Cultural Influences on Human-Robot Interaction: Effects of Robot Appearance and Control Modes

by

Amr Hamdi

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Mechanical and Mechatronics Engineering

Waterloo, Ontario, Canada, 2025

© Amr Hamdi 2025

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

This thesis investigates how cultural differences affect human-robot interaction (HRI), with a focus on physical HRI (pHRI) in collaborative tasks. It presents two empirical studies exploring the influence of cultural background on user perceptions, comfort, and acceptance when interacting with robots. Specifically, three cultural groups were considered: Western, Middle Eastern, and Chinese.

The first study examines how the robot's physical form—humanoid (TALOS) versus non-humanoid (MOVO)—shapes user experiences among participants from the three cultural backgrounds. Results indicate that the TALOS robot tended to be perceived as more comfortable and approachable, especially during independent tasks, possibly due to its human-like features and movements. However, cultural background played a significant role in moderating these perceptions: Western and Chinese participants responded more positively to TALOS, while Middle Eastern participants showed more caution and discomfort. Physiological data, including heart rate and galvanic skin response, indicated higher stress levels during collaborative tasks across all groups, with lower stress levels typically observed when interacting with TALOS.

The second study explores how cultural background affects user preferences for robot control modes—autonomous versus human-controlled—and their responses to robot errors. Using the Sawyer collaborative robotic arm, participants performed tasks in both modes (users were deceived in believing the modes were different, while in reality, they were the same), with an intentional error introduced during one of the interactions. Most participants preferred the autonomous mode and reported higher comfort when the robot operated independently. However, Western and Chinese participants generally demonstrated higher trust in autonomous systems, whereas Middle Eastern participants tended to exhibit greater caution, particularly following errors in the human-controlled condition. Interestingly, users were generally more forgiving of mistakes in the human-controlled condition, often attributing errors to the human operator. Physiological responses supported these observations, showing increased stress during error conditions, with Western and Chinese participants recovering more quickly than Middle Eastern participants.

Together, these studies highlight the importance of considering cultural background in the design and deployment of robots in several applications, like healthcare, rehabilitation, and industrial automation. This work aims to support incorporating cultural awareness into robotic systems, leading to the future development of more inclusive, trustworthy, and user-friendly technologies.

Acknowledgements

I sincerely thank my lab mates at the A.I.R. (Active and Interactive Robotics) Lab and the S.I.R.R.L. (Social and Intelligent Robotics Research) Lab for their support. I have a special appreciation for my supervisors for their constant guidance and support. Many thanks to RoboHub for providing the space and equipment that made my work possible. A special thanks to the people who have been there with me throughout my master's, Philopatear Ashraf, Marian Bocktor, Menna Rihan, and Michael Riad. A very special thanks goes to Philopatear Ashraf, who not only pushed me to complete my master's degree but also stood by me through every step of the way. Your unwavering support, motivation, and belief in me meant more than words can express. This achievement is as much yours as it is mine.

To my beloved fiancée, Your love, patience, and endless belief in me have been my anchor throughout this journey. In moments of doubt, you reminded me of my strength. In times of exhaustion, your words lifted me. Thank you for being my constant source of light, inspiration, and peace. This milestone would not have been possible without your unconditional support — I am forever grateful to walk this path with you by my side.

Dedication

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my readers.

Table of Contents

Author's Declaration ii

Abstract iii

Acknowledgements iv

Dedication v

List of Figures xi

List of Tables xiv

List of Abbreviations xvi

1 Introduction 1

1.1 Research Questions 2

1.2 Cross-Cultural Studies in Robot Appearance and Control Modes 3

1.3 Background 4

1.4 Publications 5

1.5 Roadmap of the Thesis 6

2 Related Work **7**

2.1 Robot Form and Participant Interactions in pHRI 7

2.2 Collaborative Tasks in HRI 9

2.3 Cultural Influences in Human-Robot Interaction (HRI) 9

2.4 Autonomy vs. Human Control in HRI 11

2.5 Error Tolerance in HRI 12

2.6 Research Gaps and Emerging Questions 13

2.6.1 Cultural Differences in Physical Human-Robot Interaction (pHRI) Beyond Social/Verbal Contexts 13

2.6.2 Impact of Robot Form on Physical Collaboration in Real-World Tasks 14

2.6.3 Physiological Responses Across Cultures in pHRI 15

2.6.4 Preference for Autonomy vs. Human Control Across Cultures . . . 15

2.7 Statistical Terms and Interpretation Guidelines 15

3 The Impact of Robot Appearance in Collaborative Tasks Across Multiple Culture Groups **18**

3.1 Introduction 18

3.2 Research Questions and Hypothesis 19

3.3 Methodology 21

3.3.1 Participants 21

3.3.2 Robots 21

3.3.3 Technical Implementation 22

3.3.4 Experimental Setup 26

3.3.5 Tasks 27

3.3.6 Measures 27

3.3.7 Experimental Protocol 30

3.3.8 Data Analysis 31

3.4 Results 34

3.4.1 Interaction Analysis on Pre-Experiment and Post-Experiment Questionnaire . . . 34
3.4.2 Linear Mixed Model (LMM) Results for General Attitudes Toward Robots Scale (GAToRS) Scales . . . 37
3.4.3 Effect Size Analysis for Linear Mixed Models . . . 39
3.4.4 Spearman Correlation Analysis . . . 40
3.4.5 Wilcoxon Signed-Rank test-Cultural Compariosns . . . 44
3.4.6 Mann-Whitney U Test — Cultural Comparisons . . . 44
3.4.7 Physiological data analysis . . . 47
3.4.8 Body Movements and Posture Analysis . . . 52
3.4.9 Qualitative Feedback by Cultural Groups . . . 55
3.5 Discussion . . . 60

4 The Role of Culture in Perceptions of Control Modes in pHRI 65
4.1 Introduction . . . 65
4.2 Methodology . . . 68
4.2.1 Participants . . . 68
4.2.2 Robot . . . 69
4.2.3 Technical Implementation . . . 70
4.3 Trajectory Generation and Execution . . . 70
4.4 Real-Time Visualization . . . 70
4.5 Intentional Error Injection . . . 70
4.6 Cartesian Impedance Control . . . 71
4.6.1 Control Modes . . . 72
4.6.2 Experimental Setup . . . 73
4.6.3 Tasks . . . 74
4.6.4 Measures . . . 75
4.7 Experimental Protocol . . . 76

4.8 Results . . . 78
4.8.1 LMM Analysis for each Culture group . . . 78
4.8.2 Western Participants . . . 83
4.8.3 Chinese Participants . . . 85
4.8.4 SAM Post and Pre Task Comparisons . . . 87
4.8.5 Physiological data results . . . 90
4.8.6 Body Pose Data . . . 94
4.8.7 Mode Preference and Trust by Culture . . . 99
4.9 Discussion . . . 101
4.10 Limitations . . . 104

5 Conclusion **105**
5.1 Summary of Findings . . . 105
5.2 Contribution to Knowledge . . . 108
5.3 Limitations . . . 108
5.3.1 Participant Selection and Representation: . . . 109
5.3.2 Methodological Constraints: . . . 109
5.4 Future Work . . . 109

References **111**

APPENDICES **123**

A Questionnaires **124**
A.1 Screening Questionnaire for Study 1 and 2 . . . 124
A.2 Demographic Questions for study 1 and 2 . . . 125
A.3 Pre-Experiment Questionnaires for Study 1 . . . 126
A.4 Task Related Questionnaires for Study 1 . . . 126
A.5 Open Questions for Study 1 . . . 127
A.6 Pre-Experiment and Familiarity Questions from Study 2 . . . 128
A.7 Open Questions for Study 2 . . . 128

B Ethics Approvals **129**

List of Figures

1.1 The CARESSES protocol [78] is a project that aims to take culture adaptation into HRI . . . 2

3.1 Kinematic Model of JACO 7-DOF Arm . . . 23

3.2 Kinematic Model of PAL TALOS Humanoid . . . 25

3.3 a) Overview of the experiment setup with the table placement and how the participants are standing (b) MOVO during Task 2 Handover (c) TALOS during Task 1 Pick & Place. Cameras are positioned to capture participants' body movements. . . . 26

3.4 Sensor placement on the particpants . . . 28

3.5 Photoplethysmography (PPG) signal with detected peaks. The peaks correspond to detected heartbeats, which were used to compute the interbeat intervals and subsequently estimate the heart rate. . . . 32

3.6 MediaPipe Keypoints . . . 33

3.7 Sperman Correlation analysis for each robot . . . 41

3.8 Sperman Correlation analysis for Middle Eastern Culture Category . . . 42

3.9 Sperman Correlation analysis for Western Culture Category . . . 42

3.10 Sperman Correlation analysis for Chinese Culture Category . . . 43

3.11 Comparison across all the GAToRS scales (P+ (Comfort and Enjoyment), P- (Unease and Anxiety) , S+ (Societal Optimism), S- (Societal Concerns), for each culture and robot . . . 45

3.12 Task Metric distribution across robots . . . 46

3.13 Comparison of User Position, HR, and GSR across robot type . . . 48

3.14 Comparison of Heart Rate (HR) across robot type and culture category . . 49
3.15 Comparison of HR across robot type and culture category 51
3.16 Comparison of different body poses across robot type and culture category for Pick & Place 53
3.17 Comparison of different body poses across robot type and culture category for Handover. 53

4.1 The Sawyer Sdk Robot arm 69
4.2 The red line is what the participant see displayed on the screen the black line is the actual trajectory 71
4.3 The following represents the Task done by the participants, where task (a) is a test task in which they test the system to see their ideal positioning . 73
4.4 The overview of the experiment setup displaying the robotics arm , screen, and the two cameras. 74
4.5 Comparison across all the GAToRS scales (P+ (Comfort and Enjoyment), P- (Unease and Anxiety) , S+ (Societal Optimism), S- (Societal Concerns) for each mode 79
4.6 The Task question data distribution among culture groups 80
4.7 The Perceived safety data distribution among culture groups 80
4.8 The SAM scales data distribution among culture groups for Pre Task, Task 1, 2, and 3 89
4.9 The SAM scales data distribution among culture groups for Pre Task, Task 1 , 2, and 3 90
4.10 Realtive Heart Rate(%) across the Culture categories 93
4.11 Realtive GSR(%) across the Culture categories 94
4.12 Boxplots of body pose variables across cultural categories for Task 1 96
4.13 Boxplots of body pose variables across cultural categories for Task 2 (Task with Induced Error). 97
4.14 Boxplots of body pose variables across cultural categories for Task 3. . . . 98
4.15 Mode preference by cultural category: Number of participants selecting autonomous, human-controlled, or neither mode. 100

4.16 Trust in mode by cultural category: Number of participants who expressed trust in each control mode. 100

List of Tables

3.1 Participants Summary 21
3.2 Grouped ANOVA Results for Significant Effects 35
3.3 **Effect** refers to the interaction term tested in ANOVA that significantly influenced the dependent variable. **p** values indicates statistical significance when $p < .05$ 35
3.4 LMM Results for GAToRS Scales 38
3.5 LMM Effect Size Summary for Perception Models 39
3.6 Wilcoxon Rank-Sign Test Results for MOVO vs. TALOS 44
3.7 Mann-Whitney Test Results for MOVO and TALOS 47
3.8 Grouped Wilcoxon Ranked Test Results for Heart Rate, Galvanic Skin Response (GSR), and User Position 48
3.9 Wilcoxon Ranked Test Results for Heart Rate Across Cultures 50
3.10 Wilcoxon Ranked Test Results for Galvanic Skin Response Across Cultures 51
3.11 Wilcoxon Ranked Test Results for Torso and User Position for MOVO vs TALOS 52
3.12 Mann-Whitney Test Body Posture A Cross Culture - MOVO 54
3.13 Mann-Whitney Test Body Posture A Cross Culture - TALOS 55

4.1 Participants Summary 69
4.2 LMM Results for Middle Eastern Model – P+ 81
4.3 LMM Results for Middle Eastern Model – P- 81
4.4 LMM Results for Middle Eastern Model – S+ 82

4.5 LMM Results for Middle Eastern Model – S- 82
4.6 LMM Results for Western Model – P+ 83
4.7 LMM Results for Western Model – P- 84
4.8 LMM Results for Western Model – S+ 84
4.9 LMM Results for Western Model – S+ 85
4.10 LMM Results for Chinese Model – P+ (Positive Perceptions) 86
4.11 LMM Results for Chinese Model – P+ 86
4.12 LMM Results for Chinese Model – S+ 87
4.13 LMM Results for Chinese Model – S) 88
4.14 Wilcoxon Test Results for Autonomous Mode – SAM Comparisons 91
4.15 Wilcoxon Test Results for Human-Controlled Mode – SAM Comparisons . 92
4.16 Heart Rate Differences Between Autonomous and Human-Controlled Sessions 93
4.17 Comparison of Body Pose Variables Across Mode, Culture, and Sessions . 95
4.18 Mann-Whitney Test Results for Cultural Comparisons in Autonomous Mode 96
4.19 Mann-Whitney Test Results for Cultural Comparisons in Human-Controlled Mode 98

A.1 Screening Questionnaire for Both Studies 124
A.2 Demographic Information Collected 125
A.3 Pre-Interaction Comfort and Expectation Survey for Study 1 126
A.4 Task Question for Study 1 126
A.5 Familiarity and Expectations 128

List of Abbreviations

GAToRS General Attitudes Toward Robots Scale viii, xi, xii, xiv, 27, 29–31, 34, 35, 37, 38, 44, 45, 75–77, 79, 105

GSR Galvanic Skin Response xiv, 3, 13, 15, 18–20, 29–31, 34, 47, 48, 50, 51, 61, 62, 68, 76, 77, 90, 92, 94, 103, 106, 107

HR Heart Rate xii, 3, 13, 15, 18–20, 29–31, 34, 47–49, 51, 61, 62, 68, 76, 90, 91, 106, 107

HRI Human-Robot Interaction vii, 1, 2, 4, 5, 7–9, 11–13, 15, 19, 29, 32, 66, 106, 108–110

LMM Linear Mixed Model viii, xiv, 34, 37–39

pHRI Physical Human-Robot Interaction vii, 1–3, 11, 13–15, 18, 67, 101, 105, 106, 108–110

PPG Photoplethysmography xi, 29–32, 77

SAM Self - Assessment Mankin 75–77, 87, 88, 104

sHRI Social Human-Robot Interaction 18, 108

TIPI Ten-Item Personality Inventory 75–77

Chapter 1

Introduction

As robots become increasingly embedded in human environments, understanding the dynamics of Human-Robot Interaction HRI is essential for their effective integration. Robots are now present in various domains, including healthcare, education, industrial automation, and domestic assistance, where they enhance productivity, efficiency, and user experiences [86]. The field of HRI focuses on optimizing these interactions to ensure seamless collaboration between humans and robots, making it a fundamental research area for advancing robotics technology [14].

One of the critical yet often underexplored aspects of HRI is the role of culture in shaping human perceptions and interactions with robots. Culture influences fundamental aspects of human behavior, including trust, expectations, communication styles, and emotional responses [64]. Studies have shown that cultural background significantly impacts how people perceive and engage with robots. For example, Western users tend to exhibit higher initial trust and likability toward robots than East Asian users [44], while collectivist cultures show greater comfort when interacting with multiple robots rather than a single robot [61, 44]. Such variations highlight the need for culturally adaptive robotic designs that accommodate different societal expectations and interaction patterns [77].

Despite the growing interest in cultural influences on HRI, including different applications and demographics [78], most research has focused on social robots - designed primarily for communication and companionship - rather than robots that require direct physical collaboration with humans [15]. Physical human robot interaction pHRI presents unique challenges, as it involves shared physical space, coordination, and trust in robot actions [23]. Cultural attitudes toward personal space, touch, and collaborative work dynamics can significantly impact pHRI outcomes [79]. However, cross-cultural studies in this area

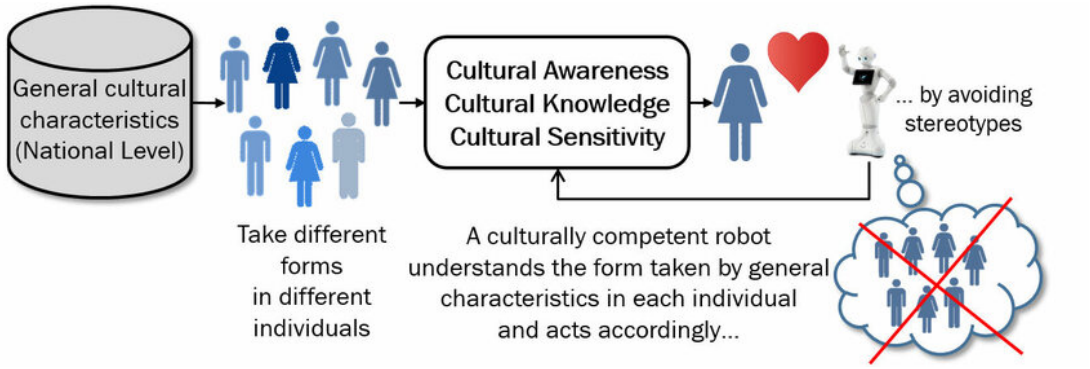


Figure 1.1: The CARESSES protocol [78] is a project that aims to take culture adaptation into HRI

remain limited, creating a gap in understanding how users from different cultural backgrounds interact with physically assistive robots [64].

Recent systematic reviews indicate that while social robots have been widely studied in various cultural contexts, relatively few studies have examined pHRI across cultures [64, 36]. For instance, a review by [80] identified 22 studies focusing on socially assistive robots in healthcare settings, but very few considered cross-cultural variations [79]. Similarly [103], conducted a study comparing Chinese, Korean, and German participants in human-robot interactions, revealing significant cultural differences in trust and engagement [60]. These findings suggest that cultural factors play a crucial role in shaping user experiences with robots, but more research is needed to apply these insights to physically collaborative robotics [77].

1.1 Research Questions

This thesis addresses the following research questions to explore the cultural influences in HRI:

I. **Robot Form Factors and Physical Interaction Acceptance:** How do robot form factors influence user acceptance and comfort during physical interactions with the robots?

II. **Cultural Influences on Robot Form Acceptance:** In a collaborative physical task, how does the robot's form factor—such as being humanoid vs. non-humanoid,

its size, and limb design—influence the efficiency of the collaboration, the human's trust and comfort, and overall task safety?

III. **Physiological Response Differences:** How do physiological responses (e.g., HR, GSR) during physical human-robot interactions vary across different cultural backgrounds, and how can these differences help shape robotic systems to suit varied cultural preferences?

IV. **Preference for Autonomy vs. Human Control Across Cultures:** How do cultural differences influence user preferences for autonomous versus human-controlled robotic systems in the presence of errors, and what implications do these preferences have for the design and deployment of robots in multicultural environments?

These research questions address critical gaps in understanding how culture shapes pHRI, focusing on user behavior, robot appearance, physiological responses, and control mode preferences. A detailed discussion of these questions, including their formulation and relevance to existing literature, is provided in Chapter 2

1.2 Cross-Cultural Studies in Robot Appearance and Control Modes

To address the research questions, this thesis presents two empirical studies investigating the impact of cultural differences in pHRI:

- **Cross-Cultural Interaction with Robots with Different Appearances:** This study examines how participants from diverse cultural backgrounds interact with two physically collaborative robots—a humanoid robot (TALOS) and a non-humanoid robotic platform (MOVO). The study assesses the effect of a robot's physical shape on general participants and the cultural influences on user preferences, comfort levels, and perceptions of robot form and functionality in physically interactive tasks such as pick-and-place and handover operations.
- **Cultural Responses to Robot Control Modes and Error Behaviors:** This study explores how users from different cultural backgrounds respond to robot errors and varying control modes (autonomous vs. human-controlled) during physical collaboration tasks. This study provides information on robot user perceptions based on control modes.

The findings of these studies contribute to the broader discourse on culturally aware human-robot interaction.

1.3 Background

HRI is deeply influenced by the cultural context of its users. It is essential to recognize the role of culture, which encompasses shared values, beliefs, norms, and behavioral patterns learned from social environments rather than inherited biologically or defined strictly by nationality. According to Hofstede [50], culture is "the collective programming of the mind which distinguishes the members of one group or category of people from another." This definition highlights that culture transcends legal or geopolitical boundaries and reflects learned social behaviors shaped by upbringing, education, community norms, and media exposure. Thus, individuals from the same nationality may differ culturally, and individuals of different nationalities may share cultural values if they have similar life experiences.

In this thesis, we examined the influence of culture through the lens of three broad cultural groups: Western, Middle Eastern, and Chinese. These groups were selected to represent a spectrum of cultural orientations, including individualism vs. collectivism, attitudes toward authority, communication styles, and exposure to robotic technologies. These groups were selected based on participants' self-declared cultural background and to represent a spectrum of orientations grounded in Hofstede's Cultural Dimensions Theory [50] , including individualism vs. collectivism, attitudes toward authority (power distance), communication styles, and exposure to robotic technologies Canada is a highly multicultural country, and institutions such as the University of Waterloo host students and researchers from diverse cultural backgrounds, providing a unique opportunity to examine these differences in a controlled yet inclusive environment.

Western cultures (e.g., Canada, USA, European descent) are often characterized by individualism, low power distance, and a task-oriented communication style [50, 98]. Prior studies suggest that individuals from Western backgrounds tend to be early adopters of new technologies and may display higher initial trust in autonomous systems [62, 75].

Middle Eastern cultures (e.g., Egypt, Jordan, Saudi Arabia) are generally more collectivist, with a strong emphasis on family, social harmony, and traditional authority structures [50, 6]. While technology adoption is increasing in these regions, attitudes toward robots—particularly in physically collaborative scenarios—can be more cautious due to cultural preferences surrounding trust, safety, and physical space [99]. Importantly, Middle Eastern perspectives remain significantly underrepresented in current HRI research,

especially in studies focusing on HRI, making their inclusion in this work both timely and necessary.

Chinese culture eg. (China, Hong Kong), reflects high collectivism, hierarchical social structures, and an emphasis on harmony and long-term orientation [50, 74]. Despite rapid technological advancement, Chinese participants may exhibit nuanced trust patterns in HRI, often shaped by social context, group affiliation, and situational factors [33].

Recognizing these cultural dimensions is essential for designing robots that are not only functional but also socially and emotionally acceptable across diverse user populations. This thesis highlights that cultural background is considered independent of national identity and is assessed based on people's lived experiences, cultural exposure, and self-identified affiliation with cultural norms and values.

