Prior robot experience is reported by 10 participants, and 12 reported performance experience.

4.2 Usability and Workload

Participants using the new model reported significantly higher usability scores ($M = 66.35$) compared to the baseline group ($M = 51.73$). An independent-samples t-test confirmed this difference was significant ($t(24) = 2.22$, $p = .036$), with a large effect size ($Cohen's\ d = 0.87$, 95% CI [1.03, 28.20]).

NASA–TLX workload scores showed a non-significant trend toward lower workload for the improved model ($M = 40.88$) compared to the baseline ($M = 49.72$), $t(24) = -1.81$, $p = .082$, Cohen's $d = -0.71$, 95% CI [-18.88, 1.22].

Figure 4 visualizes these distributions, showing a significant difference in usability (SUS; *) and a non-significant difference in perceived workload (NASA–TLX; n.s.).

4.3 Task Performance

Task success rates differed significantly for Task 1, with participants using the improved system achieving a higher success rate than those using the baseline (100% vs. 53.8%, p = .015). For Task 3, success rates did not differ significantly between conditions. No significant differences were found in completion times for Tasks 1 or 3. Task 2 was excluded from quantitative analysis due to the absence of a reliable success criterion.

4.4 Qualitative Findings

Participants' open-ended feedback highlighted several recurring themes across both system versions. The most common requested additions were improved synchronisation, voice control, and general design refinements, with baseline users especially emphasising navigation and clearer feedback, while new model users requested

finer alignment and more dexterous hand control. Across both versions, participants found the robot-as-interface intuitive, and new model users specifically valued the pause function and arm-locking mechanism. Unnecessary elements differed by system: baseline users pointed to control complexity and some hardware features, whereas new model users focused on head-related design issues and overly sensitive sensors.

4.5 Lessons Learned

A series of practical insights emerged from observations during the study. Misalignment between puppet and performer was common, driven by hardware limits, required movement thresholds, collision-avoidance interruptions, and the performer's maximum replication speed. Subtle gestures were often lost, fast motions increased desynchronisation, and some users triggered instability or shutdowns with abrupt movements. Operators who adopted a gentler, more deliberate controle style avoided many of these issues, indicating that personal interaction style strongly affects puppeteering performance. Spatial positioning also influenced task success: some participants inadvertently blocked the robots or restricted their own maneuverability, reflecting a need for clearer guidance and possibly alternative navigation controls, such as hybrid control, when physical space is limited.

Additional challenges stemmed from the interface and the experimental setup. The new model's rear pause button was sometimes unintentionally pressed without the operator realising it, while the baseline model's lack of such a feature caused participants to physically reposition the puppet when it was obstructed. System shutdowns and instability, sometimes unexplained, prevented some trials from completing.

These observations suggest several directions for refinement: (i) Provide operators with spatial positioning guidelines; (ii) Explore navigation-control alternatives that require less physical space; (iii) Implement acceleration-limiting features to prevent system failures; (iv) Ensure pause and lock mechanisms are appropriately placed and clearly indicated; (v) Provide operator training to familiarise participants with the system and its limitations.

5 Discussion and Conclusion

This study demonstrates that a "Robot-as-Interface" paradigm can provide an intuitive method for social robot teleoperation, provided the interface is designed to manage complexity. The improved puppeteering model, developed through a co-design process, significantly improved usability scores (SUS) compared to the baseline system. Although workload scores (NASA–TLX) were lower for the new model, this difference did not reach statistical significance.

The improvement in usability scores suggests that the added features successfully mitigated some of the inherent difficulties of high-DoF control. The joint locking mechanism, in particular, allowed operators to decompose complex whole-body tasks into manageable segments, addressing the dimensionality challenge highlighted in prior literature.

The observations and qualitative findings indicate that hardware and synchronization constraints, stemming from Pepper's design, motion streaming thresholds, and collision avoidance behaviour, remain barriers to fine-grained control. These constraints

may also interact with operator behaviour, as gentler, more deliberate handling appeared to improve performance, echoing findings that simplified and constrained control strategies can enhance teleoperation outcomes [7]. This highlights the need for both technical optimisations and structured operator training to ensure consistent results across different interaction styles.

Environmental and spatial constraints also emerged as influential factors. Limited operating space in the lab sometimes hindered navigation, a challenge likely to extend to real-world Wizard-of-Oz deployments. Hybrid control schemes that decouple navigation from full-body puppet manipulation, or integrate joystick-based navigation, may help mitigate such constraints.

Overall, this study provides exploratory evidence that a robot-as-interface paradigm is a promising approach to teleoperation in social robotics, offering intuitive mapping between operator action and robot behaviour. While quantitative improvements were observed in usability and specific task performance, the primary contribution lies in the observed interaction patterns and lessons learned regarding hardware constraints, operator strategies, and spatial setup. Future research should therefore build on these findings through iterative prototyping, feature-specific testing, and expanded qualitative evaluation, while also systematically comparing this approach to alternative teleoperation modalities, assessing its scalability in more complex tasks, and integrating multimodal feedback to further enhance intuitiveness and situational awareness.

