

Evaluating Mental State in Active Physically Assistive Human-Robot Interaction

Nourhan Abdulazeem¹, Nils Sichert², Ji Yuan Feng¹, Yue Hu¹

Abstract—As robots become more common in homes, understanding how their physical behaviors affect human mental states is essential. This study compares the effects of active (unanticipated) and passive (expected) robot behaviors during a physical balance task. Results show that while passive behaviors are more cognitively ergonomic, active behaviors help reduce imbalance but increase mental effort. Users’ perceptions of the robot remain unchanged across behavior types. We also find that peripheral skin temperature, combined with age and personality traits, can help robots infer users’ cognitive load and belief levels. The work highlights the potential of minimal, easily accessible data for assessing human mental states in physically assistive interactions.

I. INTRODUCTION

With the global aging population and lack of caregivers [1], physical Human-Robot Interaction (pHRI) can play a vital role in caregiving, particularly for older adults requiring physical assistance. Robots could intervene to prevent risks like loss of balance. In such situations, the robot’s actions should be executed considering not only the users’ physical states but also their mental states.

In previous research, *active pHRI* is defined as unanticipated physical actions initiated by the robot while physically coupled. These actions occur during a shared task without prior notification, leaving the human partner unaware of their timing or nature (such as physical help provided for an unanticipated fall). In contrast, *passive pHRI* aligns with human expectations, ensuring smoother collaboration [2], [3], [4]. Prior studies have primarily explored passive pHRI, with a limited investigation into unanticipated (active) robot actions in assistive scenarios [5].

Research on cognitive load in human-robot collaboration has highlighted online physiological indicators such as heart rate, Electrocardiography (ECG), Electroencephalography (EEG), pupil diameter, and blinking rates [6], [7]. [8] link peripheral skin temperature to cognitive stress, suggesting it as a potential metric for online assessment. Although these works have shown the potential of physiological signals to reflect cognitive states in human-robot collaboration, most studies focus on non-physical tasks in domains like manufacturing, gaming, or social interaction. Applications in



(a) Front view

(b) Side view

Fig. 1: A participant holds the robot, awaiting perturbation, while an experimenter prepares to drop the sandbag. The participant wears a PPG earlobe electrode, a left fifth-finger temperature probe, and a waist IMU. A second IMU is located on the board.

physically assistive scenarios—particularly those involving active robot behavior—remain limited. Furthermore, existing studies often depend on complex sensor setups that are impractical for real-world use, and few explore minimal sensing approaches.

Our study examines how active versus passive robotic behavior influences user cognitive ergonomics and perception in balance maintenance tasks with a minimal sensor setup. We assess whether peripheral skin temperature and pre-gathered data (demographics, personality traits) can infer mental state. As robots become more prevalent in assistive settings, it is crucial to develop interaction strategies that accommodate both safety and cognitive workload. Understanding these cognitive factors will inform the design control frameworks for robots that dynamically adapt their behavior to user needs.

This paper is a concise and shortened version of the IEEE-RAS Robotics and Automation Letters (RA-L) article, for more details and analysis, the reader can refer to [9].

II. METHODOLOGY

To address the identified gaps in active pHRI for physical assistance, we formulate the following research questions:

RQ1: Will active robot behavior impact participants’ mental state compared to passive robot behavior in an assistive pHRI scenario?

RQ2: What correlations exist between participants’ demographics, physical data, personality, physiological data, and mental state during a physical assistive pHRI scenario?