1.4 Publications

To lay the groundwork for this research, we investigated the impact of robot physical appearance and user culture on human-robot interaction (HRI) in collaborative physical tasks, with findings presented in Chapter 3. The findings from Chapter 3 are being prepared for publication in two separate articles. The first, focusing on general trends in user perception and physiological responses to robot morphology (independent of cultural differences), has been submitted to IEEE Robotics and Automation Letters (RAL). The second, examining cross-cultural differences in perceptions of robot appearance and physical interaction, is being prepared for submission to Advanced Robotics.

To further explore cultural influences in physical human-robot interaction (pHRI), we conducted a second user study focused on user responses to robot control modes and error perception, detailed in Chapter 4. This work examines how cultural background affects trust, comfort, and acceptability in response to autonomous and human-controlled robots under induced error conditions, providing insights for developing culturally aware robotic systems.

Each user study was designed as a standalone experiment to address distinct aspects of cultural influences in physical human-robot interaction (pHRI). Together, they offer valuable insights into the interplay between robot appearance, control modes, and user culture. Chapter 5 concludes the thesis by summarizing the contributions, highlighting critical gaps in the literature, and suggesting future research directions for culturally inclusive and trustworthy robotic systems

1.5 Roadmap of the Thesis

This thesis is organized into five chapters, each highlighting specific aspects of cultural influences in physical Human-Robot Interaction (pHRI):

Chapter 2: Related Work reviews the existing literature on robot appearance, cultural differences, robot control modes, and error handling in human-robot interactions. It identifies gaps in prior research and formulates specific research questions the thesis addresses.

Chapter 3: The Impact of Robot Appearance in Collaborative Tasks Across Multiple Cultural Groups presents an empirical study examining how different robot appearances—humanoid versus non-humanoid—influence user comfort, trust, and physiological responses across Western, Middle Eastern, and Chinese cultural groups. The study's methodology, key findings, and limitations are discussed.

Chapter 4: The Role of Culture in Perceptions of Control Modes in pHRI details another empirical study investigating how cultural background shapes user preferences and trust in autonomous versus human-controlled robot modes, particularly during error conditions. The methodology, results, and implications for robot design are analyzed in-depth.

Chapter 5: Conclusion integrates findings from both empirical studies, revisiting the initial research questions and aims presented in Chapter 2. It summarizes the main contributions, acknowledges research limitations, and offers directions for future work to support the design of robots that respect and accommodate cultural diversity.

Chapter 2

Related Work

2.1 Robot Form and Participant Interactions in pHRI

Former studies showed that a robot's appearance should match its functionality to avoid misdirecting the user's expectations of the robot [38]. It should demonstrate the ability to comprehend and adapt to its current situation [19], and its behavior should be legible for the user [55]. Physical appearance can influence HRI. Studies suggest that anthropomorphic robots, such as humanoids, may be perceived as more comfortable and approachable, potentially due to their human-like features and movements, which align with human expectations of motion and appearance [16, 40]. Similarly, a recent review reported a "general consensus that human-like robots are perceived to be more appropriate for human-like occupations" and are seen as more likable and intelligent than animal-like or machine-like robots [36]. These characteristics enhance user acceptance, particularly in collaborative and social tasks. In contrast, non-anthropomorphic robot designs often prioritize functionality and simplicity, aligning with industrial settings where efficiency is prioritized over social interaction. However, this trend may not apply universally, particularly for collaborative robots designed for human-robot interaction [81, 106]. Despite these findings, limited research directly compares how physical appearance impacts user perception during physical assistance tasks. More human-like (humanoid) robots tend to elicit more positive social responses. For example, [95] found that participants evaluated robots more favorably and even desired more contact with them when they have human-like resemblance.

A study by [82] investigated the impact of robotic aesthetics on anticipated human trust using the Canbot U03 robot, involving 74 participants from diverse global regions. The findings revealed that anthropomorphic features could both increase and decrease trust

depending on initial user inclinations, with physical appearance sometimes overriding the correctness of the robot's responses.Another study by [68] explored emotion perception in robot faces using the Furhat robot platform, with 305 participants in two conditions (full-face and eye-region). It found that robots with human-like faces express emotions comparably to humans when the full face is visible, but recognition accuracy drops significantly for mechanical-looking robots when only the eye region is shown.

Factors like noise level, prior experience with robots, or task difficulty can affect the result of HRI studies [49]. For instance, gender can also play a significant role in shaping perceptions of robot interactions, as proxemic zones—defined as the physical distance individuals maintain in social interactions—tend to be larger for females compared to males. This difference may be influenced by relative heights between the human and the robot, which can affect the sense of personal space and comfort [84]. Studying the acceptability of robots as interaction partners also requires considering participants' use of technology and pet ownership. A fervent technology user may be more inclined to accept robots in their daily life [81], while a pet owner may exhibit higher trust towards a robotic assistant due to a predisposition for forming bonds with non-human entities [49].

People often attribute greater mental capacity, agency, and intelligence to robots with human-like features [45, 20], which can increase users' trust and acceptance up to a point. However, designers must be aware of the uncanny valley effect: if a robot appears too human-like yet imperfect, it can trigger discomfort [70, 84]. Indeed, studies suggest that slightly cartoonish or simplified human forms may maximize comfort – one consensus is that "cartoon-like faces or faceless robots are most appropriate in terms of perceived comfort", while highly realistic humanoids risk eeriness [36]. Thus, the optimal robot form may balance human resemblance with simplicity to gain user trust and usability.

In physical collaboration scenarios such as pick-and-place tasks or object handovers, the robot's physical form significantly impacts interaction quality and user experience [37]. While much of the existing research on handovers emphasizes motion timing and safety, robot form also plays a crucial role [17]. Studies have shown that more human-like timing in robot motions—such as smooth, anticipatory handovers or synchronized pauses—can enhance user trust, predictability, and perceived safety, especially in collaborative tasks. For example, predictable and legible motion improves human understanding of robot intent, reducing hesitation and improving coordination [29]

For instance, research by [69] indicates that anthropomorphic robot features, such as a head with eyes, can enhance communicative gaze cues, improving user perception of timing and friendliness during handovers. Humanoid robots or robotic arms equipped with facial features or heads tend to offer more intuitive interactions by mimicking natural human

behaviors, including eye contact and head movements [1].

Conversely, faceless industrial arms often rely on alternative signaling methods, such as auditory cues or visual indicators, to convey their intent [48]. Although direct comparative experimental research between humanoid and non-humanoid robots performing identical tasks remains scarce, related studies suggest humanoid robots frequently result in higher user engagement and perceptions of naturalness in cooperative tasks like assembly or tool handovers [90].

2.2 Collaborative Tasks in HRI

Task complexity further moderates user responses in HRI, with collaborative tasks like handovers requiring close coordination and shared goals, often leading to higher cognitive load and stress [2]. Independent tasks, where the person and the robot work in parallel toward separate but related goals without direct collaboration or shared control, are less demanding but still involve interaction in the robot's presence, which can influence user comfort and engagement [46, 49]. Traditionally, comfort has mostly been measured via subjective questionnaires, which are prone to biases [27, 88]. Studies have shown that objective ones, such as physiological data and body behavior (e.g. user posture and movement patterns) can complement questionnaires [97, 53].

While task complexity has been examined in isolation, its interaction with robot design—and the combined effect on the physiological and subjective responses of participants—remains underexplored.

2.3 Cultural Influences in HRI

Culture significantly influences human-robot interaction HRI, shaping how individuals perceive, accept, and interact with robots. The commonly assumed East-versus-West contrast—characterizing Eastern countries like Japan as robot-friendly and Western countries as skeptical—is overly simplistic [109]. Studies by [10] [9] demonstrate that Japanese participants can exhibit neutral or even negative attitudes toward robots, whereas American participants sometimes express comparatively positive attitudes. For instance, a cross-national survey indicated that American participants held the most pro-robot attitudes, challenging stereotypes regarding Japanese affinity for robots [65]. A study by [101] compared speed adaptation in HRI tasks between Japanese and Italian participants, finding

no cultural differences in motor resonance. Both groups adapted their movement speeds to match the robot, suggesting some low-level interaction dynamics may be universal. Another study [43] showed Australians had more positive perceptions of a Japanese female robot and rated it as more anthropomorphic than Japanese participants after interaction. This underscores the role of cultural background in shaping responses to robot design.

Cultural background also shapes preferences regarding robot form and interaction styles. For example, Korean users strongly prefer humanoid robots, American users favor mechanical-looking robots, and Turkish participants show no strong preference either way [62]. Interestingly, Japanese users have reported less positive feelings toward humanoid or android robots than individuals from other cultures, possibly due to distinctive portrayals of robots in their media [75]. Anthropomorphic features notably influence robot perceptions in Japan, underscoring the impact of cultural expectations informed by popular culture [18].

Interaction styles and trust dynamics also reflect deeper cultural values, such as individualism and collectivism. A study by [33] found that Americans were more willing to trust and comply with robots compared to Chinese participants; However, when robots were presented as part of the participant's " ingroup "—for example, by emphasizing shared team membership, cultural affiliation, or mutual goals—trust among Chinese participants increased significantly. This finding highlights the nuanced interplay between cultural values and dynamics.

Interpretations of robot social cues and proxemics also differ culturally. Chinese users, for example, interpreted robotic facial expressions using a broader emotional vocabulary compared to Japanese users [99]. Preferences for robotic faces similarly vary, with some Japanese older adults finding cartoon-like faces comforting, while certain American users prefer faceless robots to avoid ambiguity [64].

Physical interactions, such as collaborative tasks or robot assistance, have received limited cross-cultural investigation. An exception is a Middle Eastern research indicating that Gulf Arab participants responded more positively to robots displaying culturally congruent cues compared to participants from African backgrounds [112, 64]. Contextual factors further influence acceptance; Asian participants showed greater acceptance of robots in medical scenarios compared to Middle Eastern or Western counterparts [5]regarding personal space significantly affect robot interactions, as evidenced by differences in proxemics preferences between Arabic and German participants, aligning with their respective cultural communication norms [32, 57].

A review by [64] synthesized 50 studies on cultural influences in HRI, finding that cultural background significantly affects users' expectations, attitudes, and behaviors. It

noted that culturally sensitive robot designs promote acceptance, but more significant exposure to robots doesn't always lead to positive attitudes, and communication styles (implicit vs. explicit) vary by culture.

Despite recognizing culture's role, the nuances of intra-cultural variations and specific socio-cultural mechanisms remain underexplored, particularly concerning pHRI. Although the literature recommends designers " localize " robot behaviors, precise guidelines for adjusting form, behavior, or interaction protocols based on cultural nuances still require deeper investigation [108].While physiological measures are increasingly used to assess user responses in HRI, cross-cultural research in this area is limited. A study [67]on joint attention with robot faces suggests cultural impacts on engagement, potentially influencing physiological responses. Cultural differences in joint attention and engagement in mutual gaze with a robot face. Future research should address these gaps, offering more granular insights into culturally informed robot design. This gap is especially evident in physical human-robot interaction settings, which involve direct collaboration: most prior cross-cultural HRI studies examine social or service contexts (often via surveys or short interactions), rather than hands-on collaborative tasks.

2.4 Autonomy vs. Human Control in HRI

The preference for autonomous versus human-controlled robots has been a central topic in HRI research, with studies exploring how users perceive and interact with robots based on their level of autonomy. A study by Hinds et al. [47] conducted a seminal study on this topic, examining how people perceive robots in collaborative tasks. They found that participants preferred human-controlled robots in tasks requiring high levels of precision, trust, and decision-making, as they felt more confident in the human operator's ability to handle complex or unpredictable situations. In contrast, autonomous robots were preferred for repetitive, mundane, or well-defined tasks, where their efficiency and consistency were valued.

This preference for human control in high-stakes tasks is supported by the results found in the stud by Hancock et al.[42], who conducted a meta-analysis of trust in automation. They found that users are more likely to trust human-controlled systems in situations where errors could have significant consequences, as they perceive humans as more capable of adapting to unexpected challenges. However, this preference may vary across cultures, as cultural norms influence perceptions of autonomy and control.Insights from the autonomous vehicle domain suggest that cultural factors influence acceptance of autonomous systems. For instance, Chinese participants were more receptive to AV decisions, while

Japanese participants were more cautious, rejecting decisions that might interfere with others

For example,in a study conducted by Eyssel et al. [34] explored cultural differences in the perception of robot autonomy, comparing German and Japanese participants. They found that German participants who value individualism and personal agency, were more comfortable with autonomous robots, perceiving them as independent and efficient. In contrast, Japanese participants, who emphasize social harmony and interdependence, preferred robots that could be controlled or guided by humans, as this aligned with their cultural values of collaboration and shared responsibility [31].

These findings suggest that cultural background influences preferences for autonomy versus human control in HRI. However, there is limited research on how these preferences are affected by the presence of errors or task complexity, particularly across different cultural groups.

2.5 Error Tolerance in HRI

The impact of errors on user trust and preference in HRI has been widely studied, with researchers examining how users respond to robotic mistakes and how these responses vary across contexts.A study by [89] conducted a series of experiments to investigate how errors affect user trust and acceptance of robots. They found that errors significantly reduced user trust, particularly when the errors were perceived as avoidable or when they occurred in critical tasks. However, they also found that users were more forgiving of mistakes when the robot demonstrated transparency in explaining its actions or exhibited human-like behaviors, such as apologizing for mistakes.

Cultural differences also play a role in how errors are perceived and tolerated. A Study by Rau et al.[85] compared Chinese and German participants' responses to errors in human-robot collaboration tasks. They found that Chinese participants were more tolerant of errors, attributing them to external factors or situational constraints. In contrast, German participants were more critical and likely to blame the robot for its mistakes. This difference was linked to cultural variations in attribution styles, with collectivist cultures (e.g., China) emphasizing situational factors and individualist cultures (e.g., Germany) focusing on personal responsibility.

Moreover,the following stud by [93] systematically examines how human trust dynamics and system design are influenced by robot errors. Utilizing a three-stage survey, the research reveals that participants' trust in robotic technologies significantly increased when

robots acknowledged their errors or limitations. This finding underscores the importance of error acknowledgment in enhancing user trust and suggests design considerations for more empathetic and reliable robotic systems. [35] investigates how various error types affect trust in human-robot interaction through two laboratory studies involving 160 participants. The results indicate a general effect of errors on trust levels.

Similarly, [76] explored how Japanese and British participants responded to robot errors in social interaction tasks. They found that Japanese participants were more likely to attribute errors to the robot's lack of social awareness, while British participants focused on the robot's technical limitations. These findings suggest that cultural background influences not only the tolerance for errors but also the interpretation of their causes.

Despite these insights, there is a lack of research on how cultural differences in error tolerance interact with preferences for autonomy versus human control. Additionally, few studies have examined the physiological responses (e.g., HR(Heart Rate) and GSR (Galvanic Skin Response) associated with error perception across cultures, which could provide deeper insights into the emotional and cognitive processes underlying these differences.

2.6 Research Gaps and Emerging Questions

In this section, we identify key research gaps corresponding to the earlier themes and propose specific research questions. These gaps are presented as subsections, outlining what is missing or inconclusive from prior work, followed by research questions to be addressed in this thesis. The intent is to highlight opportunities for deeper investigations, such as cultural nuances in physical interaction, the role of robot form in real-world collaboration effectiveness, cross-cultural differences in handling robot errors, and improved methods for measuring human responses in pHRI.

2.6.1 Cultural Differences in pHRI Beyond Social/Verbal Contexts

Much of the cross-cultural HRI research has focused on social interaction paradigms, such as the appropriateness of a robot's speech politeness or survey-based attitude comparisons across countries. Far fewer studies have examined cultural differences in the context of direct physical collaboration. While there are anecdotes of differences—such as [92] variations in proximity preferences between Italian and Dutch children interacting with

robots—these studies often lack explicit cultural controls and are based on incidental observations. Moreover, many cross-cultural studies suffer from varying methodologies and small, non-representative samples, making generalization challenging. For example, a review by [64] noted that the field is hindered by heterogeneous approaches and often low statistical power when comparing cultures. Most current knowledge about the cultural impact of pHRI comes from self-reported comfort levels or hypothetical scenarios rather than empirical task performance comparisons. In summary, we lack detailed, experimental insights into how culture influences physical collaboration dynamics with robots.

Research Question I: How do users from different cultural backgrounds behave and feel during physical interactions with robots?

2.6.2 Impact of Robot Form on Physical Collaboration in Real-World Tasks

While many studies have investigated how robot form (e.g., humanoid vs. non-humanoid) affects user perceptions, less research has isolated the effect of form on objective collaboration outcomes in realistic task settings. There is consensus that human-like robots are seen as more suitable for social roles and are anthropomorphized more [30], but it remains unclear how those perceptions translate to actual collaborative performance, error rates, or safety incidents during physical tasks. To the best of our knowledge, there are limited studies that directly compare different robot forms in physical collaborative tasks, which underscores the importance of our research

Additionally, certain form factors, like very large robots or those with unconventional shapes, have been understudied in pHRI; most work utilizes humanoids or commercial arms that are human-sized [102]. As robots enter domains like construction or healthcare, form factors will diversify and could impact collaboration unexpectedly. Therefore, the impact of robot morphology on the practical efficacy and safety of human-robot physical teamwork is not fully understood. It is one thing to survey users on trust in a humanoid robot, and another to measure how form affects actual task performance and bi-directional trust over time.

Research Question II: In a collaborative physical task, how does the robot's form factor—such as being humanoid vs. non-humanoid, its size, and limb design—influence the efficiency of the collaboration, the human's trust and comfort, and overall task safety?

2.6.3 Physiological Responses Across Cultures in pHRI

The use of physiological measures to assess cultural differences in HRI is still an under-explored area. While physiological indicators like (HR) and galvanic skin response (GSR) have been utilized to assess cognitive load and emotional states during human-robot interactions, more research is needed to understand how these measures vary across different cultural groups. Such understanding could provide objective data on user stress levels, engagement, and overall comfort during interactions with robots, leading to more tailored and effective robot designs.

Research Question III: How do physiological responses (e.g., HR, GSR) during physical human-robot interactions vary across different cultural backgrounds, and how can these differences help shape robotic systems to suit varied cultural preferences?

2.6.4 Preference for Autonomy vs. Human Control Across Cultures

Although the preference for autonomous versus human-controlled robots has been studied, there is a lack of research on how this preference varies across different cultural contexts, especially in the presence of errors. Cultural norms and values may influence comfort levels with automation and control, impacting the acceptance and effectiveness of robotic systems in various settings.

Research Question IV: How do cultural differences influence user preferences for autonomous versus human-controlled robotic systems in the presence of errors, and what implications do these preferences have for the design and deployment of robots in multicultural environments?

2.7 Statistical Terms and Interpretation Guidelines

This section provides an overview of key statistical terms and metrics used in Chapters 3 and 4. Understanding these concepts is essential for interpreting the results and assessing the significance and strength of the findings.

P-value

The *p-value* is a probability that measures the evidence against a null hypothesis. Formally, it is defined as the probability of obtaining a result equal to or more extreme than what was actually observed, under the assumption that the null hypothesis is true. A smaller p-value indicates stronger evidence against the null hypothesis, but it does not measure the size or importance of an effect [104, 39]. Typical thresholds for interpreting p-values include:

- $p < 0.05$: statistically significant
- $p < 0.01$: highly significant
- $p < 0.001$: very highly significant

It is important to interpret p-values in context and not as a definitive indicator of truth or scientific importance [105].

W Statistic (Wilcoxon Signed-Rank Test)

The Wilcoxon signed-rank test is a non-parametric test used for comparing two related samples or repeated measurements on a single sample. The W statistic summarizes the magnitude and direction of differences between paired observations and is useful when data are not normally distributed [107].

U Statistic (Mann-Whitney U Test)

The Mann-Whitney U test is a non-parametric test used to compare differences between two independent groups. It evaluates whether one group tends to have larger values than the other and is used when the assumptions of the t-test are not met [66].

Linear Mixed Models (LMM)

Linear Mixed Models are used to analyze data with both fixed effects (e.g., experimental conditions) and random effects (e.g., participant variability). They are particularly useful for handling repeated-measures or hierarchical data structures.

- **Fixed effects**: Experimental variables of interest.
- **Random effects**: Account for non-independence (e.g., subjects, trials).
- **Marginal** R^2: Proportion of variance explained by fixed effects alone [72].
- **Conditional** R^2: Proportion of variance explained by both fixed and random effects.

Spearman's Rank Correlation (ρ)

Spearman's ρ is a non-parametric measure of monotonic association between two variables. It assesses how well the relationship between two variables can be described using a monotonic function, and is used when the assumptions of Pearson correlation are violated [94].

Effect Sizes

Effect sizes quantify the magnitude of observed effects and complement p-values by offering insight into practical significance. Commonly used metrics include:

- **Cohen's** d: Measures the standardized difference between two means.
- **Cramér's** V: Used with chi-square tests for associations between categorical variables.
- R^2: Represents the proportion of explained variance in regression or mixed models.

Interpretation guidelines (adapted from [25]):

Effect Size Metric	**Small**	**Medium**	**Large**
Cohen's d	0.2	0.5	0.8
Cramér's V	0.1	0.3	0.5
R^2	0.02	0.13	0.26

Chapter 3

The Impact of Robot Appearance in Collaborative Tasks Across Multiple Culture Groups

3.1 Introduction

Robotic assistants have the potential to enhance convenience, efficiency, and autonomy across diverse settings, including homes, workplaces, and healthcare environments. For example, in geriatric care, robots can assist with mobility, reminders, and household chores [12], helping individuals maintain independence while reducing strain on caregivers. More broadly, they can support individuals with temporary or long-term physical challenges, improve productivity, and minimize manual effort, making them a practical solution for a wide range of users.

While much of the existing literature has emphasized Social Human-Robot Interaction (sHRI), there is a growing need to understand user responses in physically collaborative scenarios. Specifically, how a robot's form—its size, number of limbs, and body configuration—affects trust, comfort, and task efficiency is poorly understood. This gap is further complemented by over-dependence on subjective self-reports in prior work, with limited use of objective physiological measures HR and GSR to assess implicit emotional and cognitive responses. Additionally, the cultural background remains an underexplored dimension in pHRI. However, studies suggest that cultural values influence attitudes toward robots; most of this work is based on hypothetical or social settings rather than real-time physical collaboration.