Acknowledgments

References

[1] Alexandra Bejarano, Saad Elbeleidy, Terran Mott, Sebastian Negrete-Alamillo, Luis Angel Armenta, and Tom Williams. 2024. Hardships in the land of oz: Robot control challenges faced by hri researchers and real-world teleoperators. In *2024 33rd IEEE International Conference on Robot and Human Interactive Communication (ROMAN)*. IEEE, 1914–1921.

[2] John Brooke. 1996. SUS: A "Quick and Dirty" Usability Scale. In *Usability Evaluation in Industry*, Patrick W. Jordan, Bruce Thomas, Bernard A. Weerdmeester, and Ian L. McClelland (Eds.). Taylor & Francis, London, 189–194.

[3] Kourosh Darvish, Luigi Penco, Joao Ramos, Rafael Cisneros, Jerry Pratt, Eiichi Yoshida, Serena Ivaldi, and Daniele Pucci. 2023. Teleoperation of humanoid robots: A survey. *IEEE Transactions on Robotics* 39, 3 (2023), 1706–1727.

[4] Anna Eilering, Giulia Franchi, and Kris Hauser. 2014. Robopuppet: Low-cost, 3d printed miniatures for teleoperating full-size robots. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 1248–1254.

[5] Zipeng Fu, Tony Z Zhao, and Chelsea Finn. 2024. Mobile aloha: Learning bimanual mobile manipulation with low-cost whole-body teleoperation. *arXiv preprint arXiv:2401.02117* (2024).

[6] Sandra G Hart. 2006. NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, Vol. 50. Sage publications Sage CA: Los Angeles, CA, 904–908.

[7] David Kent, Carl Saldanha, and Sonia Chernova. 2017. A comparison of remote robot teleoperation interfaces for general object manipulation. In *Proceedings of the 2017 ACM/IEEE international conference on human-robot interaction*. 371–379.

[8] Tobias Klamt, Max Schwarz, Christian Lenz, Lorenzo Baccelliere, Domenico Buongiorno, Torben Cichon, Antonio DiGuardo, David Droeschel, Massimiliano Gabardi, Malgorzata Kamedula, et al. 2020. Remote mobile manipulation with

the centauro robot: Full-body telepresence and autonomous operator assistance. *Journal of Field Robotics* 37, 5 (2020), 889–919.

[9] Tsung-Chi Lin, Achyuthan Unni Krishnan, and Zhi Li. 2024. Perception and Action Augmentation for Teleoperation Assistance in Freeform Telemanipulation. *ACM Transactions on Human-Robot Interaction* 13, 1 (2024), 1–40.

[10] Daniel J Rea and Stela H Seo. 2022. Still not solved: A call for renewed focus on user-centered teleoperation interfaces. *Frontiers in Robotics and AI* 9 (2022), 704225.

[11] Laurel D Riek. 2012. Wizard of oz studies in hri: a systematic review and new reporting guidelines. *Journal of Human-Robot Interaction* 1, 1 (2012), 119–136.

[12] Shane Romero, Saad Elbeleidy, and Tom Williams. 2024. Understanding Barriers to Entry and Invisible Labor for Educational Care Wizards. In *2024 33rd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 1287–1293.

[13] Nienke Schrage-Prent, Daniel F Preciado Vanegas, and Kim Baraka. 2024. Interactive Robot Programming Inspired by Dog Training: An Exploratory Study. In *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*. 965–969.

[14] Emily Sheetz, Mina Kian, Misha Savchenko, Steven Jens Jorgensen, Evan Laske, Shaun Azimi, Maja Mataric, and Odest Chadwicke Jenkins. 2022. Comparing human input in rviz and virtual reality interfaces for 3d object manipulation tasks. In *First International Workshop on Horizons of an Extended Robotics Reality at IEEE IROS 2022*.

[15] Mehdi Sobhani, Manuel Giuliani, Alex Smith, and Anthony G Pipe. 2020. Robot teleoperation through virtual reality interfaces: Comparing effects of fixed and moving cameras. In *3rd International Workshop on Virtual, Augmented, and Mixed Reality for Human-Robot Interaction. ACM*.

[16] Social AI Lab, Vrije Universiteit Amsterdam. 2023. Social Interaction Cloud (SIC) V2. https://socialrobotics.atlassian.net/wiki/spaces/CBSR. Accessed: 15-07-2025.

[17] Susumu Tachi, Yasuyuki Inoue, and Fumihiro Kato. 2020. Telesar vi: Telexistence surrogate anthropomorphic robot vi. *International Journal of Humanoid Robotics* 17, 05 (2020), 2050019.

[18] Yuchuang Tong, Haotian Liu, and Zhengtao Zhang. 2024. Advancements in humanoid robots: A comprehensive review and future prospects. *IEEE/CAA Journal of Automatica Sinica* 11, 2 (2024), 301–328.

[19] Eleftherios Triantafyllidis, Christopher Mcgreavy, Jiacheng Gu, and Zhibin Li. 2020. Study of multimodal interfaces and the improvements on teleoperation. *IEEE Access* 8 (2020), 78213–78227.

[20] Irene Alcubilla Troughton, Kim Baraka, Koen Hindriks, and Maaike Bleeker. 2022. Robotic improvisers: Rule-based improvisation and emergent behaviour in hri. In *2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 561–569.