A FRAMEWORK FOR EVALUATING PHYSICAL HUMAN-ROBOT INTERFACE DESIGN

A Thesis Proposal

by

Aleksandar B. Bošković

Committee:

Dr. Aaron Young, Advisor
George W. Woodruff School of Mechanical Engineering
Georgia Institute of Technology

Dr. Gregory Sawicki
George W. Woodruff School of Mechanical Engineering
Georgia Institute of Technology

BACKGROUND AND STATE-OF-THE-ART

Human-robot interaction (HRI) researchers have long sought to quantify how effectively humans and robotic systems work together. Foundational efforts, such as those by Steinfeld et al. (2006) and Goodrich and Schultz (2007), provided general metrics for HRI evaluation. Building on this work, Weiss et al. (2009) developed a more targeted framework, focusing specifically on human factors of usability, social acceptance, user experience, and societal impact. While valuable, these existing evaluation approaches have primarily been applicable to teleoperated or social robots. Missing from this discourse are biomedical applications of HRI, which often require physical interfaces that enable robots to measure and transfer forces to the human body. Consequently, to the best of the proposer's knowledge, there is no existing framework that fully addresses the prolonged physical contact that is central to many medical robotic systems.

Wearable robotic exoskeletons are a prime example of medical robotic systems that require prolonged physical contact. With the immense growth of the field of wearable robotics over the last several decades, researchers have begun to recognize the need for standardized approaches to evaluating these devices. Despite the development of exoskeleton systems across industrial, medical, and military applications, researchers still lack a widely adopted framework for comparing device performance across different studies and device types (Li-Baboud et al., 2023). Nevertheless, some targeted frameworks that evaluate specific aspects of exoskeleton performance have emerged. Notably, the PoLoTAE framework proposed by Bostelman et al. (2019) provides standardized test methods for evaluating load handling in upper-limb exoskeletons, and Virts et al. (2022) later applied this framework to assess exoskeleton effects on user precision and posture.

The success of targeted frameworks like PoLoTAE demonstrates that standardized evaluation methods for specific aspects of exoskeleton performance are both feasible and valuable. Yet methods for evaluating an exoskeleton's usability remain understudied; Meyer et al. (2021) found that dedicated guidelines for evaluating the usability of wearable robotic systems remain scarce. Moreover, they observed that researchers rarely incorporate user feedback during the early stages of design, arguing that usability and ergonomics should inform design iteration from the beginning rather than serve as post-hoc considerations. These arguments were further supported by Schiele's studies on subjective and objective exoskeleton ergonomics (2009), which demonstrated how varying an exoskeleton's geometry influenced both the user's perceived comfort and the forces measured at the physical interface; thus, a poorly fitted cuff or misaligned joint can not only cause discomfort, but also influence the data recorded from sensors and outcomes of the user's task.

Beyond physical fit and comfort, user-centered design must also account for psychological factors. Stirling et al. (2024) found that users' confidence in a wearable robot affects how safely and effectively they use it, underscoring the importance of trust as a design consideration. Together, these studies illustrate the need for design and evaluation frameworks that emphasize user considerations and can be applied throughout the iteration process rather than retrospectively.

To understand how design decisions impact user considerations, some researchers have focused their attention on the physical interface where the robot and human body make physical contact. De Rossi et al. (2010) used distributed pressure sensing at lower-limb exoskeleton thigh cuffs to demonstrate that seemingly trivial changes in pad geometry and strap tension impact user comfort and transmission of force to the body. Similarly, Linnenberg and Weidner (2022) found that user comfort ratings correlated with circumferential pressure measured during upper-limb exoskeleton use. Langlois et al (2021) developed exoskeleton cuffs with integrated electromyography (EMG) sensors to provide a more complete picture of how interface design impacts muscle activity. Although these studies illustrate the value of collecting sensor data at the interface, a critical limitation persists: each study employed different approaches to collecting this data, with no consensus on how to translate findings into generalizable design principles or methods for comparing results across studies. What is needed is a standardized framework that can guide data collection at the interface and interpretation of those measurements in terms of user experience and device performance.