To address these gaps, the current study investigates how robot morphology and cultural background jointly shape user perceptions during collaborative tasks. Participants from Western, Middle Eastern, and Chinese backgrounds interact with robots of varying physical forms, and we analyze both subjective acceptance and physiological responses. This study addresses three research questions introduced in Chapter 2: **RQI**: How do users from different cultural backgrounds behave and feel during physical interactions with robots? **RQII**: How does robot form influence trust, comfort, and collaboration efficiency? **RQIII**: How do physiological responses (e.g., HR, GSR) vary across cultures in physical HRI?

While our study's population includes healthy younger adults, it provides foundational insights into user preferences and a controlled baseline for. It guides robot design and future research with diverse demographics, such as older adults. By incorporating cultural dimensions, we aim to uncover how cultural background influences user comfort, trust, and acceptance of robots in collaborative tasks.

3.2 Research Questions and Hypothesis

While prior research has explored diverse aspects of HRI, including robot design, autonomy, task complexity, and physiological responses, critical gaps remain regarding how these elements collectively influence user acceptance, stress, and comfort during physically interactive scenarios. This study addresses these gaps through an investigation (**Study 1**) focusing on robot morphology (humanoid vs. non-humanoid), cultural background (Western, Middle Eastern, and Chinese), and task type (independent pick-and-place vs. collaborative object handover). The following research questions, derived from the primary research questions of the thesis of chapter 2, guide this study:

RQ1: Robot Form Factors and Physical Interaction Acceptance. Which robot form factors (e.g., anthropomorphic traits, limb number, overall morphology) most strongly influence user acceptance and comfort during close-proximity physical interactions? Specifically, how do users perceive and accept humanoid (TALOS) compared to non-humanoid (MOVO) robots during collaborative physical tasks?

RQ2: Cultural Influences on Robot Form Acceptance. How do cultural backgrounds influence user perceptions and acceptance of robot form during physical interactions? Do cultural norms and expectations affect user preferences for humanoid versus non-humanoid robots?

RQ3: Task Type and Physiological Response Differences. How does the type of

collaborative task (independent vs. collaborative) influence users' physiological responses (HR, GSR) and perceived stress levels across different robot form factors? Additionally, are these physiological and emotional responses consistent across cultural groups?

To address these research questions, the following hypotheses are formulated based on prior literature outlined in Chapter 2:

Regarding Robot Form Factors (RQ1):

- **H1:** Participants will rate the humanoid robot (TALOS) higher in acceptance and perceive it as less stressful to interact with compared to the non-humanoid robot (MOVO) across both collaborative and independent tasks.

Regarding Cultural Influences (RQ2):

- **H4:** Cultural background will influence user acceptance, with participants from cultures more accustomed to robotic technology (Western participants and Chinese Participants) demonstrating greater acceptance and lower physiological stress when interacting with both robot types compared to participants from less robotically-experienced cultures Middle Eastern.
- **H5**: Western Culture particpants will have more positive opinions on the Humanoid robot compared to the Chinese and Middle Eastern participants.

Regarding Task Type (RQ3):

- **H2:** Collaborative tasks (e.g., object handovers involving direct human-robot interaction) will elicit higher physiological stress responses (HR and GSR) compared to independent tasks (e.g., pick-and-place tasks), irrespective of robot form.
- **H3:** The interaction between task type and robot appearance will influence user perceptions significantly, with TALOS' humanoid design eliciting more positive perceptions specifically in collaborative tasks compared to independent tasks.

By addressing these research questions and testing these hypotheses, the study aims to yield robust insights into the complex interplay among robot design, cultural factors, and interaction types, informing the development of culturally sensitive robotic systems optimized for physical collaboration.

3.3 Methodology

3.3.1 Participants

A total of 52 participants (25 males and 27 females), self-reported, took part in this study. A summary is provided in Table 3.1. Inclusion criteria required participants to identify with one or more of three cultural backgrounds representing collectivist vs. individualist cultures and varying levels of familiarity with robots. All participants were free of any known physical impairments that could affect their task performance. The participants are mostly university students and employees, recruited via posters distributed throughout the campus of the University of Waterloo. Each participant provided informed consent before the study began. None of the participants had previous experience interacting with the robots used in this experiment. To minimize potential biases, participants were not informed of the specific goals of the study until after all experimental sessions had ended. The study was approved by the University of Waterloo Ethics Human Research Ethics Board.

Table 3.1: Participants Summary

Culture Category	**Female**	**Male**
Chinese	8	8
Mean Age	21.41	25.33
SD Age	3.91	4.93
Middle Eastern	7	9
Mean Age	24.00	25.91
SD Age	5.00	4.61
Western	10	7
Mean Age	20.78	21.09
SD Age	2.11	2.02

The table presents the number of participants by culture category with Mean and SD Age of each Gender.

3.3.2 Robots

Two robots were selected for their distinctive features that could be useful in their application as robotic assistants:

- **TALOS** (PAL Robotics), shown in Fig. 3.3, is a bipedal humanoid robot consisting of a torso, head, two arms, and two legs, with 32 degrees of freedom (DOF), a weight of 95 kg and a height of 175 cm. The arms have an anthropomorphic joint layout similar to human arms (7 DOF with shoulder, elbow, and wrist joints) and have a total length of 0.89 m. As the end-effector of each arm, it has a 3-fingered pinch gripper.
- **MOVO** (Kinova Robotics), shown in Fig. 3.17, is a wheeled-base robot consisting of a telescoping torso, head with a screen, and two arms, with 20 DOF, a weight of 120 kg, and an adjustable height of 110 cm to 158 cm. The arms have 7 DOF each and have a non-anthropomorphic joint layout that is more similar to that of a serial robotic manipulator. The MOVO arms are longer than TALOS' with a length of 1.26 m. Similar to TALOS, the end-effectors of MOVO are also 3-fingered pinch grippers.

Both robots were controlled using the Robot Operating System (ROS) [83], and their motion trajectories were generated and implemented using MoveIt [22] to create the Pick & Place and Handover motions. The motions were recorded and replayed to ensure consistency between all participants.

3.3.3 Technical Implementation

Technical Implementation of Kinova MOVO Manipulator

In this experiment, the Kinova Dual Arm Manipulator MOVO was employed to perform two primary tasks: Pick-and-Place and Handover interactions. ROS (Robot Operating System) alongside MoveIt was utilized for the robot's motion planning and control.

Continuous Joints in JACO 7-DOF Arm

The JACO 7-DOF arm, shown in Fig. 3.1, features continuous rotation capabilities in specific joints, notably joints 1, 4, and 6. These continuous joints provide greater flexibility and dexterity but introduce complexity in kinematic modeling and planning due to the infinite rotational space. This necessitates advanced configuration in motion planning algorithms to ensure collision avoidance, accurate orientation, and pose control.

Challenges in Dual-Arm Manipulation

Working with dual-arm manipulators such as MOVO introduces additional complexities:

- **Coordination and Control:** Synchronizing movements between two arms requires complex algorithms to handle inter-arm dynamics effectively, ensuring precise control and timing.
- **Kinematic Complexity:** Dual-arm setups, particularly those with continuous joints, complicate the inverse kinematic solutions due to higher degrees of freedom and potentially infinite rotational configurations.

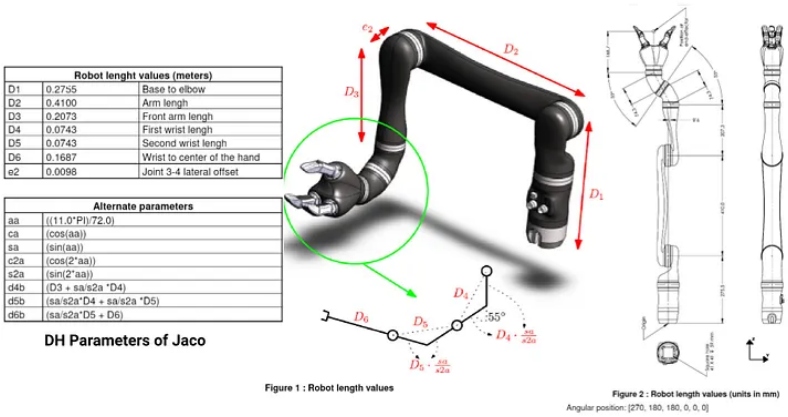


Figure 3.1: Kinematic Model of JACO 7-DOF Arm

Implementation

ROS and MoveIt were used for motion planning, trajectory execution, and interaction management for both robots. The Open Motion Planning Library (OMPL) was integrated via MoveIt to handle motion planning. Several OMPL planners including PRMStar, RRT, and RRTConnect were tested. Parameter tuning focused on handling continuous joints, planning time, and sampling bias to improve performance.

Initialization and Scene Setup

- **ROS and MoveIt Initialization:** ROS nodes managed communication, while MoveIt handled planning and collision checking. Each robot's relevant planning groups (e.g., arm, gripper) were configured with limits on planning time and number of attempts.
- **Robot and Environment Configuration:** The environment included fixed objects such as boxes and shelves, whose positions and orientations were set using quaternions. Robot and object configurations were defined modularly for scene flexibility.
- **Coordinate Frame Transformations:** Rotation matrices were defined to transform poses between wrist and base frames, allowing grasp and placement orientations to be computed in the robot's planning frame.

Pick-and-Place and Handover Task Implementation

- **Pick-and-Place Pipeline:** Both robots followed a pipeline with intermediate poses including approach, grasp, and retreat. The object was attached to the gripper within the planning scene once grasped. Cartesian paths were used when possible for smoother transitions.
- **Gripper Control:** Gripper opening and closing were commanded during relevant stages in the pipeline.
- **Handover Task with Force Detection:** Both robots released the object when a pulling force was detected. MOVO estimated external forces using joint torque sensors; TALOS used wrist-mounted force-torque sensors.

Optimization of OMPL Planning Parameters

OMPL parameters were tuned for both robots to improve path generation under configuration constraints. Optimization objectives included reducing mechanical work and avoiding unnecessary joint movement. Parameters such as goal bias and exploration range were adjusted per planner to improve planning success rates.

Technical Implementation of TALOS Humanoid

The same pick-and-place and handover tasks were implemented on the TALOS humanoid. TALOS differs from MOVO by having an additional rotational joint in the torso, which allows the robot to rotate its upper body separately from its arms Figure 3.2 shows the full kinematic model of TALOS.

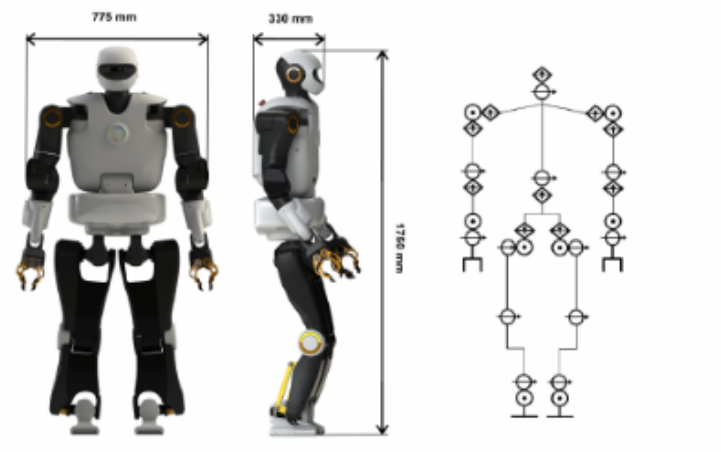


Figure 3.2: Kinematic Model of PAL TALOS Humanoid

Torso Degree of Freedom Advantage

The torso rotation joint in TALOS enabled the robot to align its body with target objects without relying solely on arm configurations. This reduced the complexity of motion planning, as the torso rotation absorbed orientation demands that would otherwise be distributed across multiple arm joints.

Scene and Robot Setup

Objects were placed in known locations using predefined positions and orientations. These were described in terms of Euler angles converted to quaternions. The robot's interaction space included tables and shelves positioned within reach of its manipulators.

Coordinate Frame Transformations

Transformations between wrist and base frames were applied to compute the correct end-effector poses during pick and place. These transformations were based on predefined rotation matrices and ensured consistency across both robots.

3.3.4 Experimental Setup

For both tasks, depicted in Fig. 3.3, the user and the robot each stand between two tables, one on their right and one on their left, with the participant facing the robot. The execution of the tasks can be seen in the video attachment. Two cameras were placed behind the tables to capture and record the participants' body behaviors in both tasks.

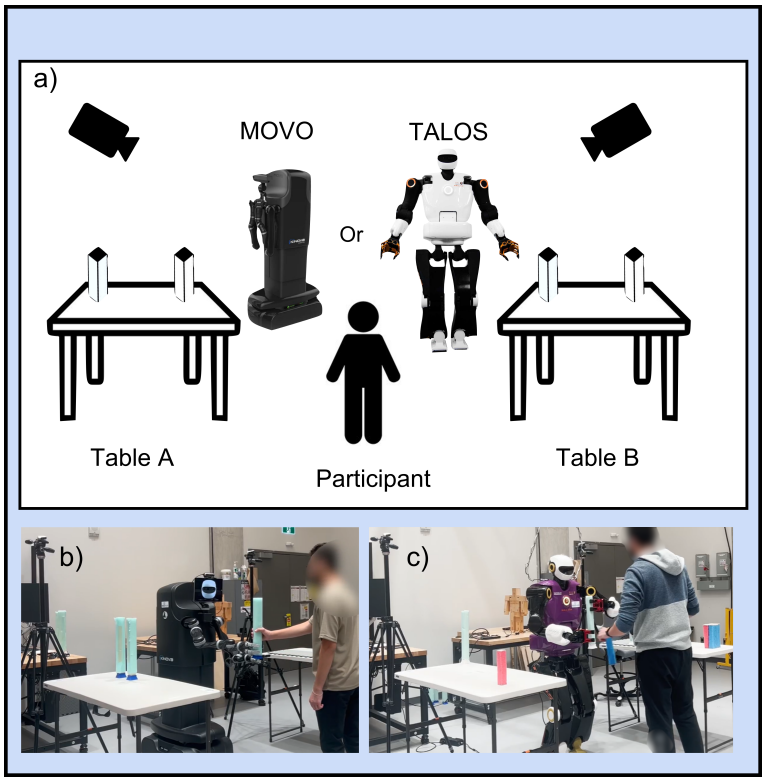


Figure 3.3: a) Overview of the experiment setup with the table placement and how the participants are standing (b) MOVO during Task 2 Handover (c) TALOS during Task 1 Pick & Place. Cameras are positioned to capture participants' body movements.

A within-participant design was used, with participants performing both tasks with each robot. The order of robots that participants encountered was randomized to avoid order bias; however, the order of the tasks was not random; participants always started with Task 1, Pick & Place, and then Task 2, Handover.

3.3.5 Tasks

Task 1 - Independent Pick & Place

In Task 1, the user and robot work independently on a Pick & Place task, where they each have to pick objects one by one from one of the tables adjacent to them and place them on the other adjoining table. The user and robot both have their own set of similar rectangular-shaped objects placed on the table, as can be seen in Fig. 3.3. Considering the robot's slower movements with respect to the participants, the user and robot are given different numbers of objects (8 for the user, 3 for the robot) to ensure that both complete the task in a similar time frame, this was only done for the independent task. The objective is to observe the participant's reaction while simply being in the robot's presence without any direct interaction.

Task 2: Collaborative handover

The robot has to pick the objects (3 in total as in the Pick & Place task) one at a time from the start table and give them to the user, as shown in Fig. 3.3. The user grabs the object from the robot, which detects the pulling force and releases it. Then, the user places the object on the end table. Thus, the user and the robot work collaboratively, with a direct physical interaction.

3.3.6 Measures

Subjective Measures:

Participants completed the **General Attitudes Towards Robots Scale (GAToRS)** [58] pre- and post-task, along with task-specific comfort and usability questions. GAToRS is a multidimensional instrument designed to assess attitudes toward robots across four distinct subscales. These subscales measure personal and societal perceptions separately for

positive and negative attitudes, which, despite their naming (+ and -), are not direct opposites. The subscales include: ($P+$), measuring personal comfort and enjoyment; (P-), measuring personal anxiety and discomfort; ($S+$), measuring societal optimism regarding robot integration; and (S-), measuring societal concerns about potential risks of robots. The following are the questions that were asked before the start of the study.

Part 1: Culture Familiarity and Percentage Western Life Participants were asked how familiar they are with their cultural backgrounds, as well as how much time they spent living in Canada. This was later converted into a percentage based on their age.

Part 2: Familiarity with New Technology Participants were asked about their familiarity with new technologies such as Artificial Intelligence (AI) and robots, as well as how familiar they are with Humanoids and Robot manipulators.

Part 3: Expectations of the robot and the comfort level Participants rated their expectations regarding the robot's efficiency, ease of use, safety, and comfort level.

The open-ended questions, along with Pre-experiment questionnaires, can be found in the AppendixA

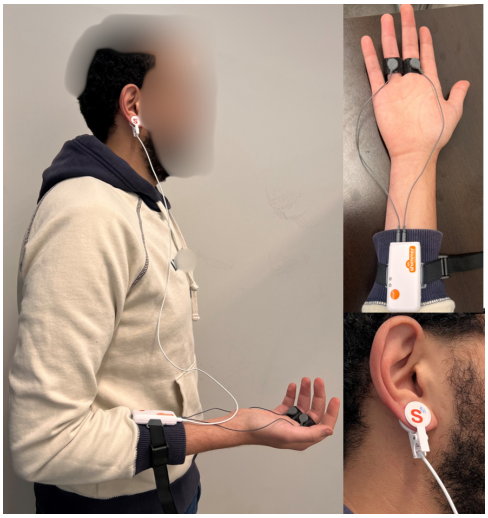


Figure 3.4: Sensor placement on the particpants

The following questions represent the post-study questionnaires:

Part 1: GAToRS Participants answer the GAToRS for each robot.

Part 2: Task Related Questionnaire Questions about Task ease, comfort, and understanding were asked after each robot for each task.

Part 3: Open Ended Questions Participants answer open-ended questions about robot preference, expectations, cultural impact, and their acceptance of the robots

The open-ended questions, along with Pre-experiment questionnaires, can be found in the Appendix A

Objective measures

Objective measures in this study are defined as the ones that can be quantified with physiological sensors, i.e., subconscious responses that are not controllable by participants and do not require subjective interpretation. Physiological signals are objective measures that can be used to measure trust, cognitive load, anxiety, and safety perception [51, 96]. In a previous study on physical human-robot interaction, GSR and HR were used to quantify cognitive load, anxiety, and emotional arousal [53, 41]. Additionally, body posture and body movements in reaction to the robot, e.g., stepping back, extending one's arm, leaning forward or backward, are also considered objective measures in the context of this paper.

Based on work in the state of the art, we selected the following objective measures for this study, which have been effectively employed in existing HRI studies to evaluate user engagement and emotional states during interactions with robots [53, 91]:

- Photoplethysmography (PPG): measures changes in blood volume and converted to heart rate for assessing anxiety level [111].
- Galvanic Skin Response (GSR): measures skin conductance changes associated with autonomic nervous system activity and has been widely used as an indicator of emotional arousal in HRI studies [100, 41]. While GSR alone does not directly quantify cognitive load or emotional intensity, it provides information on physiological arousal when combined with other measures [100].
- Body movements: retrieved from video recordings as the position of the user, arm extension, and alignment of the torso, which are analyzed to assess emotional state and ease of interaction with the robot in both tasks, such as the level of trust and confidence in the robot [4].

PPG and GSR data are collected using the Shimmer3 GSR+ sensor[1]. GSR is measured via two electrodes placed on the fingers of the participant's non-dominant hand to avoid interference when the participant is handling the objects. An optical pulse sensor in the earlobe is used to measure PPG, which is then converted to heart rate after analysis. The main Shimmer3 GSR+ unit is placed on the shoulder of the participant as seen in Fig. 3.4. Body movements is measured using two cameras and inferred using computer vision algorithms (details in Sec. 3.3.8).

3.3.7 Experimental Protocol

For each participant, the experimental session is described below. The experiment was a between-participant study. Participants were assigned the robots randomly, but we made sure that it was equally distributed (eg, Half of the particpants started with TALOS and the other half with MOVO). Note that Task 1 and Task 2 are always performed in order, whereas the order of the robots (MOVO first, then TALOS; or TALOS first, then MOVO) is randomized. Each participant repeats Tasks 1 and 2 with each robot.

1. Prior to the in-person experiment, an information letter containing all pertinent information is sent to the participants for them to read beforehand on a tablet.
2. Each participant signs the consent form using the same tablet.
3. All questionnaires were answered via the tablet, we also provided a keyboard to make it easy for the particpants.
4. The height of the participant is measured and they are asked to fill out an initial form to collect personal information, familiarity with robots, use of technology, and pet ownership, and how much they have lived in Canada.
5. The participant completes the pre-experiment questions.
6. The Shimmer3 GSR+ sensor is placed on the participant to measure GSR and HR (via PPG). The baseline data (which will be used for comparisons) is collected for 5 min. while the participant is seated outside the experimental room, prior to the interaction with the robots.
7. The participant performs Tasks 1 and 2 with either robot (MOVO or TALOS) and then answers the post-experiment and GAToRS questionnaires.

[1]https://www.shimmersensing.com/product/shimmer3-gsr-unit/

8. The participant is asked to step outside for 10 - 15 mins while the robot is switched and the experiment is reset.
9. The participant reenters the room and then performs Tasks 1 and 2 again with the other Robot and then answers the post-experiment and GAToRS questionnaires.
10. The Shimmer3 GSR+ sensor is removed.
11. The participant answers the open-ended questions.

The duration of a full experiment with one participant is approximately one and a half hours.

3.3.8 Data Analysis

This section outlines the methods used to analyze the data collected and describes the specific data types employed to derive the study's results.

Physiological data

Physiological data, including HR and GSR were analyzed to assess participants' responses during interactions with the robot. To process the HR data, we extracted it from the PPG signal using MATLAB. The PPG signal was first loaded into MATLAB with the `load` function for .mat files or other suitable functions depending on the file format, following standard signal processing practices [110].

Next, the signal underwent preprocessing to improve peak detection accuracy. A low-pass filter was applied to eliminate high-frequency noise, as recommended by [96], ensuring that only relevant physiological components remained. Peak detection was then performed using MATLAB's `findpeaks` function, with parameters adjusted to identify systolic peaks corresponding to heartbeats [91]. These peaks were converted into HR in beats per minute (bpm) by calculating the interbeat intervals, a widely accepted method for HR estimation [53] [110].