We conducted a user study where participants performed a balance maintenance task while holding onto a robot for

¹ Department of Mechanical and Mechatronics Engineering, University of Waterloo, ON, Canada (e-mail: nourhan.abdulazeem@uwaterloo.ca; jy5feng@uwaterloo.ca; yue.hu@uwaterloo.ca)

² Technical University of Hamburg, Germany (e-mail: nils.sichert@tuhh.de)

We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), reference RGPIN-2022-03857, and the National Research Council (NRC), reference AiP-302-1. We thank the Waterloo RoboHub for their support with the equipment.

assistance, as shown in Fig. 1. Participants stood on a half-ball balance board while optionally holding a 7-DoF Rethink Robotics Sawyer robot for support. A 4.2 kg sand kettlebell was dropped at fixed intervals using a pulley system to perturb the board. Participants were randomly assigned to one of two between-subject conditions:

- **Passive behavior:** The robot complied with participants' movements without intervention. In this condition, the robot was controlled via low-stiffness impedance control (20 N/m stiffness, 1 Ns/m damping on all joints) and did not apply active counterforces during perturbations.
- **Active behavior:** The robot applied corrective forces during perturbations. In this condition, when a perturbation occurred, the robot switched to joint position control for 5 seconds to simulate corrective forces. Only joint 2 moved (incrementing by 0.01 rad up to 0.4 rad) to apply an upward counterbalancing motion before reverting to impedance control.

This between-subject design helped eliminate learning effects and bias, particularly those seen in pilot tests with within-subject trials. Participants were assigned based on gender, age, weight, height, arm span, and balancing ability. Skin temperature (finger probe) and IMUs (one on the waist, one on the balance board) were used for physiological and balance assessment. Participants were not informed of the condition they were assigned before the experiment.

To ensure consistency, all participants completed a brief familiarization phase on the balance board without robot assistance, allowing them to adjust to the task requirements. Each session then lasted 140 seconds, beginning with 40 seconds of initial perturbations without a secondary task, followed by a 100-second "actual session" involving four timed perturbations and a concurrent auditory math task (running sum of spoken digits 1–6). This secondary task was intended to simulate real-world multitasking demands and served to quantify cognitive load. Perturbation events were precisely timed to ensure uniform exposure across participants and conditions. Their performance in the secondary task is evaluated with the following index:

$$\text{Performance Index} = 0.5(\text{Accuracy} + \text{Efficiency})$$

$$\text{Accuracy} = \frac{\text{Total number of correct answers}}{\text{Total number of attempted calculations}}$$

$$\text{Efficiency} = \text{Total number of correct answers}$$

Aside from personal background information, personality via the Ten Items Personality Measure (TIPI) questionnaire [10], IMUs, and temperature, data collection also included subjective self-reports via questionnaires and open-ended questions. Specifically, we adopted the StuMMB-Q (Student Mental Load and Mental Effort Biology Questionnaire) from educational science, which evaluates both Mental Load and Mental Effort [11], for robot perception, we used the well-established RoSAS (Robot Social Attribute Scale) [12] and

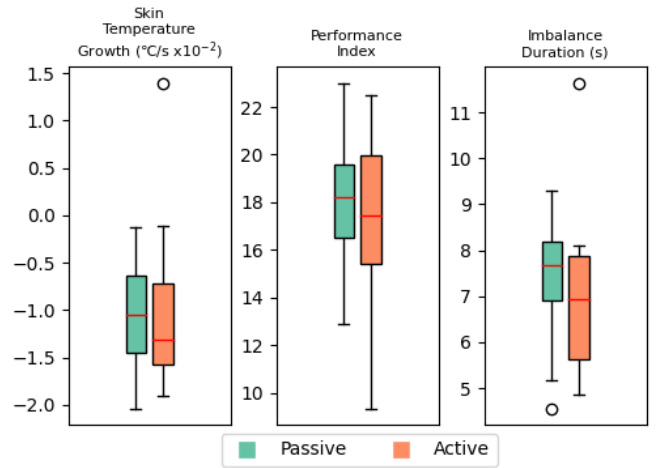


Fig. 2: Objective measures during the actual session. The sign of skin temperature indicates increases or decreases from baseline.

complement it with Perceived Safety from the Godspeed Series Questionnaire (GSQ) as recommended by the literature [13].