These interface design challenges are not unique to wearable robots. Biomechanical measurement devices, such as dynamometers, despite not being considered “robots” in the traditional sense, face remarkably similar problems at the physical interface between human and machine. Studies like Tsaopoulos et al. (2011) and Anderson et al. (2010) found that even minor misalignments between the user and the dynamometer significantly distorted measured joint torques. McDowell et al. (2012) demonstrated that slight changes in the geometry of a dynamometer handle significantly affected recorded grip strength values. Thus, it can be concluded that the accuracy and usability of biomechanical measurement devices depend on similar interface design factors that govern wearable robotics, making a physical interface design evaluation framework that spans both domains potentially feasible.

To summarize, the current landscape of human-robot interface evaluation is fragmented. General HRI frameworks have not yet been adapted and refined to contexts involving physical human-robot contact. Across the various domains involving such contact, researchers measure comfort, pressure, alignment, and usability in many different ways, making it difficult to compare findings and extract consistent design principles. However, these interfaces share the common overarching goal of safely and effectively linking human biomechanics with electromechanical systems. Recognizing this commonality presents an opportunity to unify these scattered efforts through a generalized design evaluation framework.

AIMS

This study aims to develop a framework for informing and evaluating the design of physical human-robot interfaces (pHRI), focusing on robotic systems used for mobility augmentation (i.e. exoskeletons) and measurement of human biomechanics. I hypothesize that despite differences in control architecture and intended function, these systems share fundamental design requirements for their physical interfaces.

This study will:

1. Identify relevant evaluation dimensions that can comprehensively measure the quality of a physical human-robot interface (e.g. comfort, fit, safety, trust).

2. Map each of these dimensions to performance metrics drawn from prior studies on ergonomics and human-robot interaction.

3. Determine which performance metrics are most applicable to biomedical forms of pHRI.

4. Evaluate whether this single set of metrics can effectively be applied to evaluate both wearable robotics and biomechanics measurement devices.

5. Develop and test an evaluation framework that both guides interface design and assesses interface performance across different systems.

The overarching question is whether a generalized evaluation framework is feasible or whether device-specific evaluation strategies are the only way to effectively characterize interfaces across biomedical applications.

APPROACHES

The approaches of this thesis are twofold: after conducting a literature review, the feasibility of a unified evaluation framework will be tested through case studies on biomedical robotic systems. These approaches can be further broken down into the following five steps:

1. <u>Literature Review</u>: Beginning with a structured literature review is a long-held and common theme for research on human-robot interaction (HRI), as seen in foundational works like Steinfeld et al. (2006). It is the preferred approach for identifying common trends and directions in this broad, diverse field. Existing approaches for evaluating physical human-robot interfaces will be identified by searching databases such as *IEEE Xplore* and *PubMed* using keywords like *human-robot interaction, wearable robotics, ergonomics,* and *human factors*. Steinfeld et al. and other early works will introduce common HRI metrics, while more recent reviews, such as Samarathunga et al. (2025), will explore state-of-the-art techniques for contact modeling and safety evaluation. However, since much of the HRI literature focuses on industrial domains, additional works on ergonomics, trust, and human factors, as well as studies focusing on dynamometers and wearable exoskeletons, will be used to bridge this gap into biomedical contexts.

2. <u>Performance Metric Identification</u>: The performance metrics or evaluation approaches for each selected study will be identified and organized into categories. For instance, objective

and subjective measures will be separated and grouped based on device type (exoskeleton, dynamometer, etc.); this will help identify which metrics are most adaptable across devices.

3. Framework Synthesis: Insights from the literature review will be synthesized into a proposed evaluation framework for physical human-robot interfaces, which recommends appropriate performance metrics and measurement techniques.

4. Case Studies: The scope of this proposed framework will be explored by using it to evaluate two human-interfacing robotic systems I have helped develop through my prior undergraduate research: the thigh interface of a robotic hip exoskeleton and the foot mount of a robotic dynamometer. Both systems are well within the biomedical domain, but they are distinct in that one device is wearable, and the other is not. These case studies will illustrate whether the proposed framework works across various device types or if researchers must tailor their evaluation approach on a case-by-case basis. As seen in Virts et al. (2022), case studies are an effective way to test the real-world applicability of proposed device evaluation frameworks.

5. Discussion: A final discussion will draw conclusions about the framework's generalizability, identify limitations of this study (e.g., the need for more case studies), and outline future directions for integrating standard design evaluation practices across biomedical robotic systems.