The heart rate in beats per minute (BPM) was then determined using the formula:

$$\text{Heart Rate (BPM)} = \frac{60}{\text{Average Interbeat Interval (seconds)}}$$

To validate the accuracy of the extracted heart rate, the PPG signal was visualized along with the detected peaks using the `plot` function. As shown in Figure 3.5, the detected peaks in the PPG signal were used to compute the interbeat intervals, which were then used to estimate the heart rate. The results confirmed the effectiveness of the implemented approach in accurately detecting heartbeats and estimating heart rate.

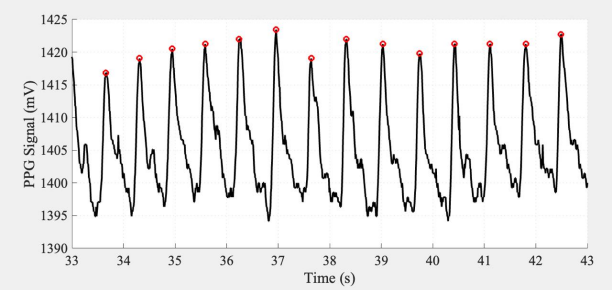


Figure 3.5: PPG signal with detected peaks. The peaks correspond to detected heartbeats, which were used to compute the interbeat intervals and subsequently estimate the heart rate.

For GSR, we measured the mean conductance for each session, following the approach in [41], and calculated the relative difference from a baseline session established before the interaction. This relative change in conductance indicates physiological arousal, consistent with established protocols for assessing autonomic responses in human-robot interaction [**?**, 41].

Stress Inference: Throughout the study, we refer to "stress" as an elevated physiological response inferred from increased HR and GSR above baseline levels. Elevated HR typically indicates heightened arousal, reflecting stress, excitement, or anxiety [96]. Similarly, increased GSR conductance suggests greater skin sweating linked to sympathetic nervous system activation, often associated with stress and potentially excitement or emotional intensity [100]. To interpret these as stress specifically, we relied on contextual cues from the experimental design, where close interaction with the robot, like handover, was likely to induce stress rather than positive arousal, like excitement. This interpretation aligns with prior HRI studies linking events to stress responses [**?**, 3]. However, we acknowledge that without additional self-report measures (e.g., questionnaires), distinguishing stress from excitement or anxiety remains an inference based on task context and physiological trends.

These physiological metrics were chosen for their established reliability in capturing real-time user responses in HRI settings [91] [53], providing an objective complement to

subjective data collected via questionnaires.

Body behaviour

To analyze the body movements of the user from the recorded videos, we use MediaPipe Pose[2] [11]. As can be seen in Fig. 3.6, we use twelve 3D key points, namely both wrists, elbows, shoulders, hips, knees, and ankles. Since our tasks do not require the user to walk, only the upper body is considered for analysis.

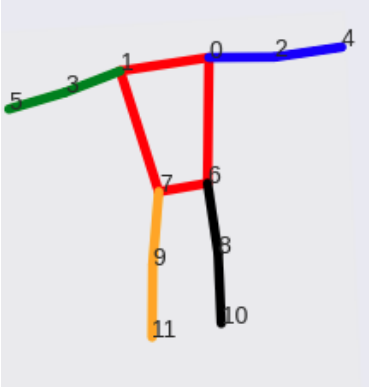


Figure 3.6: MediaPipe Keypoints

The three following distinct body movements were analyzed:

- Arm extension: the distance between the wrist and the shoulder keypoints, e.g. the distance between keypoints 1 and 5 for the right arm. According to [54], when the participant has to grasp the object from the robot, the more they feel uneasy directly interacting with the robot, the more their arm would be extended.
- Torso alignment: this corresponds to the vector's orientation joining both shoulder keypoints, i.e. keypoints 0 and 1. This allows us to see what the user's torso is turned to. Based on [26], users who have their torso turned towards it could fear the robot.
- User position: for this the position of the middle of the line joining both hip keypoints is used, i.e. keypoints 6 and 7. It serves to know to which distance from the robot the user places themselves when interacting with it. Someone afraid of the robot will tend to stand as far away from the robot as possible [71].

[2]https://google.github.io/mediapipe/solutions/pose.html

3.4 Results

The results of this study aim to provide a detailed analysis of the collaborative tasks performed with both humanoid (TALOS) and non-humanoid (MOVO), focusing on subjective perceptions, physiological responses, body movementss, and culture.

Initially, a Two-way ANOVA was used to examine the effects of robot type (MOVO vs. TALOS) and pre-experiment expectations on participant perceptions and task performance. This analysis focused on the influence of robot appearance and design, independently of cultural background.

Linear Mixed Models (LMM) [56] are statistical models that allow researchers to analyze data involving both fixed and random effects. Fixed effects represent systematic influences, such as experimental conditions or specific variables of interest, while random effects account for variations arising from random factors, such as individual differences among participants.

LMM was utilized to identify significant predictors influencing the GAToRS scales ($P+$, P-, $S+$, S-), as shown in Table 3.4. We validated the models using standard diagnostic tests, including examining residuals through histograms and Q-Q plots for assessing normality and scatter plots of residuals versus fitted values to check for homoscedasticity, ensuring that the assumptions underlying linear regression were met.

Cultural effects were further explored using Wilcoxon Signed-Rank and Mann-Whitney U tests, which compared robot interactions and participant responses across different cultural groups. The objective measures were not normally distributed based on the Shapiro-Wilk test; the Wilcoxon Ranked Test was also used to compare physiological measures (HR and GSR) and body movements metrics across different experimental conditions including the effect of Robot appearance and then moving on to the culture effect

These analyses allow for a structured understanding of how robot type, culture, and individual characteristics influenced user perceptions and responses during human-robot interaction. Only results with significance (i.e. $p < 0.05$) are reported and discussed in the remainder of this paper. The qualitative analysis was done using Nvivo software [28],

3.4.1 Interaction Analysis on Pre-Experiment and Post-Experiment Questionnaire

Two-way ANOVA examined the effects of pre-experiment expectations and robot type (TALOS and MOVO) on participant perceptions and task performance. To identify the

more significant robot, we compared mean values for statistically significant variables, with the higher mean indicating a stronger influence on user experience independent of culture. The results, summarized in Table 3.2, highlight significant interactions, particularly how familiarity with robots, expected ease of use, and perceived safety influenced both the GAToRS scale and task metrics based on robot appearance.

Table 3.2: Grouped ANOVA Results for Significant Effects

Dep. Variable	Effect	Robot	p
P-	Expect_Robots_Safe	MOVO	0.047812
P-	Comfort_Level_Robots	MOVO	0.000904
P-	Expect_Robots_Easy_Use	MOVO	0.021563
P+	Expect_Robots_Assist	TALOS	0.003214
P+	Expect_Robots_Easy_Use	TALOS	0.008976
P+	Expect_Robots_Safe	TALOS	0.014231
P+	Comfort_Level_Robots	TALOS	0.025634
S-	Expect_Robots_Assist	MOVO	0.035704
S-	Expect_Robots_Safe	MOVO	0.000001
S+	Familiarity_Robotic_Arms	MOVO	0.015313
S+	Expect_Robots_Assist	MOVO	0.020156
S+	Expect_Robots_Easy_Use	MOVO	0.045678
S+	Comfort_Level_Robots	MOVO	0.032145
S+	Familiarity_Technology	MOVO	0.038761
Task_Comfort_1	Expect_Robots_Easy_Use	TALOS	0.028764
Task_Comfort_1	Comfort_Level_Robots	TALOS	0.000005
Task_Comfort_1	Expect_Robots_Safe	TALOS	0.000012
Task_Comfort_2	Comfort_Level_Robots	TALOS	0.029384
Task_Comfort_2	Expect_Robots_Easy_Use	MOVO	0.005673
Task_Ease_1	Comfort_Level_Robots	TALOS	0.012847
Task_Ease_2	Familiarity_Humanoids	TALOS	0.039874
Task_Ease_2	Comfort_Level_Robots	TALOS	0.041762
Task_Understanding_1	Comfort_Level_Robots	TALOS	0.019832
Task_Understanding_2	Familiarity_Robotic_Arms	TALOS	0.043215

Table 3.3: **Effect** refers to the interaction term tested in ANOVA that significantly influenced the dependent variable. **p** values indicates statistical significance when $p < .05$.

P+: Expectations of robot assistance, ease of use, safety, and comfort were associated

with more positive perceptions, emphasizing the importance of pre-interaction attitudes. Furthermore, a higher level of comfort with robots was associated with improved perceptions, particularly in interactions with TALOS,

P-: Interactions between robot type and expectations of safety, ease of use, and comfort with robots significantly influenced negative perceptions (P-). Participants with higher expectations of robot ease of use and safety, or those who felt more comfortable with robots, reported greater negative perceptions when interacting with MOVO compared to TALOS.

S+: Societal optimism was significantly influenced by expectations of robot assistance, ease of use, and safety, as well as familiarity with technology. Participants with higher expectations and familiarity with robots and technology expressed stronger societal optimism about robots, particularly evident in interactions with MOVO.

S-: Expectations of robot assistance and safety were associated with increased societal concerns (S-) when interacting with MOVO. Participants who had higher expectations for robots assisting in tasks or ensuring safety reported stronger societal concerns when engaging with MOVO

Task Metrics: Participants who had higher expectations of robot ease of use and safety, as well as higher comfort levels with robots, reported greater task comfort, ease, and understanding when interacting with TALOS, Familiarity with technology also contributed to improved perceptions of task comfort and ease. An exception was observed for *Expect_Robots_Easy_Use*, where higher safety expectations were linked to increased comfort when interacting with MOVO. However, across most metrics, TALOS was perceived as the more effective and comfortable robot for task execution.

Analysis of Variance (ANOVA) Results

The two-way ANOVA analysis revealed several statistically significant interaction effects between participants' pre-experiment perceptions and robot type (TALOS vs. MOVO) on various user outcomes, including comfort, perceived task ease, and multiple dimensions of the General Attitudes Toward Robots Scale (GAToRS).

To assess the practical significance of these effects, partial eta squared (η^2) values were computed. Based on Cohen's guidelines for interpreting effect sizes:

- Approximately 38% of significant effects fell within the medium effect range ($\eta^2 \approx$ 0.06–0.14),

- Around 27% of effects exceeded the threshold for a large effect size ($\eta^2 > 0.14$),
- The remaining 35% were classified as small effects ($\eta^2 \approx 0.01$–0.06).

The average effect size across all significant results was moderate ($\eta^2 \approx 0.10$), indicating that robot appearance and users' pre-interaction expectations had a noticeable, though not overwhelming, influence on participants' responses.

These findings suggest that robot morphology and individual expectations can influence perceived comfort, emotional responses, and trust during physical interaction tasks. Nonetheless, given the moderate effect sizes and limitations in sample size, these results should be considered preliminary. Future work with larger samples and more diverse task settings will be valuable in verifying and extending these observations.

3.4.2 LMM Results for GAToRS Scales

LMM results for each subscale are presented in Table 3.4. The significant predictors are summarized below.

Comfort and Enjoyment Around Robots (*P+*)

Height negatively impacted personal positive perceptions, indicating taller participants had lower positive perceptions. Cultural background significantly influenced perceptions, with Chinese (*Estimate* = 0.748, $p = 0.0094$) and Western participants (*Estimate* = 1.349, $p = 0.0030$) expressing higher positive perceptions compared to Middle Eastern participants. Interestingly, higher percentages of life spent in Western culture negatively influenced positive perceptions . Additionally, interactions with TALOS (humanoid robot) increased positive perceptions compared to MOVO . Familiarity with robotic arms positively influenced *P+* , whereas greater familiarity with humanoids decreased positive perceptions. Expecting robots to be safe, higher comfort levels around robots, and tasks perceived as easier all significantly contributed to higher positive perceptions.

Anxiety and Discomfort Around Robots (*P-*)

Comfort levels with robots and comfort in tasks significantly reduced negative perceptions, suggesting that increased comfort reduces anxiety. Male participants reported significantly lower negative perceptions compared to females. Height was positively associated with negative perceptions, indicating taller participants felt higher discomfort around robots.

Table 3.4: LMM Results for GAToRS Scales

Variable	Estimate	S.E.	T	p-value
P+ (Positive Perceptions)				
Height	-0.037	0.011	-3.294	0.0018
Culture (Chinese vs. Middle Eastern)	0.748	0.277	2.699	0.0094
Culture (Western vs. Middle Eastern)	1.349	0.433	3.119	0.0030
Percentage Western Life	-0.013	0.004	-3.125	0.0029
Robot Collaborated (TALOS vs. MOVO)	0.096	0.026	3.636	0.0004
Familiarity Humanoids	-0.490	0.180	-2.716	0.0090
Familiarity Robotic Arms	0.328	0.147	2.235	0.0298
Expect Robots Safe	0.522	0.178	2.924	0.0051
Comfort Level Robots	0.300	0.145	2.070	0.0435
Task Ease	0.066	0.024	2.738	0.0068
P- (Negative Perceptions)				
Comfort Level Robots	-0.597	0.139	-4.304	0.0001
Task Comfort	-0.273	0.045	-6.071	¡0.0001
Gender (Male vs. Female)	-0.690	0.318	-2.167	0.0348
Height	0.034	0.016	2.169	0.0347
S+ (Social Positive Perceptions)				
Pet Owned (No vs. Yes)	-0.403	0.173	-2.327	0.0239
Expect Robots Assist	0.640	0.128	4.993	¡0.0001
Comfort Level Robots	0.215	0.101	2.126	0.0383
Task Understanding	0.059	0.021	2.825	0.0053
S- (Social Negative Perceptions)				
Culture (Chinese vs. Middle Eastern)	-0.300	0.272	-1.101	0.0776
Culture (Western vs. Middle Eastern)	-1.814	0.420	-4.317	0.0001
Percentage Western Life	0.014	0.004	3.150	0.0027
Expect Robots Assist	0.439	0.179	2.451	0.0177
Expect Robots Safe	-0.393	0.169	-2.325	0.0240
Task Understanding	0.046	0.022	2.103	0.0370

Note: The **Estimate** column provides regression coefficients; **S.E.** indicates standard error; **T** values indicate direction and strength of effects; **p-values** indicate statistical significance.

Societal Hopes about Robots (*S+*)

Participants without pets showed lower social positive perceptions compared to pet owners. Higher expectations regarding robots' ability to assist significantly increased social positive perceptions. Additionally, greater comfort levels with robots and better task understanding significantly improved participants' positive societal perceptions.

Societal Concerns About Robots (*S-*)

Participants from Western cultural backgrounds reported significantly lower societal negative perceptions compared to Middle Eastern participants (*Estimate* = -1.814, $p = 0.0001$). However, a greater percentage of life spent in Western cultures was positively correlated with societal negative perceptions (*Estimate* = 0.014, $p = 0.0027$). Expecting robots to assist was associated with increased societal negative perceptions, while expecting robots to be safe significantly decreased societal concerns. Lastly, better task understanding was linked to increased societal negative perceptions.

3.4.3 Effect Size Analysis for Linear Mixed Models

To better understand the explanatory power of each linear mixed model (LMM) used in this chapter, we computed the marginal and conditional R^2 values following the approach proposed by Nakagawa and Schielzeth [72].

Marginal R^2 (R^2_m) represents the variance explained solely by the fixed effects in the model. **Conditional** R^2 (R^2_c) reflects the variance explained by both fixed and random effects (in this case, random intercepts grouped by *Participant_ID*).

Table 3.5: LMM Effect Size Summary for Perception Models

Model Outcome Variable	**Marginal** R^2	**Conditional** R^2
P– (Negative Perceptions)	0.411	0.915
P+ (Positive Perceptions)	0.494	0.965
S+ (Social Positive Perceptions)	0.404	0.928
S– (Social Negative Perceptions)	0.368	0.956

Note: Marginal R^2 quantifies variance explained by fixed effects alone; Conditional R^2 includes both fixed and random effects (participants).

The following summarize the results presnted in the table:

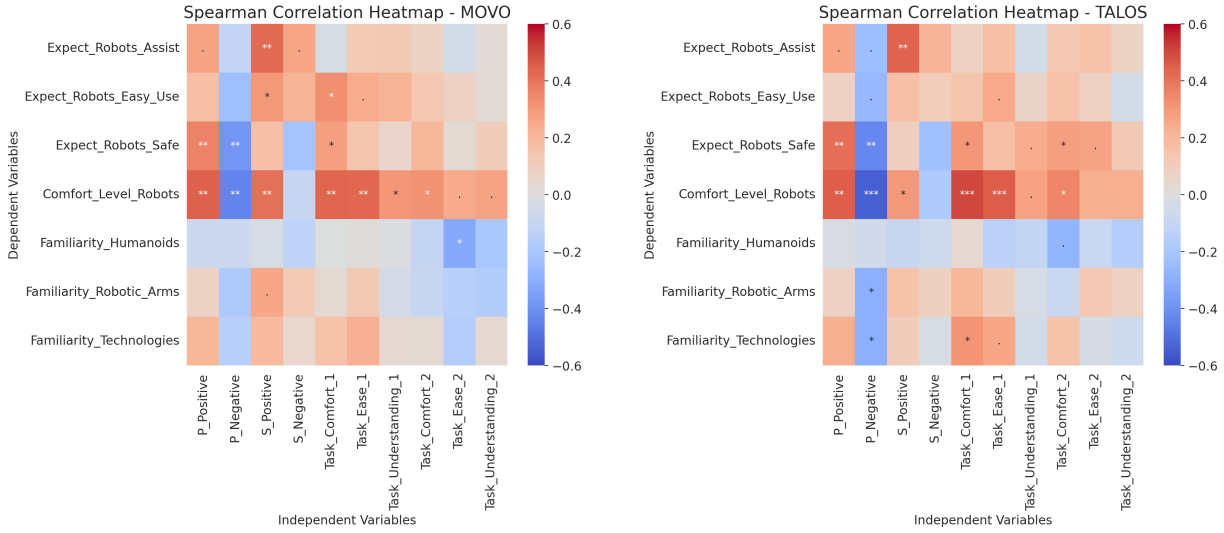
- All models show high conditional R^2 values (between 0.91 and 0.97), indicating that both fixed effects and participant-level variation jointly explain the vast majority of the variance in the outcome variables.
- The marginal R^2 values (ranging from 0.37 to 0.49) suggest that fixed effects alone explain a moderate proportion of variance. According to Cohen's criteria [25], these represent medium-to-large effect sizes.
- The relatively large gap between R^2_m and R^2_c underscores the importance of accounting for individual differences in human-robot interaction studies.

Overall, the fixed effects included in each LMM—such as comfort, expectations, familiarity, robot type, and cultural group—make meaningful contributions to explaining participants' perceptions and attitudes. However, participant-level random variation still accounts for a substantial portion of the overall variance, as expected in within-subject experimental designs.

3.4.4 Spearman Correlation Analysis

This section presents the results of the Spearman correlation analysis, examining relationships between independent(Post-experiment questions) and dependent(Pre-experiment questions) variables across two robot interaction scenarios: MOVO and TALOS. Correlation strengths and significance levels are visualized using heatmaps to illustrate variable interactions clearly and also across culture groups.

Figure 3.7a presents the Spearman correlation heatmap for the MOVO robot. Strong positive correlations ($p < 0.01$) emerged between *Comfort Level with Robots* and participants' perceptions of robot safety (*Expect Robots Safe*) as well as expectations regarding robot assistance (*Expect Robots Assist*). Additionally, *Expect Robots Safe* showed significant positive correlations with personal comfort and enjoyment (P+) and personal anxiety and discomfort (P-). Other notable correlations ($p < 0.05$) include *Expect Robots Easy Use* with societal optimism regarding robot integration (S+), and *Comfort Level with Robots* with comfort during independent tasks (*Task Comfort 1*) and understanding of the robot's actions and intentions during independent tasks (*Task Understanding 1*). Some negative correlations were observed but were generally of lower magnitude.



(a) MOVO's Sperman's Correlation results

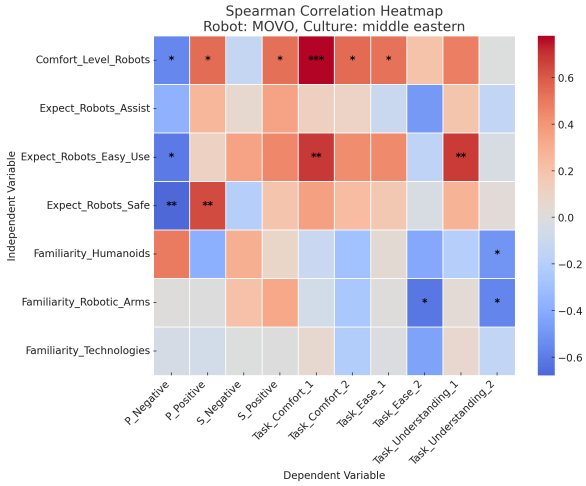
(b) TALOS's Sperman's Correlation results

Figure 3.7: Sperman Correlation analysis for each robot

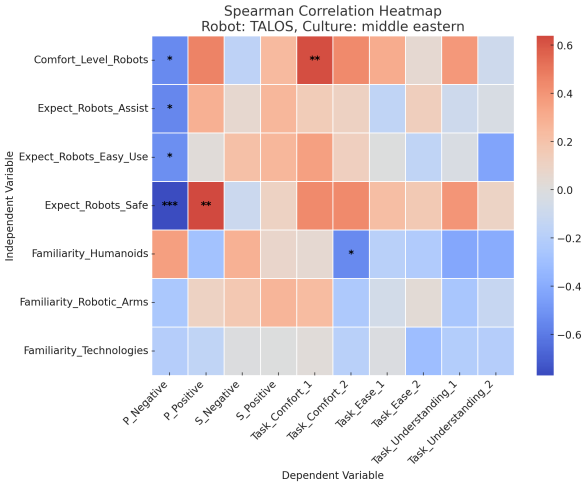
Figure 3.7b presents the Spearman correlation heatmap for the TALOS robot scenario. Similar to MOVO, strong positive correlations ($p < 0.01$ and $p < 0.001$) were evident between *Comfort Level with Robots* and perceptions of robot safety (*Expect Robots Safe*), personal comfort and enjoyment (P+), and comfort during independent tasks (*Task Comfort 1*). Significant negative correlations ($p < 0.05$) were found involving personal anxiety and discomfort (P-) with familiarity regarding robotic arms and technologies. Correlations involving expectations regarding robot assistance (*Expect Robots Assist*) exhibited some variations compared to the MOVO scenario.

