

**A WEB-BASED WIZARD-OF-OZ PLATFORM FOR
COLLABORATIVE AND REPRODUCIBLE
HUMAN-ROBOT INTERACTION RESEARCH**

by

Sean O'Connor

A Thesis

Presented to the Faculty of

Bucknell University

in Partial Fulfillment of the Requirements for the Degree of
Bachelor of Science with Honors in Computer Science

February 5, 2026

Approved:

L. Felipe Perrone
Thesis Advisor

Brian King
Second Reader

Alan Marchiori
Chair of the Department of Computer Science

Acknowledgments

(Draft Acknowledgments)

Contents

Abstract	8
1 Introduction	1
1.1 Motivation	1
1.2 HRISstudio Overview	2
1.3 Research Objectives	3
2 Background and Context	4
2.1 Human-Robot Interaction and Wizard-of-Oz	4
2.2 Prior Work	5
3 Related Work and State of the Art	7
3.1 Existing Frameworks	7

	3
3.2 General vs. Domain-Specific Tools	8
3.3 Methodological Critiques	8
3.4 Research Gaps	9
4 Reproducibility Challenges in WoZ-based HRI Research	10
4.1 Sources of Variability	10
4.2 Infrastructure and Reporting	10
4.3 Platform Requirements	10
5 System Design: HRISudio Platform	11
5.1 Design Goals	12
5.2 High-Level Architecture	12
5.3 Hierarchical Experimental Model	12
5.4 Visual Experiment Designer	12
5.5 Execution Interfaces	12
5.6 Robot Integration and Plugins	12
5.7 Data Management	12

	4
6 Implementation Details	13
6.1 Technology Stack	13
6.2 Technical Challenges	13
6.3 System Capabilities	13
7 Experimental Evaluation of HRISudio	14
7.1 Evaluation Goals	14
7.2 Study Design	14
7.3 Procedure	14
8 Results	15
8.1 Quantitative Results	15
8.2 Qualitative Findings	15
9 Discussion	16
9.1 Interpretation of Findings	16
9.2 Comparison to Prior Work	16
9.3 Limitations	16

	5
10 Conclusion and Future Work	17
10.1 Contributions	17
10.2 Future Directions	17
Study Materials	20
.1 Protocols	20
.2 IRB Materials	20
.3 Questionnaires	20
Technical Documentation	21
.4 Deployment	21
.5 Plugin Specification	21

List of Tables

List of Figures

Abstract

[Abstract goes here]

Chapter 1

Introduction

1.1 Motivation

To build the social robots of tomorrow, researchers must find ways to convincingly simulate them today. The process of designing and optimizing interactions between human and robot is essential to the field of Human-Robot Interaction (HRI), a discipline dedicated to ensuring these technologies are safe, effective, and accepted by the public. However, current practices for prototyping these interactions are often hindered by complex technical requirements and inconsistent methodologies.

In a typical social robotics interaction, a robot operates autonomously based on pre-programmed behaviors. Because human interaction is inherently unpredictable, pre-programmed autonomy often fails to respond appropriately to subtle social cues, causing the interaction to degrade. To overcome this, researchers utilize the Wizard-of-Oz (WoZ) technique, where a human operator—the “wizard”—controls the robot’s

actions in real-time, creating the illusion of autonomy. This allows for rapid prototyping and testing of interaction designs before the underlying artificial intelligence is fully matured.

Despite its versatility, WoZ research faces two critical challenges. First, a high technical barrier prevents many non-programmers, such as experts in psychology or sociology, from conducting their own studies without engineering support. Second, the hardware landscape is highly fragmented. Researchers frequently build bespoke, “one-off” control interfaces for specific robots and specific experiments. These ad-hoc tools are rarely shared, making it difficult for the scientific community to replicate studies or verify findings. This has led to a replication crisis in HRI, where a lack of standardized tooling undermines the reliability of the field’s body of knowledge.

1.2 HRIStudio Overview

To address these challenges, this thesis presents HRIStudio, a web-based platform designed to manage the entire lifecycle of a WoZ experiment: from interaction design, through live execution, to final analysis.

HRIStudio is built on three core design principles: disciplinary accessibility, scientific reproducibility, and platform sustainability. To achieve accessibility, the platform replaces complex code with a visual, drag-and-drop interface, allowing domain experts to design interaction flows much like creating a storyboard. To ensure reproducibility, HRIStudio enforces a structured experimental workflow that acts as a “smart co-pilot” for the wizard. It guides them through a standardized script to mini-

mize human error while automatically logging synchronized data streams for analysis. Finally, unlike tools tightly coupled to specific hardware, HRISstudio utilizes a robot-agnostic architecture to ensure sustainability. This design ensures that the platform remains a viable tool for the community even as individual robot platforms become obsolete.