SIGNIFICANCE

The adoption of a standardized framework for evaluating physical human-robot interfaces will advance the medical robotics field in a number of ways. For researchers, it will provide a foundation for comparing interface designs across domains and enable insight from exoskeleton development to inform dynamometer design, and vice versa. This framework will also help researchers avoid design mistakes already recognized in other systems, accelerating innovation by promoting best practices across device types. For users, standardized interface evaluation will ensure that designs prioritize safety, comfort, and usability from the earliest stages of development. More broadly, this work will address a critical barrier to translating medical robots from academic prototypes into real-world systems that can be adopted by society by building users' trust and confidence in these devices.

REFERENCES

[1] Steinfeld, A., Fong, T., Kaber, D., Lewis, M., Scholtz, J., Schultz, A., et al. (2006). "Common metrics for human-robot interaction," in *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction* (New York, NY, USA: Association for Computing Machinery HRI '06), 33–40. doi:10.1145/1121241.1121249

[2] Michael A. Goodrich and Alan C. Schultz. 2007. Human-robot interaction: a survey. Found. Trends Hum.-Comput. Interact. 1, 3 (February 2007), 203–275. https://doi.org/10.1561/1100000005

[3] Weiss, Astrid & Bernhaupt, Regina & Lankes, Michael & Tscheligi, Manfred. (2009). The USUS evaluation framework for human-robot interaction. Proc. of AISB 09. 4. 11-26.

[4] Li-Baboud Y-S, Virts A, Bostelman R, Yoon S, Rahman A, Rhode L, Ahmed N, Shah M. Evaluation Methods and Measurement Challenges for Industrial Exoskeletons. Sensors. 2023; 23(12):5604. https://doi.org/10.3390/s23125604

[5] Bostelman, R. , Li-Baboud, Y. , Virts, A. , Yoon, S. and Shah, M. (2019), Towards Standard Exoskeleton Test Methods for Load Handling, WearRACon 2019, Scottsdale, AZ, US, https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=927164

[6] Virts, A. , Bostelman, R. , Yoon, S. , Li-Baboud, Y. and Shah, M. (2022), A Peg-in-Hole Test and Analysis Method for Exoskeleton Evaluation, Technical Note (NIST TN), National Institute of Standards and Technology, Gaithersburg, MD, https://doi.org/10.6028/NIST.TN.2208, https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=932548

[7] Meyer, J.T., Gassert, R. & Lambercy, O. An analysis of usability evaluation practices and contexts of use in wearable robotics. J NeuroEngineering Rehabil 18, 170 (2021). https://doi.org/10.1186/s12984-021-00963-8

[8] Schiele, André. (2009). Ergonomics of Exoskeletons: Subjective Performance Metrics. 480-485. 10.1109/IROS.2009.5354029.

[9] Schiele, André. (2009). Ergonomics of Exoskeletons: Objective Performance Metrics. 103 - 108. 10.1109/WHC.2009.4810871.

[10] Stirling, Leia & Wu, Man & Peng, Xiangyu. (2024). Measuring Trust for Exoskeleton Systems. 10.48550/arXiv.2407.07200.

[11] De Rossi, Stefano Marco Maria & Vitiello, Nicola & Lenzi, Tommaso & Ronsse, Renaud & Koopman, Bram & Persichetti, Alessandro & Vecchi, Fabrizio & Ijspeert, A.J. & Kooij, Herman & Carrozza, Maria Chiara. (2010). Sensing Pressure Distribution on a Lower-Limb Exoskeleton Physical Human-Machine Interface. Sensors. 11. 207-227. 10.3390/s110100207.

[12] Linnenberg C, Weidner R. Industrial exoskeletons for overhead work: Circumferential pressures on the upper arm caused by the physical human-machine-interface. Appl Ergon. 2022 May;101:103706. doi: 10.1016/j.apergo.2022.103706. Epub 2022 Feb 5. PMID: 35134687.