Furthermore, the Spearman correlation analysis revealed distinct trends in how different cultural groups interacted with MOVO and TALOS. Among Middle Eastern participants, the MOVO robot exhibited a higher average correlation with a strong positive correlation peaking at (0.778, $p < .01$), particularly in cases where familiarity with humanoid robots was linked to task ease and comfort. However, negative correlations as low as (-0.678, $p < .05$) suggest that higher expectations of safety and ease of use did not always lead to a positive post-task perception. For TALOS, the average correlation was lower (0.009), with a maximum positive correlation of (0.638, $p < .05$), indicating that while prior familiarity played a role, it had a more inconsistent impact compared to MOVO. Additionally, TALOS showed the strongest negative correlation across all groups (-0.770, $p < .001$), suggesting that in some cases, participants with higher familiarity experienced lower perceived ease

or comfort.

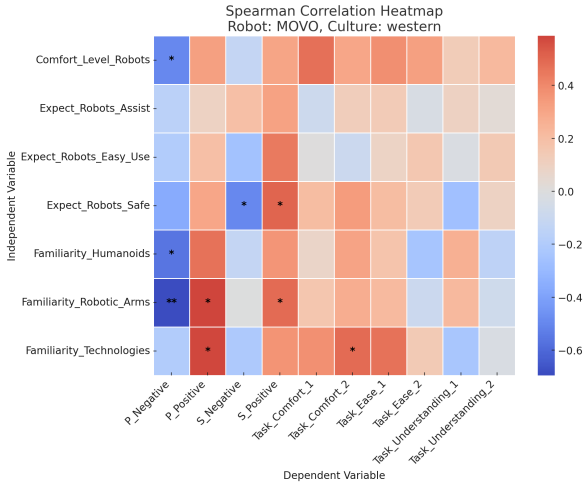


(a) Sperman's Correlation results for MOVO with Middle Eastern Culture

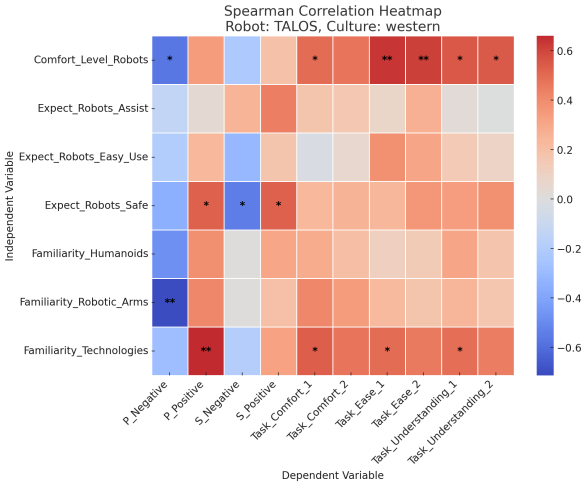


(b) Sperman's Correlation results for TALOS with Middle Eastern Culture

Figure 3.8: Sperman Correlation analysis for Middle Eastern Culture Category

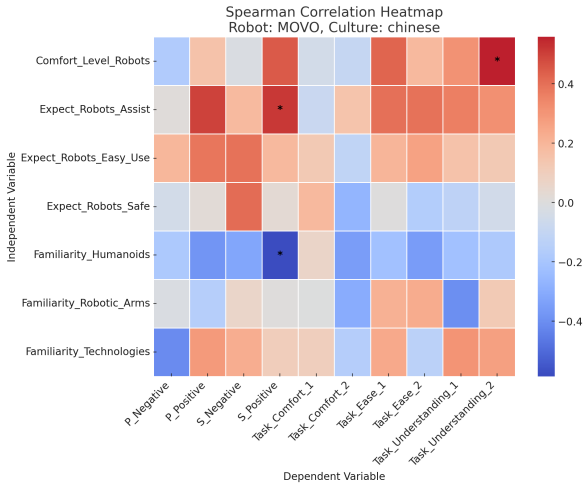


(a) Sperman's Correlation results for MOVO with Western Culture

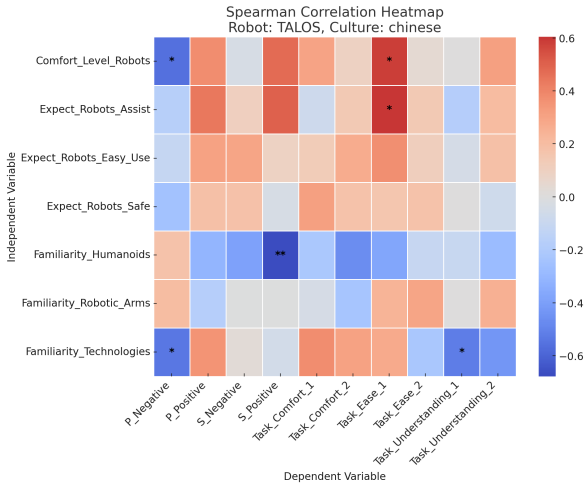


(b) Sperman's Correlation results for TALOS with Western Culture

Figure 3.9: Sperman Correlation analysis for Western Culture Category



(a) Sperman's Correlation results for MOVO with Chinese Eastern Culture



(b) Sperman's Correlation results for TALOS with Chinese Eastern Culture

Figure 3.10: Sperman Correlation analysis for Chinese Culture Category

Western participants exhibited a slightly different pattern, with MOVO showing the highest average correlation (0.101), suggesting a moderate relationship between pre-task familiarity and post-task perceptions. However, the strongest correlation recorded was (0.586, $p < .05$), which was notably lower than the peak values observed in the Middle Eastern group. Negative correlations were present as well, with a minimum value of (-0.694, $p < .05$), further reinforcing that expectations did not always align with actual experience. TALOS was not analyzed separately for this group due to its lower variation in results.

For Chinese participants, MOVO again showed stronger positive correlations compared to TALOS, with an average correlation of (0.061) and a maximum correlation of (0.558, $p < .01$), indicating that familiarity with robotic arms had a meaningful impact on perceived comfort. However, negative correlations were present (-0.586, $p < .05$), suggesting that some participants with prior familiarity still faced unexpected difficulties or disappointments. TALOS followed a similar trend, with an average correlation of (0.044) and a maximum of (0.603, $p < .05$), reinforcing that participants with prior experience generally found interactions smoother but not always perfectly aligned with their expectations.

3.4.5 Wilcoxon Signed-Rank test-Cultural Compariosns

Initially, we conducted a comparative analysis between the two robots, TALOS and MOVO, across all the GAToRS scales (P+, P-, S+, S-). The statistical tests did not reveal any significant differences between the two robots. Thus, we proceeded with further statistical tests to explore potential differences at a cultural level. The data distribution of the scales can be seen in Figure 3.11 and 3.12

Table 3.6: Wilcoxon Rank-Sign Test Results for MOVO vs. TALOS

Variable	Culture	W-Statistic	p-value
S+	Middle Eastern	174	0.0013
P+	Western	240	0.0026
S+	Western	304	0.0246
S+	Chinese	188	0.0211
S-	Chinese	232	0.0124

Note: Significant Wilcoxon rank-sign test results comparing MOVO and TALOS robots within each cultural group. Only statistically significant results ($p < .05$) are reported.

To explore whether perceptions of TALOS and MOVO varied within each cultural group, we performed the Wilcoxon Signed-Rank Test separately for each group. The results, shown in Table 3.6, revealed several significant differences across the four dimensions of robot perception. Among Middle Eastern participants, TALOS was rated significantly higher than MOVO in terms of social positivity (S+). Western participants showed a stronger preference for TALOS in both positive perceptions (P+) and social positivity, with consistently higher ratings than MOVO. In contrast, Chinese participants rated MOVO higher than TALOS in social positivity, while also perceiving MOVO as less concerning (S-) than TALOS. These findings suggest culturally influenced differences in how each robot was evaluated, particularly with regard to social attributes and perceived positivity.

3.4.6 Mann-Whitney U Test — Cultural Comparisons

To investigate whether cultural background influenced participant attitudes towards TALOS and MOVO, we conducted Mann-Whitney U tests comparing cultural groups. The results are summarized in Table 3.7.

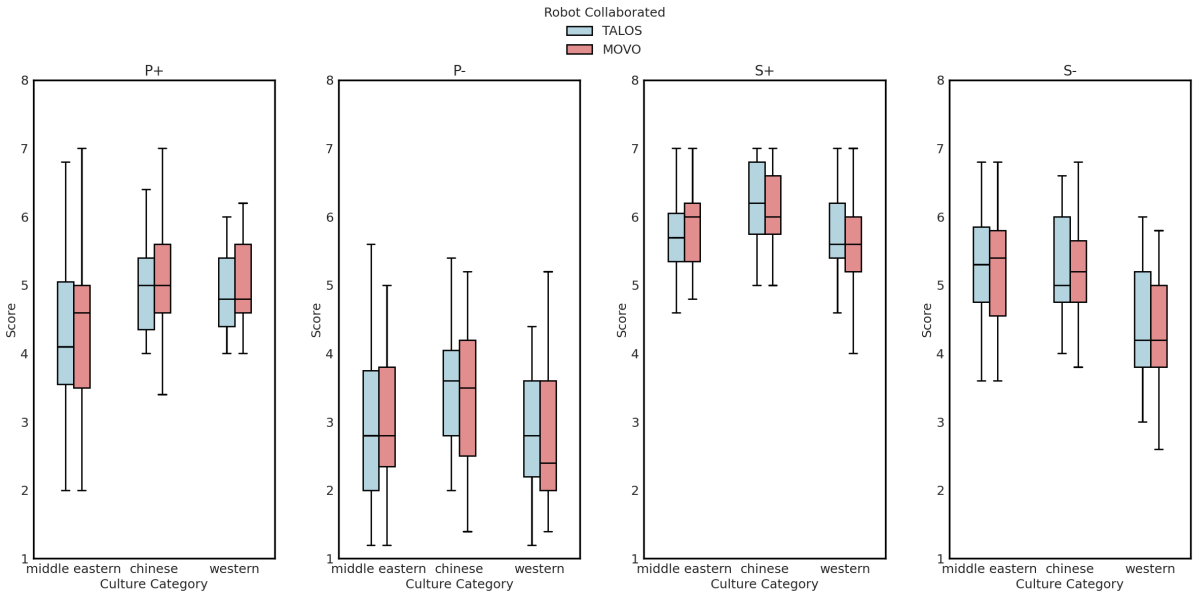


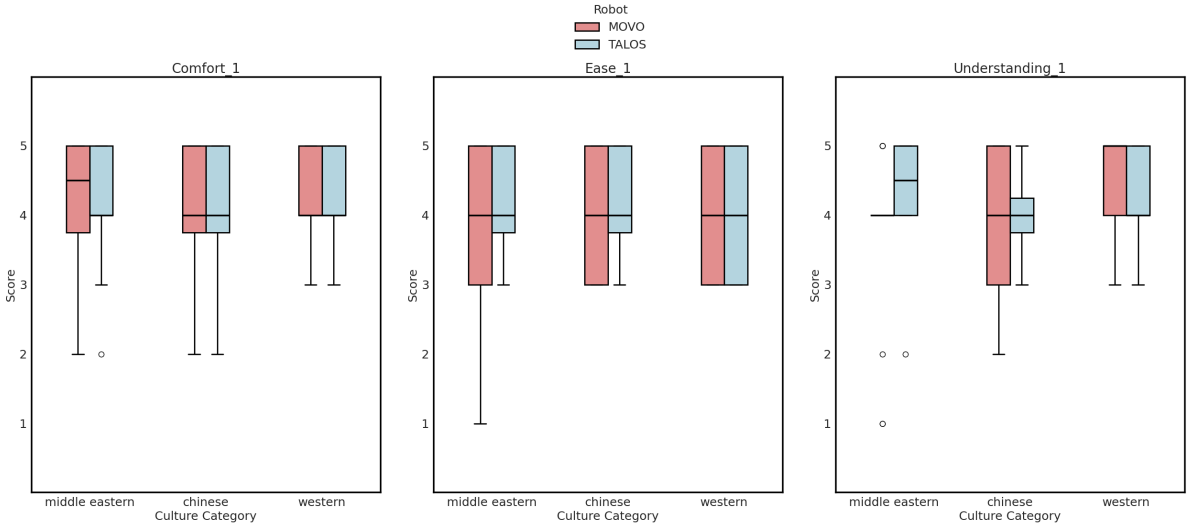
Figure 3.11: Comparison across all the GAToRS scales (P+ (Comfort and Enjoyment), P- (Unease and Anxiety) , S+ (Societal Optimism), S- (Societal Concerns), for each culture and robot

Comfort and Enjoyment Around Robots (P+)

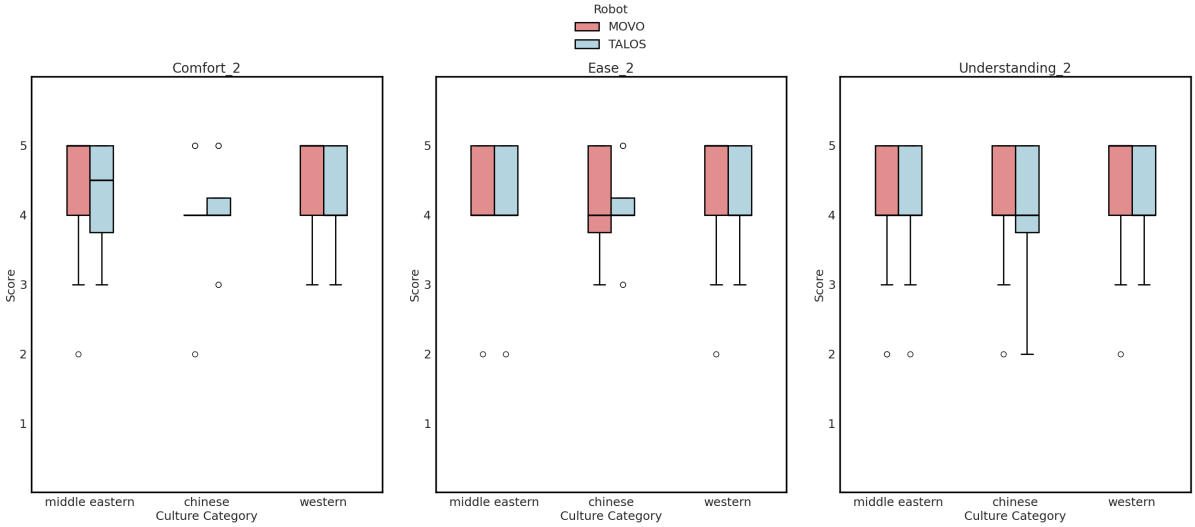
Cultural background significantly influenced participants' positive perceptions of MOVO. Western participants reported higher positive scores compared to Middle Eastern participants, indicating greater comfort and enjoyment in their interactions with MOVO. Similarly, Chinese participants rated MOVO more positively than Middle Eastern participants. These findings suggest that cultural familiarity with robotic systems may contribute to variations in positive perceptions.

Unease and Anxiety Around Robots (P-)

Negative perceptions also varied across cultural groups, particularly in their interactions with MOVO. Western participants exhibited lower negative scores than Chinese participants, suggesting reduced unease. This difference highlights how previous exposure to technology and cultural expectations about automation might contribute to variations in discomfort levels when engaging with robotic systems.



(a) Task-related questionnaire - Pick & Place Task



(b) Task-related questionnaire - Handover

Figure 3.12: Task Metric distribution across robots

Rational Hopes About Robots at a Societal Level (S+)

Cross-cultural differences were also evident in perceptions of societal benefits associated with robots. Chinese participants reported the highest ratings when interacting with

MOVO, surpassing both Middle Eastern and Western participants. These findings indicate that participants from cultures with greater exposure to robotics may hold stronger expectations regarding their positive societal contributions.

Rational Worries About Robots at a Societal Level (S-)

Concerns about the societal impact of robots varied between cultural groups. Western participants exhibited the lowest levels of societal concern, followed by Middle Eastern participants. In contrast, Chinese participants reported the highest concerns about the broader implications of robotic integration. These results underscore the influence of cultural perspectives on rational concerns regarding automation and its societal effects.

Table 3.7: Mann-Whitney Test Results for MOVO and TALOS

Scale	Comparison	Robot	p-value
P+	Middle Eastern vs. Western	MOVO	0.0312
P+	Middle Eastern vs. Chinese	MOVO	0.0210
P+	Middle Eastern vs. Chinese	TALOS	0.0203
P-	Western vs. Chinese	MOVO	0.0462
S+	Middle Eastern vs. Chinese	MOVO	0.0222
S+	Western vs. Chinese	MOVO	0.0141
S+	Western vs. Chinese	TALOS	0.0141
S-	Middle Eastern vs. Western	MOVO	0.0002
S-	Middle Eastern vs. Western	TALOS	0.0004
S-	Western vs. Chinese	MOVO	0.0004
S-	Western vs. Chinese	TALOS	0.0006

Note: Significant Mann-Whitney test comparisons between cultural groups for each robot. Only statistically significant results ($p < .05$) are shown.

3.4.7 Physiological data analysis

The results indicate distinct physiological responses across robot types and tasks. We compared the objective measures (GSR, HR, and Body Movements) across robot types and also tasks, which revealed significant differences.

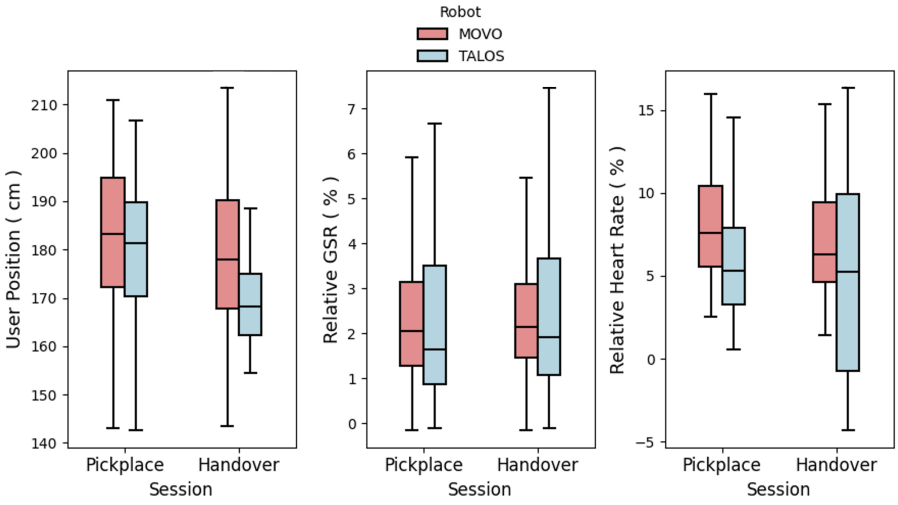


Figure 3.13: Comparison of User Position, HR, and GSR across robot type

Fig. 3.13 visually presents the comparisons across robot types, while Table 3.8 provides a detailed summary, including significant differences observed in both robot type and task comparisons. For Body Movements, only User Position is reported as the remaining factors did not yield significant differences.

Heart Rate HR Results

Table 3.8: Grouped Wilcoxon Ranked Test Results for Heart Rate, GSR, and User Position

Metric	Comparison	p-val
Heart Rate	MOVO vs TALOS Pick & Place	0.00018
	MOVO vs TALOS Handover	0.0346
	MOVO Pick & Place vs Handover	0.0089
	TALOS Pick & Place vs Handover	0.000656
GSR	MOVO vs TALOS Handover	0.00765
	MOVO Pick & Place vs Handover	0.00034
	TALOS Pick & Place vs Handover	0.000438
User Position	MOVO vs TALOS Pick & Place	0.0325
	MOVO vs TALOS Handover	0.00976

Significance levels $p < .05$.

The *Handover* task elicited higher responses when comparing the two robots and the two tasks. Within the *Pick & Place task*, participants interacting with MOVO showed higher HR than those interacting with TALOS. For *Handover* tasks, TALOS and MOVO elicited similar responses. Within each robot type, the Handover task consistently elicited higher heart rates than the Pick & Place task with more significant median differences observed for MOVO and TALOS. These findings indicate that robot type and task type both influence heart rate responses, with the Handover task generally inducing higher rates.

The *Handover* task elicited higher heart rate responses compared to the *Pick & Place task* task across all cultural groups. The most significant differences were found among Chinese participants, where MOVO pick-and-place vs. handover ($p = 0.00005$) indicated a major increase in HR during Handover tasks.

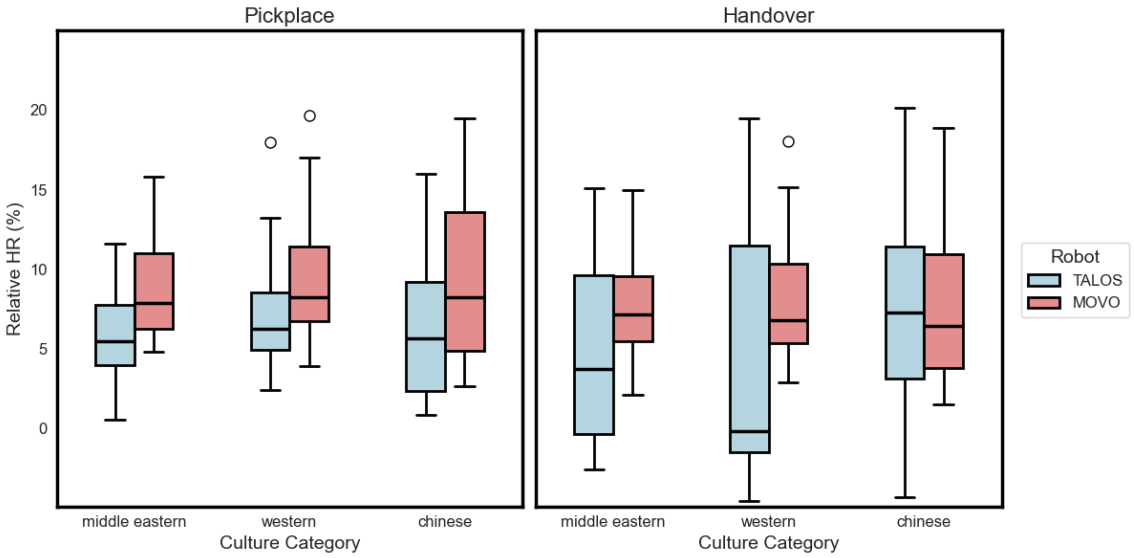


Figure 3.14: Comparison of HR across robot type and culture category .