The experiment lasted approximately 45 minutes per participant, including setup, training, data collection, and post-experiment debriefing. All participants provided informed consent before participation. The study was approved by the institutional ethics review board, and participants were debriefed about the study's objectives and potential risks. They had the opportunity to withdraw at any time without consequences.

III. RESULTS

We recruited 54 participants, mostly university students; 35 were analyzed, with ages from 17 to 35, 27 Females (F), and 8 Males (M); 16 were assigned to the passive condition (ages: 21.6 ± 2.8 , 13 F, 3 M), and 19 to the active condition (ages: 23.7 ± 3.8 , 14 F, 5 M). Three participants were excluded for ineligibility and 16 for balance failures. We refrained from recruiting older adults given hardware constraints and the chosen challenging physical task.

Objective measures (skin temperature, performance index, imbalance duration) were analyzed over 100 [sec], excluding the first 4 [sec] for task familiarity. Temperature gradients and gyroscope data were processed to assess imbalance, using participant-specific thresholds. The final scores of the subjective measurements (questionnaires) were computed by averaging items within each sub-scale.

Both objective and subjective (mental load questionnaires) measures indicate higher cognitive strain in the active condition, as shown in Figs. 2 and 3, respectively. However, robot perception dimensions (competence, warmth, discomfort, perceived safety) remained unchanged between conditions.

Participants in the active condition exhibited greater skin temperature variation, suggesting heightened stress responses. Additionally, imbalance durations were shorter in the active condition, indicating that corrective robotic forces improved balance stability at the expense of increased cog-

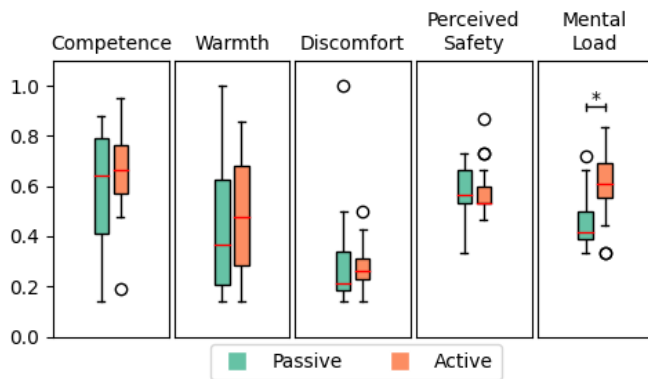


Fig. 3: Subjective measures. Each dimension ratings are normalized to fall from 0 to 1. Significance levels ($*$:= p -value < 0.05) are indicated on lines between conditions.

nitive effort. The subjective mental load ratings further supported this finding, with participants in the active group reporting higher perceived workload. Statistical analysis using the Mann-Whitney U test confirmed significant differences in mental load scores between passive and active conditions ($p < 0.01$). The passive group exhibited greater performance efficiency in the secondary math task, suggesting reduced cognitive burden compared to the active group. In response to **RQ1**, these results indicate that active robot behavior increases mental load but does not significantly affect participants’ perceptions.

In response to **RQ2**, correlation analysis, shown in Fig. 4, revealed that age, agreeableness, and conscientiousness influence cognitive ergonomics and belief.

Further analysis of postural stability metrics showed that participants in the active condition exerted more corrective movements, as indicated by higher oscillations recorded in IMU data. These findings suggest that while active assistance improves balance recovery, it introduces additional cognitive and motor demands.

IV. DISCUSSION

1) *Cognitive Ergonomics*: Passive interaction is more cognitively ergonomic than active interaction. While active behavior reduced imbalance durations, unexpected robot movements added cognitive strain. This highlights a key challenge in pHRI design: optimizing assistance while minimizing cognitive load.

2) *Perception and Belief*: Users’ perception of the robot was unaffected by behavior type, suggesting that task relevance (safety and cognitive ergonomics over enjoyment) is prioritized in assistive settings. This finding contrasts with previous work in social robotics, where unexpected behaviors often influence perceived warmth and trustworthiness.