[13] Langlois K, Geeroms J, Van De Velde G, Rodriguez-Guerrero C, Verstraten T, Vanderborght B and Lefeber D (2021) Improved Motion Classification With an Integrated Multimodal Exoskeleton Interface. Front. Neurorobot. 15:693110. doi: 10.3389/fnbot.2021.693110

[14] Tsaopoulos DE, Baltzopoulos V, Richards PJ, Maganaris CN. Mechanical correction of dynamometer moment for the effects of segment motion during isometric knee-extension tests. J Appl Physiol (1985). 2011 Jul;111(1):68-74. doi: 10.1152/japplphysiol.00821.2010. Epub 2011 Apr 7. PMID: 21474701; PMCID: PMC3137545.

[15] Anderson DE, Nussbaum MA, Madigan ML. A new method for gravity correction of dynamometer data and determining passive elastic moments at the joint. J Biomech. 2010 Apr 19;43(6):1220-3. doi: 10.1016/j.jbiomech.2009.11.036. Epub 2010 Jan 4. PMID: 20047749; PMCID: PMC2849864.

[16] Mcdowell, Thomas & Wimer, Bryan & Welcome, Daniel & Warren, Chris & Dong, Ren. (2012). Effects of handle size and shape on measured grip strength. International Journal of Industrial Ergonomics. 42. 199–205. 10.1016/j.ergon.2012.01.004.

[17] Samarathunga, Piyumal & Valori, Marcello & Legnani, Giovanni & Fassi, Irene. (2025). Assessing Safety in Physical Human-Robot Interaction in Industrial Settings: A Systematic Review of Contact Modelling and Impact Measuring Methods. Robotics. 14. 10.3390/robotics14030027.

CERTIFICATE OF PROPOSAL APPROVAL

Date: 12/01/25

Student Name	Aleksandar Blaine Boskovic
gtID#	903783727
Major	Mechanical Engineering
Faculty Mentor Name	Dr. Aaron Young
Proposed Thesis Title	A Framework for Evaluating Physical Human-Robot Interface Design

Completion of Proposal

We, the below signed, hereby state our full approval of the proposal submitted by the above student in partial fulfillment of the requirements for the Research Option.

Aaron Young		*12/1/25*
Faculty Mentor (print name)	**Signature**	**Date**
Gregory S. Sawicki		*11/30/2025*
Faculty 2nd Reader (print name)	**Signature**	**Date**

LMC 4701/4702 RESEARCH WORK PLAN

PROJECT TITLE	A Framework for Evaluating Physical Human-Robot Interface Design	**ADVISOR**	Dr. Aaron Young
NAME	Aleksandar Boskovic	**APPROVAL**	*Aaron Young*

TASK TITLE	COMPLETED	PHASE 1					PHASE 2				
Literature Review & Study Preparation		AUGUST	SEPTEMBER	OCTOBER	NOVEMBER	DECEMBER	JANUARY	FEBRUARY	MARCH	APRIL	MAY
Review general HRI papers	☑	H	T	L							
Review papers on wearable robot design	☑		H	T	L						
Review papers on bioinstrumentation design	☑		H	T	L						
Write design section of exoskeleton paper	☑			H	T	L					
Write design section of dynamometer paper	☐				H/T	L					
Determine parameters to evaluate in designs	☐				H	T					
Data Collection & Proposal											
Complete annotated bibliography	☑		H	T							
Draft proposal	☑			H	T	L					
Final proposal submission	☐				H	T					
Evaluate exoskeleton design parameters	☐					H	T	L			
Evaluate dynamometer design parameters	☐					H	T	L			
Data Analysis & Thesis											
Analyze & compare/contrast case studies	☐						H	T	L		
Collect any additional required data	☐						H	T	L		
Create figures	☐							H/T	L		
Draft thesis	☐								H	T	L
Peer review of thesis	☐								H	T	L
Thesis edits	☐								H	T	L
Prepare final thesis presentation	☐									H	T

Summary paragraph:
My thesis seeks to develop a generalized framework for evaluating physical human–robot interface (pHRI) design, particularly for biomedical robotic systems such as wearable exoskeletons and biomechanical measurement devices. I aim to identify key evaluation dimensions (comfort, fit, safety, trust), map them to objective and subjective performance metrics, and test whether a unified framework can apply across device types. Through a literature review and case studies involving a hip

Note:
I am currently in the process of co-authoring mechanical design sections in other papers about the exoskeleton and dynamometer I will be using as case studies..