Middle Eastern participants also exhibited strong HR differences between MOVO pick-and-place and MOVO handover ($p = 0.00006$), highlighting task-dependent variations in physiological arousal. Western participants displayed consistent HR variations across conditions but with smaller effect sizes.

The boxplots in Fig. 3.14 illustrate that MOVO generally induces higher HR responses, particularly in the*Pick & Place task* task for Middle Eastern and Western participants, while Chinese participants show a stronger distinction between tasks. These findings sug-

Table 3.9: Wilcoxon Ranked Test Results for Heart Rate Across Cultures

Culture	Comparison	p-val
Middle Eastern	TALOS_Pick & Place vs MOVO_Pick & Place	0.00012
Middle Eastern	MOVO_Pick & Place vs MOVO_Handover	0.00006
Western	TALOS_Pick & Place vs MOVO_Pick & Place	0.00021
Western	TALOS_Handover vs MOVO_Handover	0.02016
Western	TALOS_Pick & Place vs TALOS_Handover	0.02322
Chinese	MOVO_Pick & Place vs MOVO_Handover	0.00005
Chinese	TALOS_Pick & Place vs MOVO_Pick & Place	0.00043
Chinese	MOVO_Pick & Place vs MOVO_Handover	0.00065

Significance threshold: $p < 0.05$.

gest that both robot type and task type influence heart rate, with MOVO generally leading to higher physiological arousal.

GSR Results

For *Handover tasks*, MOVO elicited significantly higher GSR compared to TALOS . In *Pick & Place tasks*, no significant difference was observed between TALOS and MOVO. Within the MOVO conditions, the *Pick & Place task* elicited significantly lower GSR compared to the Handover task, while for TALOS, the same trend was observed with the Handover task eliciting higher GSR than the Pick & Place task. Comparing across robots, MOVO and TALOS showed a significant difference in GSR during the Handover task, but no significant difference during the Pick & Place task. These results indicate that on average, the Handover task tends to elicit higher arousal compared to the Pick & Place task, and MOVO generally produces stronger GSR responses than TALOS, particularly in the Handover task.

For *Handover tasks*, MOVO elicited significantly higher GSR compared to TALOS across all cultural groups. The strongest GSR variations were observed in Western participants, where MOVO pick-and-place vs. MOVO handover ($p = 0.01286$) and TALOS pick-and-place vs. handover ($p = 0.01743$) showed notable increases in GSR. Middle Eastern participants also exhibited a significant increase in GSR from pick-and-place to handover for TALOS ($p = 0.04791$), though the effect was weaker compared to Western and Chinese groups.

The Chinese participants demonstrated a significant increase in GSR between tasks for

Table 3.10: Wilcoxon Ranked Test Results for Galvanic Skin Response Across Cultures

Culture	Comparison	p-val
Middle Eastern	TALOS_Pick & Place vs TALOS_Handover	0.04791
Western	TALOS_Pick & Place vs TALOS_Handover	0.01743
Western	MOVO_Pick & Place vs MOVO_Handover	0.01286
Chinese	TALOS_Pick & Place vs TALOS_Handover	0.03534
Chinese	MOVO_Pick & Place vs MOVO_Handover	0.01245

Significance threshold: $p < 0.05$.

both TALOS ($p = 0.03534$) and MOVO ($p = 0.01245$), suggesting they were more physiologically reactive to task changes compared to other cultural groups. Fig. 3.15 confirms that MOVO generally leads to greater physiological arousal than TALOS, particularly in the Handover task.

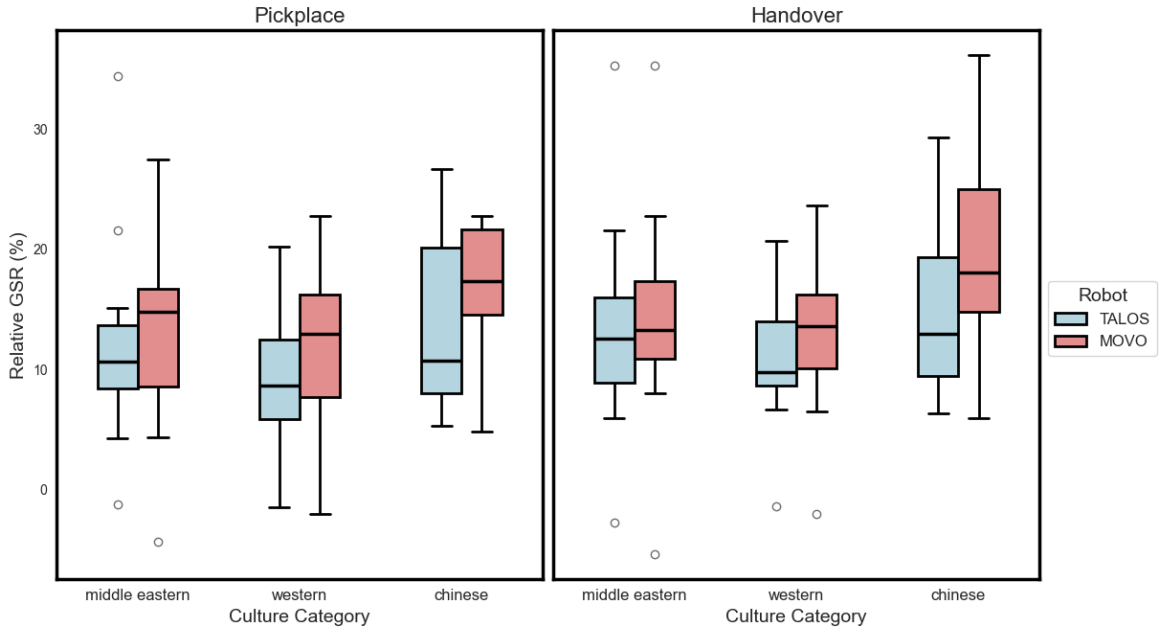


Figure 3.15: Comparison of HR across robot type and culture category .

These results indicate that on average, the Handover task tends to elicit higher arousal compared to the Pick & Place task, and MOVO generally produces stronger GSR responses

than TALOS. Cultural background appears to play a role in physiological reactivity, with Chinese participants showing the most pronounced differences across tasks.

3.4.8 Body Movements and Posture Analysis

Wilcoxon Ranked Results

Analysis of the user's positioning relative to the robot revealed significant differences between MOVO and TALOS Participants maintained a greater distance from MOVO compared to TALOS during both tasks. The results suggest that robot familiarity, perceived safety, and comfort level may influence how closely participants position themselves relative to the robot.

For Cultural differences between the robots for Torso Roll and Tilt Angles Significant differences in torso movement were observed among participants from different cultural backgrounds when interacting with MOVO and TALOS robots. Specifically, in the *Handover* task, *Western* participants exhibited a significant difference in *Torso Roll Angle* ($p = 0.00244$), where interactions with MOVO led to greater variation in torso movement. Similarly, in the *Pick & Place* task, *Chinese* participants demonstrated differences in *Torso Alignment* ($p = 0.00316$), indicating that the two robots influenced their posture differently.

Table 3.11: Wilcoxon Ranked Test Results for Torso and User Position for MOVO vs TALOS

Culture	**Task**	**Variable**	**p-val**
Western	Handover	Torso Roll Angle Mean	0.00244
Middle Eastern	Handover	User Position (Z) Mean	0.00061
Chinese	Pick & Place	Torso Alignment Angle Mean	0.00316
Chinese	Pick & Place	User Position (Z) Mean	0.03052

Significance threshold: $p < 0.05$.

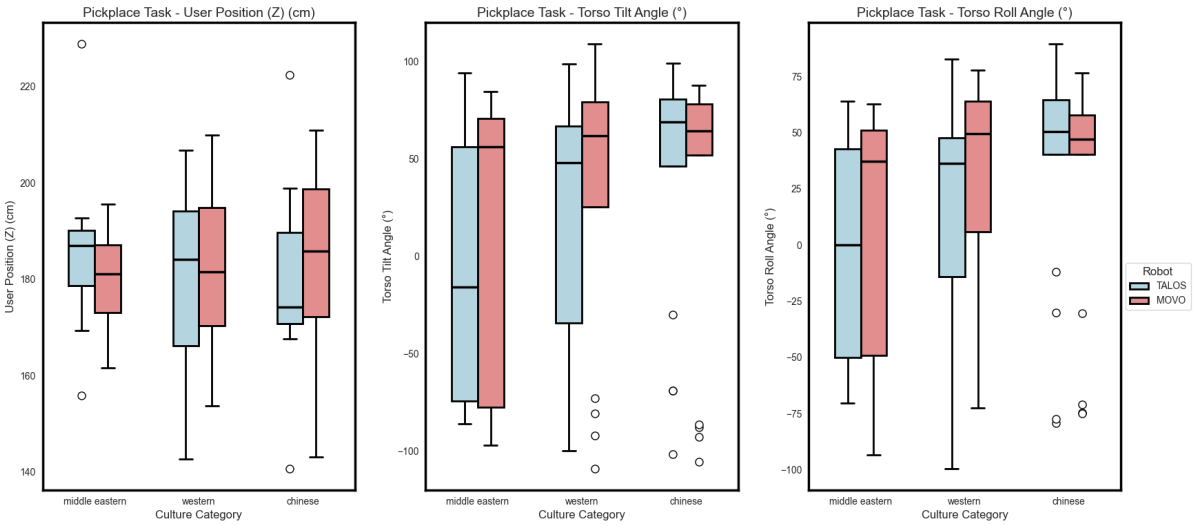


Figure 3.16: Comparison of different body poses across robot type and culture category for Pick & Place

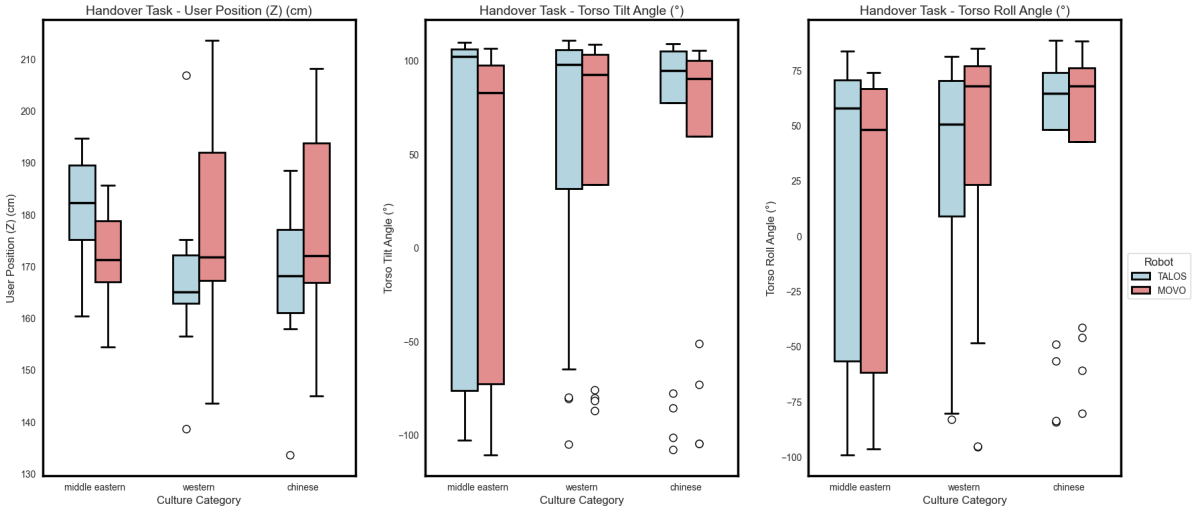


Figure 3.17: Comparison of different body poses across robot type and culture category for Handover.

For Cultural differences between the robots for User Position (Z) Mean The box plots indicate that participants interacting with *TALOS* generally positioned themselves closer to the robot compared to *MOVO*, particularly in the *Handover* task. This effect was statistically significant for *Western* ($p = 0.00061$) and *Chinese* ($p = 0.03052$) participants.

Mann-Whitney Results

Tables 3.12 and 3.13 present the *Mann-Whitney test results* for body posture differences across cultural groups when interacting with *MOVO* and *TALOS*, respectively.

Torso Roll and Tilt Angles The results indicate a significant difference in the *Torso Roll Angle* for *Western* participants during the *Handover* task ($p = 0.00244$), where interactions with MOVO resulted in greater variation in torso movement compared to TALOS. Additionally, in the *Pick & Place* task, *Chinese* participants exhibited significant differences in *Torso Alignment Angle* ($p = 0.00316$), suggesting that MOVO and TALOS influenced torso posture differently. Cross-cultural differences in *Torso Roll Angle* were also observed across both tasks.

Table 3.12: Mann-Whitney Test Body Posture A Cross Culture - MOVO

Task	Culture 1	Culture 2	Variable	p-val
Pick & Place	middle eastern	chinese	Torso Tilt Angle Mean	0.01245
Pick & Place	middle eastern	chinese	Torso Roll Angle Mean	0.02404
Handover	middle eastern	chinese	Torso Roll Angle Mean	0.00175
Handover	middle eastern	western	Torso Roll Angle Mean	0.01585

Significance threshold: $p < 0.05$.

User Position (Z) Mean The analysis of user positioning reveals that participants generally positioned themselves closer to TALOS than MOVO. This effect was observed in the *Handover* task and was statistically significant for both *Western* ($p = 0.00061$) and *Chinese* ($p = 0.03052$) participants. The data suggest that users maintained different interaction distances depending on the robot they engaged with, which varied across cultural groups.

Table 3.13: Mann-Whitney Test Body Posture A Cross Culture - TALOS

Task	Culture 1	Culture 2	Variable	p-val
Pick & Place	middle eastern	chinese	Left Arm Extension Mean	0.00574
Pick & Place	middle eastern	chinese	Torso Tilt Angle Mean	0.00148
Pick & Place	middle eastern	chinese	Torso Roll Angle Mean	0.00089
Pick & Place	chinese	western	Torso Tilt Angle Mean	0.00354
Pick & Place	chinese	western	Torso Roll Angle Mean	0.01685
Handover	middle eastern	chinese	Torso Roll Angle Mean	0.03478
Handover	chinese	western	Right Arm Extension Mean	0.03208
Handover	chinese	western	Torso Roll Angle Mean	0.03038

Significance threshold: $p < 0.05$.

3.4.9 Qualitative Feedback by Cultural Groups

Our Thematic analysis process addressed to [24] six-phase approach. Significant patterns that align with the research questions were obtained by analyzing the common elements in how a topic is discussed [24, 113].

The analysis was made by collaborating with another researcher, in which we met and made iterative revisions to refine the codebook used to identify the common themes, through discussions. An inter-rater reliability test showed good agreement (Cohen's $\kappa = 0.85$). We refined the final themes collaboratively until we agreed on the final themes.

The open-ended qualitative responses were analyzed to identify trends within each cultural group regarding robot preference, comfort levels, and perceptions of robot interactions. We identified overarching themes reflecting key cultural insights. Within each theme, sub-themes are indicated in bold text.

Out of 52 participants, 60% ($n = 31/52$) favored TALOS, while 38% ($n = 20/52$) preferred MOVO, and 2% ($n = 1/52$) of Middle Eastern participants stated they preferred neither. TALOS was preferred by 11/16 (69%) of Chinese participants, 7/16 (44%) of Middle Eastern participants, and 11/17 (65%) of Western participants. MOVO was preferred by 5/16 (31%) of Chinese participants, 8/16 (50%) of Middle Eastern participants, and 6/17 (35%) of Western participants.

Factors Affecting Robot Preference and Discomfort

This theme captures participants' preferences and dislikes regarding the robots' design, appearance, movement, efficiency, comfort, and safety concerns.

Size and Space Efficiency: MOVO was favored for its compact design, particularly in household settings where space is limited. Among those who preferred MOVO, 5/16 (31%) of Chinese, 6/16 (38%) of Middle Eastern, and 4/17 (24%) of Western participants mentioned size as a primary factor. One Middle Eastern participant noted, *"MOVO fits better in a home because it is small and doesn't take up unnecessary space."* However, TALOS was perceived as more stable and structured, making its movements feel more predictable, especially in environments where precision is crucial.

Human-like Appearance and Predictability: TALOS' humanoid appearance was a major factor for those who preferred it. Many participants (11/16, 69% of Chinese; 7/17, 41% of Western; and 5/16, 31% of Middle Eastern) noted that TALOS' human-like form and predictable movement patterns made it feel more intuitive to interact with. One Chinese participant stated, *"TALOS moved more naturally, which made it feel more like an assistant than just a machine."* Western participants echoed this sentiment, with one mentioning, *"It was easier to anticipate TALOS' movements, so I trusted it more in close interactions."* Additionally, 7/16 (44%) of Chinese and 6/17 (35%) of Western participants noted that TALOS' movements were more precise and space-efficient compared to MOVO.

Robot Shape and Appearance: MOVO's non-humanoid appearance was a key factor for those who preferred it, with 5/16 (31%) of Chinese participants, 7/16 (44%) of Middle Eastern participants, and 6/17 (35%) of Western participants explicitly stating their discomfort with humanoid robots. One participant remarked, *"I prefer robots that don't resemble humans. It makes them feel more like tools rather than artificial people."* TALOS, on the other hand, was seen as more approachable by those who valued predictability in interactions. However, some participants found its humanoid features unsettling, with 4/16 (25%) of Middle Eastern participants and 3/17 (18%) of Western participants stating, *"The head and arms on the TALOS did not make me feel extremely comfortable, but I did feel that they were unnecessary."*

Comfort and Intimidation: TALOS was seen as intimidating by 7/16 (44%) of Middle Eastern participants and 6/16 (38%) of Chinese participants, who cited its size and human-like movements as reasons for discomfort. MOVO, while perceived as less intimidating, was sometimes described as *"weird"* or *"alien-like"* due to its unique arm design by 5/16 (31%) of Middle Eastern and 4/17 (24%) of Western participants.

Movement & Control: MOVO was described as having smoother, quieter move-

ments, particularly by 7/16 (44%) of Chinese participants, 6/16 (38%) of Middle Eastern participants, and 5/17 (29%) of Western participants. However, 6/16 (38%) of Chinese participants and 5/16 (31%) of Middle Eastern participants found MOVO's arm movements unnatural. TALOS' movements were perceived as more human-like, which 8/16 (50%) of Chinese and 6/17 (35%) of Western participants preferred because it made interactions feel more intuitive.

Task Efficiency and Functionality: TALOS was viewed as more capable of handling complex tasks, particularly in industrial or structured environments. 11/16 (69%) of Chinese, 7/16 (44%) of Middle Eastern, and 11/17 (65%) of Western participants favored TALOS due to its perceived versatility. However, MOVO was praised for precision, especially in delicate object handling, with 5/16 (31%) of Chinese and 4/17 (24%) of Western participants stating, *"MOVO seemed more careful when handling small items, which made me trust it more."*

Noise, Safety, and Risk: 7/17 (41%) of Western participants mentioned TALOS was too loud, while 5/16 (31%) of Middle Eastern participants found MOVO's humming noise annoying. Additionally, 8/16 (50%) of Middle Eastern participants and 6/17 (35%) of Western participants raised safety concerns about MOVO's idle positioning. One participant stated, *"TALOS made a loud mechanical noise when moving, which was distracting."* Another mentioned, *"MOVO was quieter, but its humming sound got irritating over time."*

Robots' Helpfulness for Older Adults

Participants shared their perspectives on whether robots could assist older adults in mobility and household tasks.

Helping with Mobility and Lifting: 77% ($n = 40/52$) of participants believed robots could assist older adults. Western participants had the highest support (81%, $n = 13/17$), followed by Middle Eastern (69%, $n = 11/15$) and Chinese (56%, $n = 9/16$). One participant noted, *"A robot that can help an older adult move things or support them while walking would be very useful."*

Household Tasks and Assistance: 73% ($n = 38/52$) of participants across all cultures believed robots could assist with daily chores like cooking and cleaning. *"Robots could help older adults by handing over objects or cleaning around the house,"* one participant stated. Another added, *"I would trust a robot to help an elderly person with simple, repetitive tasks."*

Reducing Dependence and Improving Accessibility: 62% ($n = 10/15$) of Middle Eastern participants believed robots could help older adults maintain independence by assisting

with navigation, while 47% ($n = 8/17$) of Western participants raised concerns about widespread acceptance. One participant explained, *"Older adults who struggle to move around on their own could benefit from a robot helper, but I don't know if they would trust it."*

Limitations and Concerns: 50% ($n = 8/16$) of Middle Eastern participants and 41% ($n = 7/17$) of Western participants believed that robots were not advanced enough yet to provide full assistance to older adults. *"Right now, robots are too limited. They are not ready to fully replace a caregiver."*

Ideal Robotic Assistant's Appearance and Functionality

Participants described their ideal robot's appearance, size, movement, and functions. The following key themes emerged:

Preference for Non-Humanoid or Machine-Like designs: Some participants preferred robots that are clearly machine-like, avoiding human-like features. This preference was stated by 6/16 (38%) of Chinese, 7/16 (44%) of Middle Eastern, and 4/17 (24%) of Western participants. One participant noted, *"Robot should not resemble a human."* Another emphasized, *"No facial features, no head, no eyes, just a functional machine."*

Preference for Humanoid or Partially Humanoid Appearances: In contrast, some participants preferred robots with human-like features to facilitate better interaction. A balance between a humanoid appearance and robotic predictability was emphasized. 10/16 (63%) of Chinese, 9/16 (56%) of Middle Eastern, and 16/17 (94%) of Western participants favored some degree of human-like design. Participants also mentioned avoiding the uncanny valley by using clearly artificial materials. A Western participant explained, *"A more human-like design makes it easier to interact with the robot naturally."*

Size and Space Considerations: Many participants preferred compact robots to ensure they do not take up excessive space, particularly in home environments. This was noted by 5/16 (31%) of Chinese, 4/16 (25%) of Middle Eastern, and 5/17 (29%) of Western participants. A Chinese participant stated, *"Sleek, lightweight designs feel less intimidating."* Another participant mentioned, *"A robot that fits easily in small spaces would be more practical."*

Aesthetic Preferences: Participants expressed differing opinions on aesthetics. Some preferred soft edges and rounded shapes to make robots appear more harmless and friendly, mentioned by 3/16 (19%) of Chinese and 3/17 (18%) of Western participants. Color preferences also varied, with 6/17 (35%) of Western participants favoring bright colors (e.g.,

purple, green) and 3/16 (19%) of Chinese participants preferring neutral colors (black, grey, white). One participant explained, *"Bright colors can make the robot look more friendly and less intimidating."*

Movement and Structure: Participants emphasized the importance of movement and structural features in ensuring safe and effective interactions. Many preferred robots with flexible arms that could reach varying heights, which was particularly noted by 5/16 (31%) of Chinese and 4/17 (24%) of Western participants. Additionally, 6/16 (38%) of Chinese and 5/17 (29%) of Western participants favored wheels over legs for easier mobility, as it provided better balance and efficiency. Hidden joints and soft materials were also preferred, especially by 3/16 (19%) of Middle Eastern participants, to reduce potential harm during interaction. The ability to move with smooth, seamless gestures was another important feature, as 7/16 (44%) of Chinese participants indicated that fluid motion made robots feel more natural and less mechanical.