1.3 Research Objectives

The primary objective of this work is to demonstrate that a unified, web-based software framework can significantly improve both the accessibility and reproducibility of HRI research. Specifically, this thesis aims to develop a production-ready platform, validate its accessibility for non-programmers, and assess its impact on experimental rigor.

First, this work translates the foundational architecture proposed in prior publications into a stable, full-featured software platform capable of supporting real-world experiments. Second, through a formal user study, we evaluate whether HRISstudio allows participants with no robotics experience to successfully design and execute a robot interaction, comparing their performance against industry-standard software. Finally, we quantify the impact of the platform’s guided execution features on the consistency of wizard behavior and the accuracy of data collection.

This work builds upon preliminary concepts reported in two peer-reviewed publications [1, 2]. It extends that research by delivering the complete implementation of the system and a comprehensive empirical evaluation of its efficacy.

Chapter 2

Background and Context

2.1 Human-Robot Interaction and Wizard-of-Oz

HRI is a multidisciplinary field dedicated to understanding, designing, and evaluating robotic systems for use by or with humans. Unlike industrial robotics, where safety often means physical separation, social robotics envisions a future where robots operate in shared spaces, collaborating with people in roles ranging from healthcare assistants and educational tutors to customer service agents.

For these interactions to be effective, robots must exhibit social intelligence. They must recognize and respond to human social cues—such as speech, gaze, and gesture—in a manner that is natural and intuitive. However, developing the artificial intelligence required for fully autonomous social interaction is an immense technical challenge. Perception systems often struggle in noisy environments, and natural language understanding remains an area of active research.

To bridge the gap between current technical limitations and desired interaction capabilities, researchers employ the WoZ technique. In a WoZ experiment, a human operator (the “wizard”) remotely controls the robot’s behaviors, unaware to the study participant. To the participant, the robot appears to be acting autonomously. This methodology allows researchers to test hypotheses about human responses to robot behaviors without needing to solve the underlying engineering challenges first.

2.2 Prior Work

This thesis represents the culmination of a multi-year research effort to address critical infrastructure gaps in the HRI community. The ideas presented here build upon a foundational trajectory established through two peer-reviewed publications.

We first introduced the concept for HRIStudio as a Late-Breaking Report at the 2024 IEEE International Conference on Robot and Human Interactive Communication (RO-MAN) [1]. In that work, we identified the lack of accessible tooling as a primary barrier to entry in HRI and proposed the high-level vision of a web-based, collaborative platform. We established the core requirements for the system: disciplinary accessibility, robot agnosticism, and reproducibility.

Following the initial proposal, we published the detailed system architecture and preliminary prototype as a full paper at RO-MAN 2025 [2]. That publication validated the technical feasibility of our web-based approach, detailing the communication protocols and data models necessary to support real-time robot control using standard web technologies.

While those prior publications established the “what” and the “how” of HRISudio, this thesis focuses on the realization and validation of the platform. We extend our previous research in two key ways. First, we move beyond prototypes to deliver a complete, production-ready software platform (v1.0), resolving complex engineering challenges related to stability, latency, and deployment. Second, and crucially, we provide the first rigorous user study of the platform. By comparing HRISudio against industry-standard tools, this work provides empirical evidence to support our claims of improved accessibility and experimental consistency.

Chapter 3

Related Work and State of the Art

3.1 Existing Frameworks

The HRI community has a long history of developing custom tools to support WoZ studies. Early efforts focused on providing robust interfaces for technical users. For example, Polonius [3] was designed to give robotics engineers a flexible way to create experiments for their collaborators, emphasizing integrated logging to streamline analysis. Similarly, OpenWoZ [4] introduced a cloud-based, runtime-configurable architecture that allowed researchers to modify robot behaviors on the fly. These tools represented significant advancements in experimental infrastructure, moving the field away from purely hard-coded scripts. However, they largely targeted users with significant technical expertise, requiring knowledge of specific programming languages or network protocols to configure and extend.

3.2 General vs. Domain-Specific Tools

A recurring tension in the design of HRI tools is the trade-off between specialization and generalizability. Some tools prioritize usability by coupling tightly with specific hardware. WoZ4U [5], for instance, provides an intuitive graphical interface specifically for the Pepper robot, making it accessible to non-technical researchers but unusable for other platforms. Manufacturer-provided software like Choregraphe [6] for the NAO robot follows a similar pattern: it offers a powerful visual programming environment but locks the user into a single vendor’s ecosystem. Conversely, generic tools like Ozlab seek to support a wide range of devices but often struggle to maintain relevance as hardware evolves [7]. This fragmentation forces labs to constantly switch tools or reinvent infrastructure, hindering the accumulation of shared methodological knowledge.