3) *Predicting Mental States*: Results show that peripheral skin temperature correlates with mental load and perceived warmth, supporting its use in real-time pHRI adaptation. Future studies should explore multimodal sensing techniques to enhance predictive accuracy.

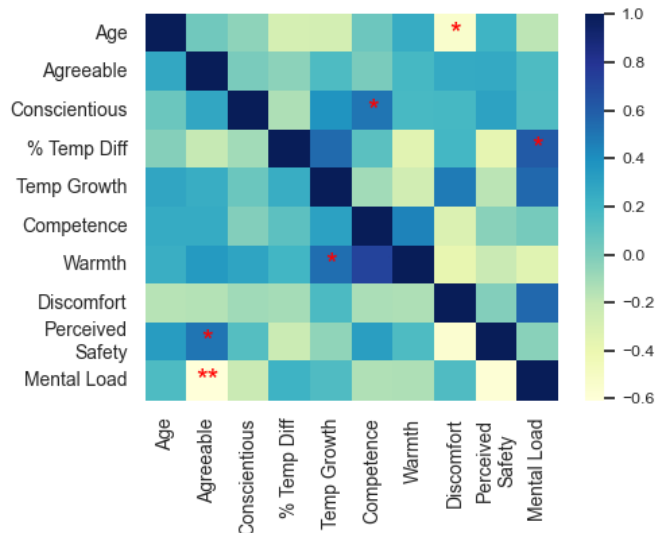


Fig. 4: Heatmap illustrates strong Spearman’s correlations ($|\rho| > 0.6$) observed for the passive behavior (upper triangle) and the active behavior (lower triangle) between subjective measures and both pre-gathered and online data. Correlations marked with $*$, and $**$ indicate a p -value < 0.05 and < 0.01, respectively.

4) *Limitations and future work*: The sample consisted mainly of university students, limiting the generalizability to older adults. Additionally, the between-subject design restricts direct comparisons of individual differences. Future work should explore a within-subject design, expand the participant pool, and explore additional physiological indicators such as galvanic skin response (GSR) and heart rate variability (HRV). Investigating long-term adaptation effects and integrating real-time machine learning models for cognitive state prediction could further improve assistive robotic systems.

While our results support the feasibility of using physiological signals for mental state inference, further validation with diverse populations and real-world applications is necessary. Importantly, our findings are specific to the particular implementation of active physical assistance used in this study, i.e. unanticipated robot actions for balance recovery, and may not generalize to all forms of active pHRI. However, we frame our claim within the context of the Hancock Hedonomics Pyramid [14], which prioritizes physical safety and ergonomics as foundational in human-robot interaction, especially in assistive contexts. As such, we posit that if the robot’s active behavior is strictly tied to improving user safety, as it is in this paper, similar cognitive trade-offs (i.e., increased mental effort for improved physical stability) may reasonably be expected. This is not a conclusion but a tentative claim that invites future work to test its boundaries in other settings and with different implementations of active assistance. Understanding the long-term effects of specific forms of active versus passive robotic behavior on cognitive ergonomics will be essential for designing effective assistive robots tailored to different care scenarios.

V. CONCLUSION

This study demonstrates that active pHRI increases cognitive load while reducing imbalance, whereas passive pHRI is more cognitively ergonomic. Robot perception remains unaffected by behavior type. Peripheral skin temperature, combined with demographic and personality traits, shows promise in inferring cognitive ergonomics and belief. Future research should refine predictive models for real-time robotic adaptation in assistive settings.