Functional Highlights: Functionality played a critical role in shaping participants' preferences for their ideal robotic assistant. Many emphasized the importance of voice-controlled features for accessibility, with 5/16 (31%) of Chinese and 4/17 (24%) of Western participants mentioning its significance. Additionally, 7/16 (44%) of Middle Eastern participants and 8/17 (47%) of Western participants expressed a strong desire for robots that could automate daily chores such as cooking and cleaning. One participant stated, *"The robot should help with repetitive household tasks to save time."* This highlights the growing expectation for robots to integrate seamlessly into daily life, making routine tasks easier and more efficient.

Cultural Background Influence on Robot Interaction and Preference

The extent to which culture influenced participants' robot preferences varied across cultural groups.

Culture as a Significant Influence: 56% ($n = 9/16$) of Chinese participants, 50% ($n = 8/16$) of Middle Eastern participants, and 35% ($n = 6/17$) of Western participants stated their culture influenced their perception of robots. Some Chinese participants noted that their exposure to robots in China influenced their comfort level, while a Middle Eastern participant explained, *"My upbringing emphasized manual labor and effort. Having a robot do things for me feels unnatural."* Among Western participants, a few cited media influence, with one stating, *"Movies always portray AI as dangerous, which makes me more hesitant."*

No Cultural Influence Reported: 38% ($n = 6/16$) of Chinese participants, 44% ($n = 7/16$) of Middle Eastern participants, and 35% ($n = 6/17$) of Western participants

stated their preferences were shaped by personal exposure to technology rather than culture. One Chinese participant noted, *"I don't think my cultural background influenced my robot preference. It's more about how efficient the robot is."* Similarly, a Western participant stated, *"I work with automation, so I view robots through a professional lens rather than a cultural one."* Among Middle Eastern participants, one mentioned, *"In my country, we don't see robots often, so I base my preference on how useful they seem."*

Alternative Reasons for Preferences: Among those who did not attribute their choices to culture, 38% ($n = 6/16$) of Chinese participants mentioned prior exposure to robots in education or media, 44% ($n = 7/16$) cited a preference for efficiency, and 31% ($n = 5/16$) referenced safety concerns as the primary factor. Additionally, 53% ($n = 8/16$) of Middle Eastern participants and 47% ($n = 8/17$) of Western participants mentioned their preference was driven by direct interaction with robots in professional or educational settings rather than cultural background.

Media and Pop-Culture Influence: 35% ($n = 6/17$) of Western participants specifically mentioned that movies, social media, and pop culture played a role in shaping their perception of robots. Some noted that portrayals in sci-fi movies made them more skeptical, while others felt more familiar and open-minded due to exposure in fiction. One participant stated, *"I've grown up watching movies about robots (usually negative), so seeing the human figure robot brought forward those fears."* Another explained, *"Media always portrays robots and AI as scary things ... makes me more hesitant."*

Religious and Ethical Considerations: 6% ($n = 1/16$) of Middle Eastern participants mentioned religious concerns about human-like robots, while another 6% ($n = 1/16$) referenced ethical concerns. One participant noted, *"I think religious beliefs play a role in how we accept robots. Some aspects of AI feel unnatural."*

These findings highlight how cultural background influences robot preferences, from design choices to perceptions of safety and usability.

3.5 Discussion

Our study provides insights into the interplay between robot appearance and task complexity and cultural background, as well as their collective impact on user stress and comfort during human-robot interactions. This discussion addresses the hypotheses and situates the findings within the broader context of robot appearance for assistive and collaborative applications.

Our findings partially confirm **H1**. Interactions with the humanoid robot TALOS consistently elicited lower negative perceptions (P-) during independent tasks like Pick & Place with respect to the non-humanoid robot MOVO. In addition, participants stood closer to TALOS compared to MOVO during both tasks, suggesting that the human-like features of TALOS might have reduced user discomfort and stress by aligning with users' expectations of motion and interaction. However, during collaborative tasks such as Handover, the difference between TALOS and MOVO was less pronounced. This indicates that task complexity and user familiarity mediate the effect of robot appearance on user comfort.

The second hypothesis **H2** predicted that collaborative tasks would elicit higher stress levels, as evidenced by physiological responses, regardless of robot type. The results of our study strongly support this hypothesis. HR and GSR were significantly higher during the Handover task than the Pick & Place task. Collaborative tasks inherently require closer human-robot interaction and coordination, leading to heightened cognitive and emotional engagement [63]. These findings emphasize the need to carefully design robot behaviors and interactions in collaborative scenarios to mitigate user stress.

H3 explored the interaction between task type and robot appearance on user perceptions, initially predicting that the humanoid robot TALOS would be perceived more positively in collaborative tasks compared to independent ones. However, the findings indicate a more significant relationship between robot appearance, task complexity, and user expectations. While TALOS was consistently evaluated more positively in independent tasks, in collaborative settings there were mixed perceptions of MOVO and TALOS. Participants reported higher trust levels and ease of interaction with TALOS, but (P+) and engagement varied depending on task complexity and prior expectations. In contrast, MOVO elicited higher negative perceptions (P-) and (S+), particularly in collaborative tasks. These results highlight the importance of tailoring robot appearance and interaction dynamics to specific task demands, strengthening findings in the existing literature [38, 40].

Furthermore, partial insights from the open-ended responses support these findings. Among the participants, 29 out of 52 clearly favored TALOS due to its humanoid features, which they associated with comfort and familiarity, expressing that it looks like a human. In contrast, 19 participants preferred MOVO, citing its non-humanoid appearance as less intimidating and more utilitarian. Notably, 4 participants reported no preference. These varied preferences indicate the nuanced role of robot appearance in shaping user perceptions, highlighting the importance of tailoring robot designs to the specific needs and expectations of target users. Additionally, 43 participants believed that both robots hold significant potential for assisting older adults in home environments, reinforcing the

relevance of these findings for assistive robotics.

From a broader perspective, the study highlights several insights into the appearance of robots and task complexity and their implications for user experience. The results suggest that TALOS's appearance may have improved comfort and trust during less demanding tasks, suggesting their suitability for assistive roles, such as those encountered by older adults or users with physical impairments. Conversely, results suggest that MOVO's functional design appears to align well with tasks emphasizing efficiency, indicating its potential preference in settings prioritizing task efficiency (e.g. industrial setting).

The elevated stress levels observed during collaborative tasks underline the need for intuitive robot behaviors and predictable motions. Simplifying these tasks and incorporating clear verbal or visual cues could enhance user acceptance and reduce anxiety. Finally, this study enhances our understanding of how robot appearance and task complexity influence user responses by integrating subjective measures with objective physiological and behavioral data. It highlights the need to balance anthropomorphic features for social acceptance with task-specific functionality, particularly in assistive applications

H4 suggested that cultural background would influence user acceptance, such that participants from cultures with greater exposure to robotic technology (Western and Chinese) were expected to show higher acceptance and lower physiological stress compared to those from cultures with less exposure (Middle Eastern). This hypothesis was robustly supported by multiple results. Western and Chinese participants reported significantly higher positive perceptions (P+) and lower negative perceptions (P-) than Middle Eastern participants. Physiological measures, such as HR andGSR, also indicated lower stress responses for Western and Chinese participants. Specifically, interactions with the non-humanoid MOVO robot elicited higher stress responses among Middle Eastern participants. Additionally, Spearman correlation analyses reinforced these results, showing stronger positive correlations between comfort levels with robots and expectations of safety for Western and Chinese participants compared to Middle Eastern participants. These findings collectively underscore the moderating role of cultural exposure to robotic technology on acceptance and physiological comfort. Interestingly, more nuanced effects of cultural exposure emerged. For example, greater exposure to Western life (quantified by the percentage of life spent in Western countries) was associated with decreased personal positive perceptions (P+) and increased societal concerns (S-). These correlations suggest that extended exposure to Western environments may heighten expectations or increase awareness of broader societal implications of robotics, adding complexity to the overall relationship between cultural background and robot acceptance, suggesting context-dependent influences of cultural integration [64]. Middle Eastern participants exhibited the most cautious attitudes toward both robots. While MOVO, with its non-humanoid form, was perceived as

safer, TALOS, despite its human-like appearance, elicited more significant skepticism.

H5 anticipated that Western participants would prefer the humanoid robot over the non-humanoid robot. The results strongly supported this hypothesis. Western participants consistently reported higher comfort and positive perceptions towards TALOS, the humanoid robot, compared to MOVO. Wilcoxon signed-rank tests and Mann-Whitney U tests further corroborated these preferences. Qualitative feedback echoed these results, highlighting TALOS's humanoid appearance and intuitive movements as primary contributors to Western participants' positive experiences [10].

Further supporting these hypotheses, physiological data revealed significant task-dependent variations across cultural groups. Handover tasks elicited greater physiological arousal compared to pick-and-place tasks, especially for Middle Eastern and Chinese participants interacting with MOVO. Body posture analyses complemented these results, revealing that participants generally maintained greater distances and exhibited more pronounced adjustments with MOVO. Specifically, Western participants showed notable variations in torso movements during MOVO interactions, indicating greater caution or discomfort.

Qualitative feedback enriched these findings, showing distinct preferences aligned with cultural backgrounds. Western participants favored TALOS's human-like predictability, Middle Eastern participants often preferred MOVO due to discomfort with humanoid appearances, and Chinese participants frequently emphasized functionality and compactness.

This study revealed distinct response variations among the Western, Middle Eastern, and Chinese participant groups, emphasizing that culture significantly shapes user interactions and perceptions of robot forms in collaborative tasks. Western participants demonstrated generally higher acceptance and comfort levels with humanoid robots, reflecting individualistic values and familiarity with anthropomorphic technology. In contrast, Middle Eastern participants showed increased caution and reduced comfort, potentially stemming from less frequent exposure to robotic technologies and higher expectations regarding robot safety and ease of use. Chinese participants displayed intermediate responses, highlighting a complex interplay of collectivist values and moderate technological familiarity. These cultural distinctions display the importance of integrating culturally adaptive interfaces into robotic appearance to enhance user acceptance and effectiveness across diverse societal contexts.

Limitations

Several limitations were identified in this study that should be considered when interpreting the findings. Firstly, the sample size of 52 participants is relatively modest, which may limit the generalizability and external validity of the results. Increasing the number and diversity of participants in future studies could provide more comprehensive insights into cultural influences. Additionally, robot motor noises might have influenced participant perceptions and responses, as acoustic factors could have independently affected their comfort and stress levels. Furthermore, the robots' color schemes may have subtly impacted emotional and perceptual responses, potentially influencing perceptions of friendliness, trustworthiness, or intimidation. Future research would benefit from more careful control of these factors to isolate the effects directly related to robot appearance. Additionally, participants' prior experience or familiarity with robotic technologies was not thoroughly controlled or assessed, potentially influencing their comfort levels and perceptions. Future studies could evaluate participants' previous exposure to robots to better understand the effects attributed purely to robot appearance. Moreover, the relatively short duration of interactions might have limited the depth of participants' perceptions and physiological responses. Prolonged interactions could reveal more nuanced responses and longer-term comfort or trust dynamics. Lastly, the study environment, being a controlled laboratory setting, might not fully capture naturalistic responses and behaviors in real-world scenarios. Future research should aim for more real-life setups to better generalize findings to everyday contexts.

Chapter 4

The Role of Culture in Perceptions of Control Modes in pHRI

4.1 Introduction

The integration of robots into human environments is accelerating, particularly in healthcare, rehabilitation, and assistive living. Collaborative robots are increasingly deployed in real-world settings to support physical tasks such as eldercare, household assistance, and rehabilitation. As these systems become more prevalent, understanding user acceptance and trust, especially when errors occur, becomes critical to successful deployment.

Culture has been shown to influence how people perceive robots' appearance, behavior, and roles. However, fewer studies have examined how cultural background affects user preferences for robot *control modes*—namely, whether robots operate autonomously or under human supervision—and how users attribute responsibility when errors occur. While our previous study focused on robot form and task context, this study shifts attention to autonomy and error handling across cultures.

We specifically investigate whether cultural background affects how users respond to errors made by either autonomous or human-controlled robots. This includes perceptions of trust, responsibility, and preferred control mode. Prior work suggests that people attribute blame differently depending on whether a robot is seen as an independent agent or a human proxy [47, 52]. Furthermore, trust in robots can persist despite errors, especially in low-risk situations, though user confidence may decline [89, 87]. Cultural factors likely moderate these effects, but this area remains underexplored.

We conducted an experimental study using the *Sawyer* (Rethink Robotics) collaborative robotic arm to explore these cultural differences. Participants from three distinct cultural backgrounds—Chinese, Middle Eastern, and Western—interacted with the robot under two control conditions: **autonomous** and **human-controlled**. During each interaction, participants performed three tasks, with the second task introducing a deliberate error. By analyzing user reactions to the errors in both control modes, we aim to provide insights into cultural variations in HRI, particularly in terms of trust, error perception, and control mode preference. This study directly addresses **RQIV**: *How do cultural differences influence user preferences for autonomous versus human-controlled robotic systems in the presence of errors, and what implications do these preferences have for the design and deployment of robots in multicultural environments?*

Research Questions

This study is guided by three secondary research questions **(RQ1, RQ2, and RQ3)** derived from the primary question in Chapter 2 **RQ III** and **RQ IV**, each addressing a critical aspect of HRI in the context of physical collaboration tasks. These questions explore the interplay between robot control modes, intentional errors, and cultural background, providing an understanding of how these factors influence user perception, acceptance, and comfort during physical interaction with robots.

Unlike the previous study, this one does not include formal hypotheses. Due to the cross-cultural focus and the introduction of intentional error, the study takes a more exploratory approach to capture potential variations across groups. This approach allows for a broader examination of patterns without being limited to specific predicted outcomes.

1. **RQ1: How does the control mode (autonomous vs. human-controlled) influence participants' acceptance and comfort during physical interaction tasks?**

   The control mode of a robot—whether it operates autonomously or is controlled by a human—plays a significant role in shaping user perceptions and interactions. Autonomous robots are often perceived as more advanced and capable, but they may also evoke feelings of uncertainty or mistrust due to their lack of direct human oversight [73]. On the other hand, human-controlled robots may provide a sense of security and predictability, as users can attribute the robot's actions to a human operator. However, this mode may also limit the robot's perceived independence and efficiency.

More specifically, this research question seeks to investigate:

- Whether participants feel more comfortable and trusting when interacting with a robot they believe is controlled by a human, compared to one they believe is autonomous.
- How the control mode affects participants' willingness to engage in physical collaboration with the robot, particularly in tasks that require close proximity and direct manipulation of the robot's end-effector.

By addressing this question, the study will provide insights into how control modes can be optimized to enhance user acceptance and comfort in pHRI scenarios.

2. **RQ2: Does introducing an intentional error in the robot's behavior affect participants' preference for one control mode over the other?**

   Errors in robot behavior, whether intentional or unintentional, can significantly impact user trust and acceptance. In this study, an intentional error is introduced during one of the tasks in both control modes to observe how participants react to such anomalies. The error is designed to simulate a common issue in robotic systems, such as unexpected movements or deviations from the expected trajectory.

   This research question aims to explore:

   - Whether the presence of an error influences participants' preference for one control mode over the other. For example, do participants become more critical of the autonomous mode when an error occurs, or do they attribute it to the human operator in the human-controlled mode?
   - Whether the introduction of an error leads to increased anxiety or discomfort during physical interaction tasks, and how this varies between the two control modes.

   By examining the impact of intentional errors on user preference and perception, this study will contribute to a deeper understanding of how robot behavior influences user trust and acceptance, particularly in scenarios where errors are likely to occur.

3. **RQ3: How does cultural background influence participants' reactions to control modes and errors in physical robot interactions??**

Cultural background shapes how people respond to technology, influenced by factors like individualism versus collectivism [64]. This question explores how culture affects reactions to control modes and errors, investigating.

This research question investigates:

- Whether cultural groups differ in their preference for autonomous versus human-controlled robots.
- How cultural factors, such as familiarity with technology or attitudes toward automation, shape participants' overall perception of the robot's reliability and trustworthiness.
- When regarding physiological measurements such as GSR and HR, as well as body movements, will they be different as well for each culture group?

This study investigates how cultural background shapes people's reactions to robot control methods and errors. Exploring these cultural influences provides a deeper understanding of human-robot interaction, which can support future efforts to make robots more relatable and easier to integrate into diverse societies, especially in areas like healthcare and rehabilitation, where user acceptance is key.

4.2 Methodology

4.2.1 Participants

A total of 60 participants (27 males and 33 females) took part in this study. The summary of the participants is represented in 4.1, similar to the previous study. The participants are mostly university students and employees, recruited via posters distributed throughout the campus of the University of Waterloo. Each participant provided their informed consent before the study began. None of the participants had previous experience interacting with the robot used in this experiment. To minimize potential biases, participants were not informed of the specific goals of the study until after all experimental sessions had ended. The Ethics Human Research Ethics Board of the University of Waterloo approved the study.

Table 4.1: Participants Summary

Culture Category	**Female**	**Male**
Chinese	14	6
Mean Age	21.41	25.33
SD Age	3.91	4.93
Middle Eastern	9	11
Mean Age	24.00	25.91
SD Age	5.00	4.61
Western	10	10
Mean Age	20.78	21.09
SD Age	2.11	2.02

Number of participants by culture category with Mean and SD Age of each Gender.

4.2.2 Robot

The Sawyer robot arm was utilized for this experiment, which is a 7-degree-of-freedom (DOF) collaborative robot arm, with a reach of 1260 mm and a payload of 4kg. The robot is interfaced via ROS.



Figure 4.1: The Sawyer Sdk Robot arm

4.2.3 Technical Implementation

The system executes a predefined end-effector trajectory using inverse kinematics, visualizes the motion in real-time, and dynamically adjusts stiffness during motion using a Cartesian impedance controller. A single-point deviation is intentionally introduced along the path to simulate a disturbance and evaluate the robot's adaptive behavior.

4.3 Trajectory Generation and Execution

The robot follows a smooth trajectory defined between a start and an end pose. A set of intermediate control points in Cartesian space is interpolated using cubic splines to generate a continuous path. For each point along the trajectory:

- Inverse kinematics is used to compute the corresponding joint angles.
- These joint positions are appended to a motion trajectory buffer.
- The trajectory is sent to the robot for execution in joint space.

This allows the robot's end-effector to move smoothly through space, following the desired motion pattern.

4.4 Real-Time Visualization

A graphical interface implemented using PyGame[1] is used to visualize the robot's end-effector position in real-time. The planned trajectory and the actual path taken by the robot are displayed on the screen. This helps users observe the motion performance and any deviations or disturbances that occur 4.3.

4.5 Intentional Error Injection

To simulate the error task scenario, a deviation is introduced at a specific waypoint along the trajectory. This is done by modifying the interpolated Z-position (vertical axis) at one

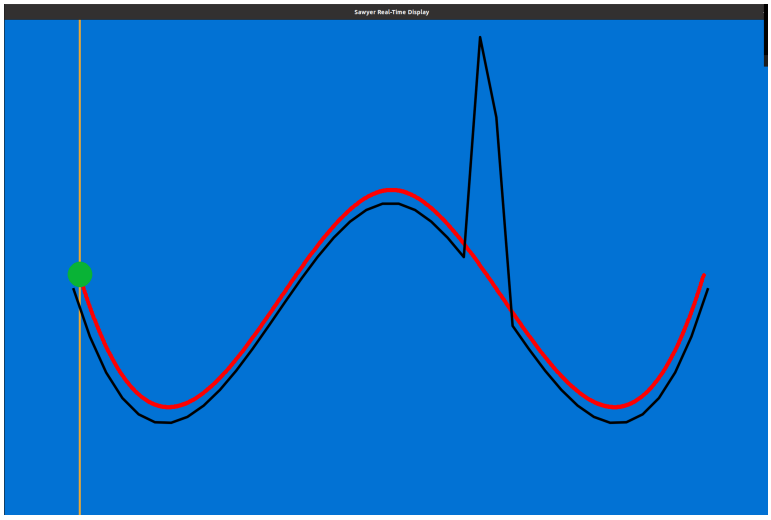


Figure 4.2: The red line is what the participant see displayed on the screen the black line is the actual trajectory

index in the trajectory. The resulting path contains a brief upward "jump" in the motion in Figure 4.2.

This deviation is treated as an intentional error, and its location is identified by its Cartesian coordinates (y-axis). These coordinates are broadcast during execution, allowing other components in the system to monitor the robot's progress and detect when it approaches the error point.

4.6 Cartesian Impedance Control

To enable safe and compliant physical interaction with participants, we used Sawyer's *Interaction Control Mode* in the Cartesian impedance configuration. This mode allows the robot's end-effector to behave like a virtual mass-spring-damper system, enabling the user to guide the arm with natural resistance and responsiveness physically.

The controller was configured with the following stiffness and damping model:

$$\mathbf{F}_{\text{imp}} = \mathbf{K}_x(\mathbf{x}_d - \mathbf{x}) + \mathbf{D}_x(\dot{\mathbf{x}}_d - \dot{\mathbf{x}}) \tag{4.1}$$

[1]https://www.pygame.org/docs/

where:

- $\mathbf{F}_{\text{imp}}$ is the virtual force applied at the end-effector,
- $\mathbf{x}_d$ and $\mathbf{x}$ are the desired and actual end-effector poses,
- $\dot{\mathbf{x}}_d$ and $\dot{\mathbf{x}}$ are the desired and actual end-effector velocities,
- $\mathbf{K}_x$ is the Cartesian stiffness matrix,
- $\mathbf{D}_x$ is the Cartesian damping matrix.

The desired force $\mathbf{F}_{\text{imp}}$ is used internally by the robot to compute joint torques that resist displacement from the equilibrium pose, thereby enabling intuitive physical interaction.