3.3 Methodological Critiques

Beyond software architecture, the methodological rigor of WoZ studies has been a subject of critical review. In a seminal systematic review, Riek [8] analyzed 54 HRI studies and uncovered a widespread lack of consistency in how wizard behaviors were controlled and reported. The review noted that very few researchers reported standardized wizard training or measured wizard error rates, raising concerns about the internal validity of many experiments. This lack of rigor is often exacerbated by the tools themselves; when interfaces are ad-hoc or poorly designed, they increase the cognitive load on the wizard, leading to inconsistent timing and behavior that can confound study results.

3.4 Research Gaps

Despite the rich landscape of existing tools, a critical gap remains for a platform that is simultaneously accessible, reproducible, and sustainable. Existing accessible tools are often too platform-specific to be widely adopted, while flexible, general-purpose frameworks often present a prohibitively high technical barrier. Furthermore, few tools directly address the methodological crisis identified by Riek by enforcing standardized protocols or actively guiding the wizard during execution. HRISudio aims to fill this void by providing a web-based, robot-agnostic platform that not only lowers the barrier to entry for interdisciplinary researchers but also embeds methodological best practices directly into the experimental workflow.

Chapter 4

Reproducibility Challenges in WoZ-based HRI Research

4.1 Sources of Variability

4.2 Infrastructure and Reporting

4.3 Platform Requirements

Chapter 5

System Design: HRISudio Platform

5.1 Design Goals

5.2 High-Level Architecture

5.3 Hierarchical Experimental Model

5.4 Visual Experiment Designer

5.5 Execution Interfaces

5.6 Robot Integration and Plugins

5.7 Data Management

Chapter 6

Implementation Details

6.1 Technology Stack

6.2 Technical Challenges

6.3 System Capabilities

Chapter 7

Experimental Evaluation of HRISudio

7.1 Evaluation Goals

7.2 Study Design

7.3 Procedure

Chapter 8

Results

8.1 Quantitative Results

8.2 Qualitative Findings

Chapter 9

Discussion

9.1 Interpretation of Findings

9.2 Comparison to Prior Work

9.3 Limitations

Chapter 10

Conclusion and Future Work

10.1 Contributions

10.2 Future Directions

References

- [1] Sean O'Connor and L. Felipe Perrone. HRISudio: A Framework for Wizard-of-Oz Experiments in Human-Robot Interaction Studies (Late Breaking Report). 2024 33rd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), 2024.
- [2] Sean O'Connor and L. Felipe Perrone. A Web-Based Wizard-of-Oz Platform for Collaborative and Reproducible Human-Robot Interaction Research. 2025 34th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), 2025.
- [3] David V. Lu and William D. Smart. Polonius: A Wizard of Oz Interface for HRI Experiments. *Proceedings of the 6th International Conference on Human-Robot Interaction*, pages 77–78, 2011.
- [4] Guy Hoffman and Cynthia Breazeal. OpenWoZ: A Runtime-Configurable Wizard-of-Oz Framework for Human-Robot Interaction. *Proceedings of the 11th ACM/IEEE International Conference on Human-Robot Interaction*, pages 117–124, 2016.
- [5] Frank Rietz and Maren Bennewitz. WoZ4U: An Open-Source Wizard-of-Oz In-

- terface for Human-Robot Interaction. *Proceedings of the 16th ACM/IEEE International Conference on Human-Robot Interaction*, pages 95–103, 2021.
- [6] E. Pot, J. Monceaux, R. Gelin, and B. Maisonnier. Choregraphe: a graphical tool for humanoid robot programming. In *RO-MAN 2009 - The 18th IEEE International Symposium on Robot and Human Interactive Communication*, pages 46–51, 2009. doi: 10.1109/ROMAN.2009.5326209.
- [7] Pettersson, John Sören and Wik, Malin. The longevity of general purpose Wizard-of-Oz tools. In *Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction*, OzCHI '15, page 422–426, New York, NY, USA, 2015. Association for Computing Machinery. ISBN 9781450336734. doi: 10.1145/2838739.2838825. URL <https://doi.org/10.1145/2838739.2838825>.
- [8] Laurel D. Riek. Wizard of Oz studies in HRI: a systematic review and new reporting guidelines. *J. Hum.-Robot Interact.*, 1(1):119–136, July 2012. doi: 10.5898/JHRI.1.1.Riek. URL <https://doi.org/10.5898/JHRI.1.1.Riek>.

Study Materials

.1 Protocols

.2 IRB Materials

.3 Questionnaires

Technical Documentation

.4 Deployment

.5 Plugin Specification