Expanding on these findings, future work should explore adaptive robotic strategies that balance cognitive ergonomics and physical support. While we found that active robotic behaviors can improve balance stability at the cost of higher cognitive load, this result stems from a specific form of active assistance. We encourage caution in generalizing this outcome to all active pHRI systems. Nevertheless, in line with the Hedonomics framework, we believe that when safety is the principal objective of the robotic assistance, as opposed to enhancing enjoyment or hedonic engagement, similar dynamics may emerge. Future longitudinal and comparative studies across varied robot behaviors and user groups will be essential to confirm or refine this hypothesis. By implementing adaptive models that consider physiological and behavioral data, future assistive robots could improve user comfort and efficiency, making them more viable for widespread adoption in healthcare and home environments.

REFERENCES

- [1] W. H. Organization, "Ageing: Global population."
- [2] Y. Hu, M. Benallegue, G. Venture, and E. Yoshida, "Interact With Me: An Exploratory Study on Interaction Factors for Active Physical Human-Robot Interaction," *IEEE Robot. Autom. Lett.*, vol. 5, no. 4, pp. 6764–6771, Oct. 2020.
- [3] Y. Hu, N. Abe, M. Benallegue, N. Yamanobe, G. Venture, and E. Yoshida, "Toward Active Physical Human-Robot Interaction," *IEEE Trans. Hum. Mach. Syst.*, 2022.
- [4] N. Abe, Y. Hu, M. Benallegue, N. Yamanobe, G. Venture, and E. Yoshida, "Human Understanding and Perception of Unanticipated Robot Action in the Context of Physical Interaction," *ACM THRI*, p. 3643458, Jan. 2024.
- [5] N. Abdulazeem and Y. Hu, "Human Factors Considerations for Quantifiable Human States in Physical Human-Robot Interaction," *Sensors*, vol. 23, no. 17, p. 7381, Jan. 2023, number: 17 Publisher: Multidisciplinary Digital Publishing Institute.
- [6] M. Lagomarsino, M. Lorenzini, P. Balatti, E. D. Momi, and A. Ajoudani, "Pick the Right Co-Worker: Online Assessment of Cognitive Ergonomics in Human-Robot Collaborative Assembly," *IEEE Transactions on Cognitive and Developmental Systems*, pp. 1–1, 2022.
- [7] M. I. Ahmad, I. Keller, D. A. Robb, and K. S. Lohan, "A framework to estimate cognitive load using physiological data," *Personal and Ubiquitous Computing*, pp. 1–15, 2023.
- [8] W. Bainbridge, S. Nozawa, R. Ueda, K. Okada, and M. Inaba, "A Methodological Outline and Utility Assessment of Sensor-based Biosignal Measurement in Human-Robot Interaction," *Int. J. Soc. Robot.*, vol. 4, no. 3, pp. 303–316, Aug. 2012.
- [9] N. Abdulazeem, N. Sichert, J. Y. Feng, and Y. Hu, "Quantifying human mental state in interactive phri: Maintaining balancing," *IEEE Robotics and Automation Letters*, vol. 10, no. 3, pp. 2958–2965, 2025.
- [10] S. D. Gosling, P. J. Rentfrow, and W. B. Swann, "A very brief measure of the Big-Five personality domains," *JRP*, vol. 37, no. 6, pp. 504–528, Dec. 2003.
- [11] M. Krell, "Evaluating an instrument to measure mental load and mental effort considering different sources of validity evidence," *Cogent Education*, vol. 4, Jan. 2017.
- [12] C. M. Carpinella, A. B. Wyman, M. A. Perez, and S. J. Stroessner, "The Robotic Social Attributes Scale (RoSAS)," in *ACM HRI*, Mar. 2017, pp. 254–262.
- [13] C. Bartneck, T. Belpaeme, F. Eyssel, T. Kanda, M. Keijsers, and S. Šabanović, *Human-Robot Interaction*, 1st ed. Cambridge University Press, Feb. 2020.
- [14] P. A. Hancock, A. A. Pepe, and L. L. Murphy, "Hedonomics: The Power of Positive and Pleasurable Ergonomics," *Ergonomics in Design*, vol. 13, no. 1, pp. 8–14, Jan. 2005.