During our study, we set $\mathbf{x}_d$ and $\dot{\mathbf{x}}_d$ to constant values, allowing participants to move the arm while experiencing elastic and viscous resistance physically. This setup supports safe, intuitive guidance without requiring participants to input commands or use external interfaces.

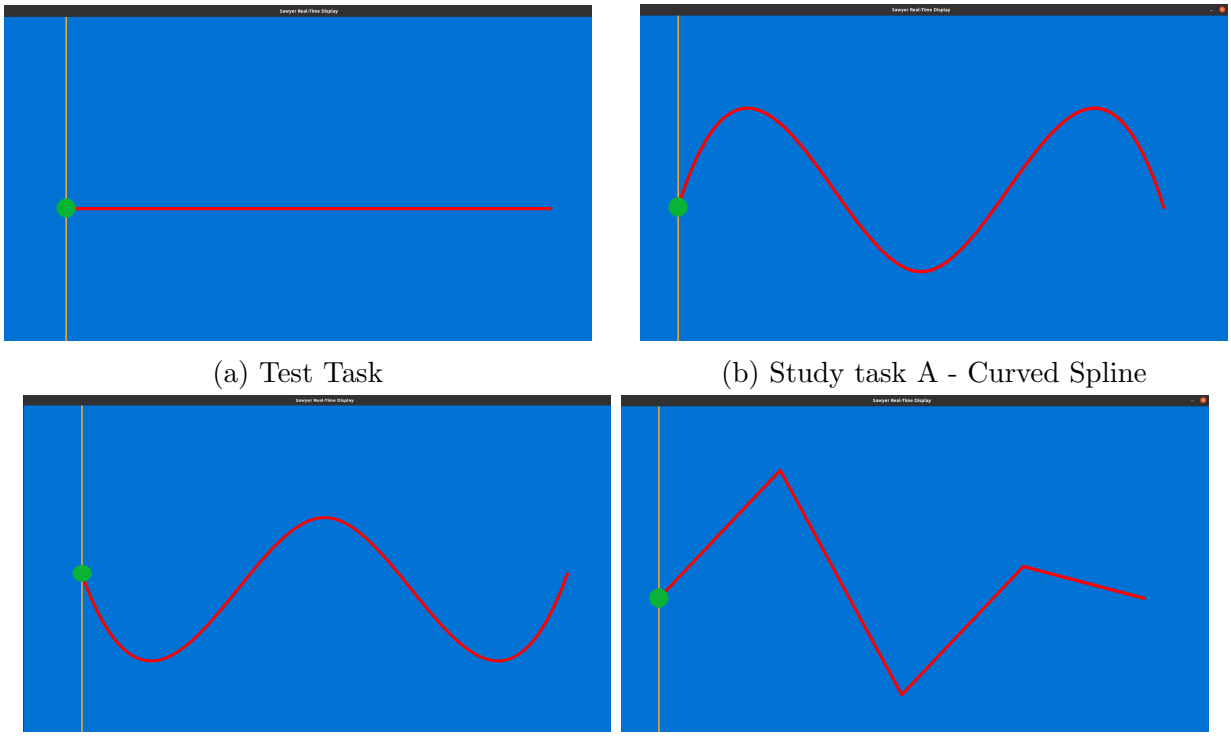
The interaction control was configured relative to the robot base frame, with no external force commands applied. The damping term was left active, and collision avoidance was suppressed to allow unimpeded physical motion by the user. This impedance-based setup allowed participants to guide the robot.

When the robot approaches the identified error location, the controller:

1. Detects the end-effector's proximity to the error point using real-time position feedback.
2. Temporarily increases stiffness in all Cartesian directions to reduce compliance and make the robot more rigid.
3. After a fixed duration, reverts back to the default, more compliant stiffness settings.

4.6.1 Control Modes

Although the study focuses on the effect of control modes, the implementation of the study did not involve an actual human control mode. To ensure that the tasks were done under the same conditions, the participants were led to believe that an actual human was controlling the robot. However, both modes were the same, i.e. the study is based on deception.



(a) Test Task

(b) Study task A - Curved Spline

(c) Study task B - curved Spline reversed

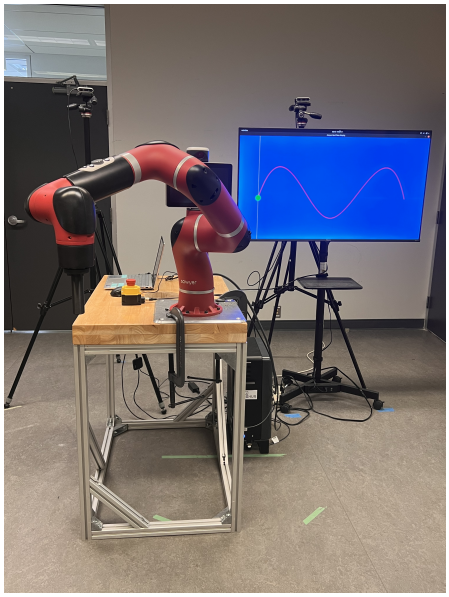
(d) Study task C - Linear spline

Figure 4.3: The following represents the Task done by the participants, where task (a) is a test task in which they test the system to see their ideal positioning

4.6.2 Experimental Setup

The experiment was conducted in a controlled environment where participants interacted with the Sawyer robotic arm while performing a series of tasks. The setup consisted of the following components shown in Fig. 4.4:

- **Display System**: A large LCD screen in front of the participants displays the game interface.
- **Camera System**: Two cameras mounted on tripods behind the LCD screen to capture participant movements. Video recordings were analyzed post-experiment to extract detailed body movement data.
- **Robot Platform**: The Sawyer robotic arm was securely fixed on a table in front of the participants to execute pre-programmed task sequences.



(a) Front view of the Experiment Setup



(b) Side view of the Experiment Setup

Figure 4.4: The overview of the experiment setup displaying the robotics arm , screen, and the two cameras.

- **Sensor System**: Participants were equipped with GSR+ Shimmer unit sensors to measure their responses during the experiment.
- **Handle**: A 3D printed handle was designed for users to move the robot comfortably during the tasks

4.6.3 Tasks

The participants completed three sequential tasks displayed on the LCD screen (see Fig. 4.3). Initially, a test task was introduced to familiarize participants with the robot interactions as seen in Fig. 4.3a. Subsequently, participants performed three specific tasks: two involving curved spline trajectories and one involving a straight spline trajectory, as illustrated in Fig. 4.3. Task 2, in both interaction modes, incorporated an intentional trajectory error, compelling participants to deviate from the originally intended path. Participants completed the Self-Assessment Manikin (SAM) questionnaire immediately following each task. The task order was randomized for each interaction mode.

The visual components of the tasks projected on the LCD screen included:

- **Green Circle:** Represented the robot's end-effector.
- **Red Line:** Indicated the trajectory participants were instructed to follow.
- **Horizontal Yellow Line:** Served as a reference cue for participants, signaling when to initiate robot movement. Participants were instructed to move the robot as soon as this line started moving.

4.6.4 Measures

Participants completed both subjective and objective measures through pre- and post-experiment questionnaires and data collection.

Subjective Measures:

To maintain consistency and allow for comparison across studies, we retained the same core subjective measures from Study 1 in Study 2—namely, the GAToRS and task-related Likert-style questions assessing comfort, trust, and perceived safety. These tools effectively captured participants' general impressions of the robots and the interaction experience. However, since Study 2 introduced a new focus on robot control modes and error perception, it became important to also capture individual differences and emotional reactions. To that end, we added two additional instruments: the Ten-Item Personality Inventory (TIPI))[21] and the Self - Assessment Mankin (SAM)[13]. TIPI allowed us to consider how personality traits might relate to participants' preferences or trust in autonomous versus human-controlled robots, while SAM provided a simple way to track emotional changes—such as stress or comfort—before and after interacting with the robot, especially in response to the intentional error introduced during the task. These additional measures helped us better understand how participants responded in the more nuanced context of Study 2, especially where control modes and error perception played a central role.

Pre-Experiment Questionnaires: Participants completed the following pre-experiment questionnaires:

- **Part 1 Cultural Familiarity and Percentage of Western Life**: Participants reported their familiarity with their cultural backgrounds and the duration of time spent living in Canada or the USA. This duration was later converted into a percentage relative to their age.

- **Part 2 Familiarity with New Technology**: Participants indicated their familiarity with emerging technologies, including Artificial Intelligence (AI), robots, and robotic manipulators.
- **Part 3 Expectations of the Robot and Comfort Level**: Participants rated their expectations regarding the robot's efficiency, ease of use, safety, and comfort level using a 5-point likert scale.
- **Part 4:** Participants completed the TIPI questionnaire to assess personality traits.
- **Part 5**: Participants rated their general emotional state before starting the experiment using SAM.

Post-Experiment Questionnaires: Following task completion, participants completed:

- **SAM Questionnaire**: Completed immediately after each individual task.
- **Task-Related Questions**: After each interaction mode, participants rated task ease, understanding, and comfort using a 5-point Likert scale.
- **GAToRS**: Completed after each interaction mode, utilizing a 7-point Likert scale.

Objective Measures

- Physiological data, including **HR** and **GSR**, were collected using Shimmer3 GSR+ sensors.
- Body behavioral data, such as posture and distance, were analyzed via video recordings processed using MediaPipe.

4.7 Experimental Protocol

For each participant, the experimental session followed the sequence described below. The study used a deception protocol where participants were told they would interact with both autonomous and human-controlled robots; however, without their knowledge, there was no real autonomous control mode—both modes were the same. To avoid order bias, the perceived interaction modes were alternated across trials. Each participant completed

the tasks under these different perceived conditions. Additionally, during Task 2 for both perceived modes, an intentional error was introduced: the robot deliberately deviated from the expected trajectory to assess participant responses to unexpected robot behavior. The lab environment did not show the "human" who supposedly controls the robot during the human-controlled mode, as the person was sitting behind a frosted glass so the participant was led to belive that this person is controlling the robot. There was a debriefing session at the end explaining the experiment.

1. Before the in-person experiment, an information letter was shown to the participants containing all the experiment details before we started the study via tablet.
2. Each participant signed the consent form on the tablet.
3. They were asked to fill out an initial form to collect personal information, familiarity with robots, use of technology, and the duration they had lived in Canada.
4. The participant completed the pre-experiment questions, including the TIPI.
5. Physiological sensors were placed on the participant to measure the galvanic skin response GSR and PPG. Baseline data, to be used for comparison, was collected for 5 minutes while the participant was seated outside the experimental room before interacting with the robot.
6. The participant performed three tasks under one interaction mode, answering SAM questionnaire after each task.
7. After completing all tasks under an interaction mode, the participant answered the post-experiment task-related questions, the Godspeed questionnaire (safety subscale), and the GAToRS.
8. The Shimmer sensors were removed.
9. The participant answered open-ended questions.
10. The participant was informed of the true nature of the experiment and signed an informed consent letter letting them know about the deception.

4.8 Results

The General Attitudes Towards Robots Scale (GAToRS) [59] was used to measure participants' attitudes toward robots, incorporating both positive and negative dimensions. The data were normally distributed across cultural categories and experimental conditions. T-tests were used for pairwise comparisons Figure 4.5 presents the distribution of scores.

For **P+** and **S+**, participants across cultural groups reported similar distributions, with no significant differences between autonomous and human-controlled modes. In contrast, **P-** shows slightly lower median scores for the Middle Eastern group than the Western and Chinese groups. The distributions of **S-** remain consistent across cultural categories, with no significant variations.

Perceived safety, assessed using the Godspeed questionnaire [7], was normally distributed across all cultural categories. Figure 4.7 illustrates that safety scores were similar for autonomous and human-controlled modes. The Western and Middle Eastern groups exhibited comparable distributions, while the Chinese group had slightly lower median scores for the autonomous mode. However, no statistically significant differences were found between conditions.

Task-related metrics, including Task Comfort, Task Ease, and Task Understanding, were analyzed. Figure 4.6 illustrates the distributions of scores across cultural categories and control modes.

Task Comfort and Task Ease showed similar distributions among all groups, with no significant differences between autonomous and human-controlled conditions. Task Understanding exhibited slightly higher variation in the Middle Eastern and Western groups, while the Chinese group maintained relatively consistent scores across conditions.

4.8.1 LMM Analysis for each Culture group

The tables presented the outcomes of LMM analyses conducted separately for each cultural category (Middle Eastern, Western, and Chinese). Each table lists the predictor variables, their corresponding parameter estimates (Estimate), standard errors (S.E.), t-values, and associated p-values. The estimate indicates the magnitude and direction of the predictor's influence on the dependent variable; positive estimates reflect positive relationships, while negative estimates indicate inverse relationships

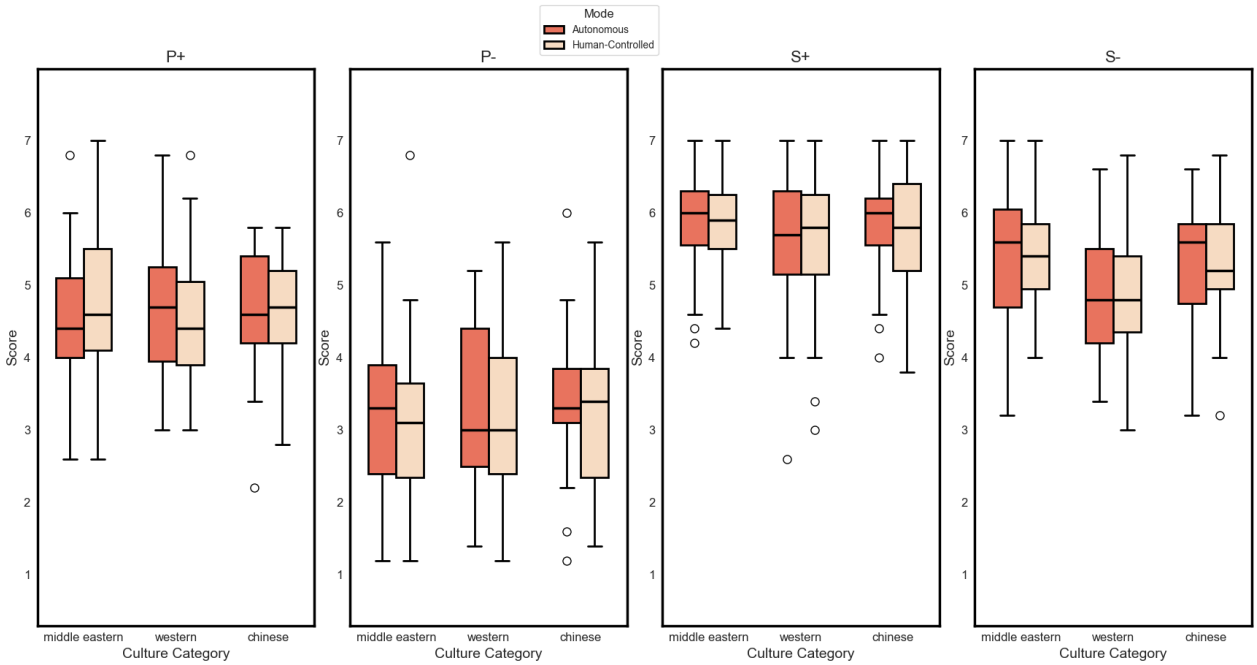


Figure 4.5: Comparison across all the GAToRS scales (P+ (Comfort and Enjoyment), P- (Unease and Anxiety) , S+ (Societal Optimism), S- (Societal Concerns) for each mode

Middle Eastern Participants

Personal-Level Positive Attitudes (P+) The analysis identified several significant predictors influencing personal-level positive attitudes toward robots. Expectations of safety and ease of use, perceived safety, extraversion, emotional stability, and openness positively influenced attitudes. Conversely, higher familiarity with new technology and agreeableness negatively influenced these perceptions.

Personal-Level Negative Attitudes (P-) Gender was a significant predictor, with males showing greater negative attitudes ($p = 0.014$). Familiarity with robotic arms, expected ease , comfort level during tasks, conscientiousness, and emotional stability were negatively correlated with negative attitudes, suggesting higher levels of these factors reduce discomfort. Agreeableness, however, was positively correlated ($p < 0.001$), suggesting that more agreeable individuals experienced more significant discomfort.

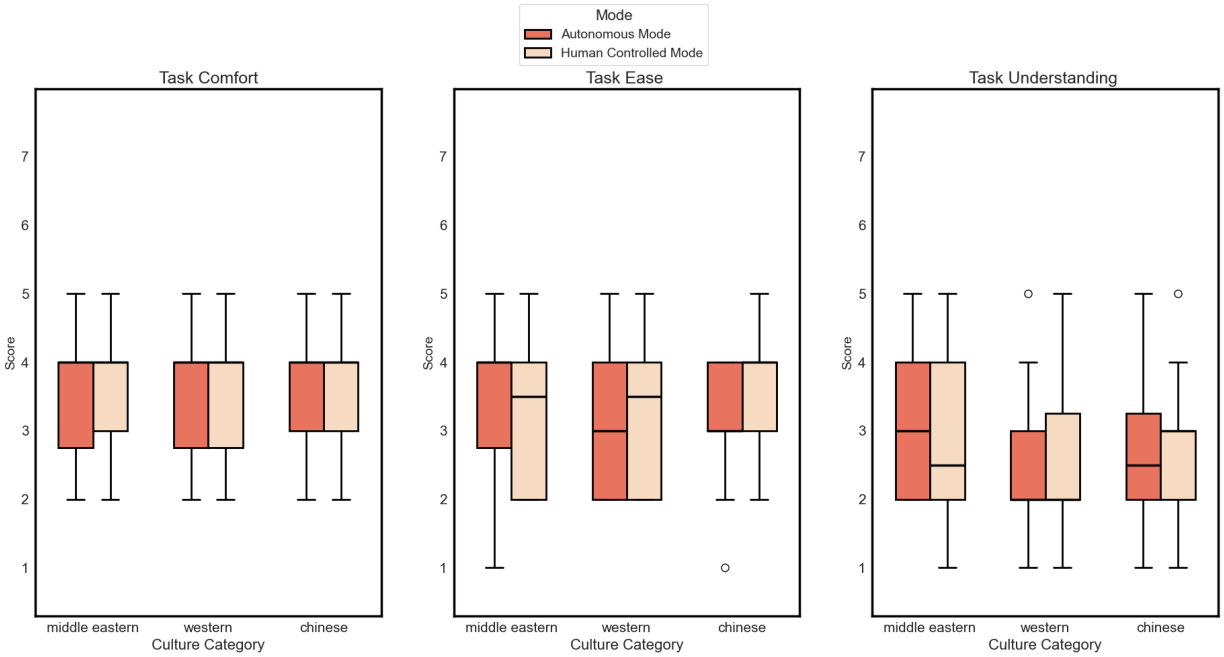


Figure 4.6: The Task question data distribution among culture groups

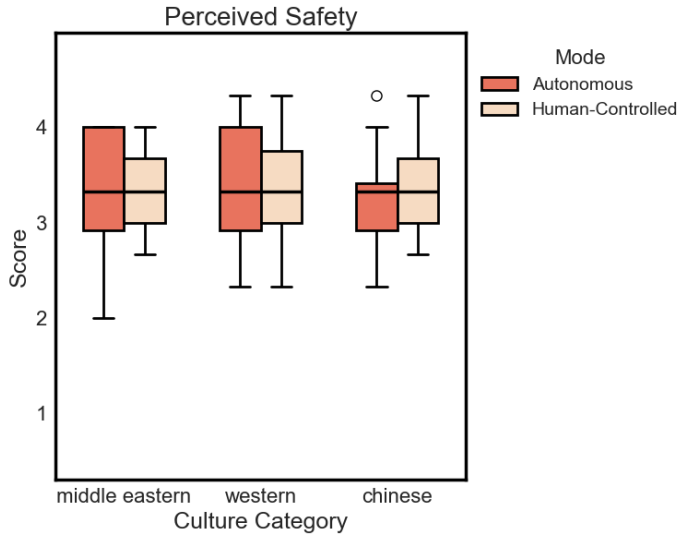


Figure 4.7: The Perceived safety data distribution among culture groups

Table 4.2: LMM Results for Middle Eastern Model – P+

Variable	**Estimate**	**S.E.**	**T**	**p-value**
Familiarity with New Technology	-0.235	0.072	-3.289	0.0032
Perceived Safety	0.424	0.153	2.778	0.0107
Expect Ease	0.763	0.214	3.555	0.0017
Expect Safe	0.682	0.134	5.109	< 0.0001
Extraversion	0.502	0.107	4.696	< 0.0001
Agreeableness	-0.558	0.132	-4.213	¡0.001
Emotional Stability	0.351	0.130	2.701	0.0128
Openness	0.656	0.142	4.609	¡0.001

Only statistically significant results ($p < .05$) are shown. Estimate indicates the regression coefficient; S.E. is the standard error.

Table 4.3: LMM Results for Middle Eastern Model – P-

Variable	**Estimate**	**S.E.**	**T**	**p-value**
Gender (Male)	0.682	0.259	2.631	0.0144
Familiarity with Robotic Arms	-0.408	0.085	-4.784	< 0.0001
Expect Ease	-1.262	0.216	-5.849	< 0.0001
Comfort Level Task	-0.702	0.153	-4.591	< 0.0001
Agreeableness	0.758	0.152	4.978	< 0.0001
Conscientiousness	-0.304	0.089	-3.422	0.0021
Emotional Stability	-0.631	0.083	-7.642	< 0.0001

Only statistically significant results ($p < .05$) are shown. Estimate indicates the regression coefficient; S.E. is the standard error.

Societal-Level Positive Attitudes (S+) Participants' expectations of robotic assistance, extraversion, familiarity with robotic arms, familiarity with new technology, height, and percentage of life spent in Western cultures ($p < 0.001$) positively impacted societal-level positive attitudes. In contrast, expected ease negatively influenced societal-level positive attitudes. Emotional stability, openness, and being male ($p = 0.021$) were also negatively correlated.

Societal-Level Negative Attitudes (S-) Height, familiarity with robotic arms, ease during tasks, extraversion, agreeableness ($p = 0.020$), and conscientiousness ($p < 0.001$)

Table 4.4: LMM Results for Middle Eastern Model – S+

Variable	**Estimate**	**S.E.**	**T**	**p-value**
Height	0.027	0.007	3.800	0.0008
Gender (Male)	-0.330	0.135	-2.453	0.0215
Percentage Western Life	0.015	0.002	8.223	< 0.0001
Familiarity with Robotic Arms	0.443	0.044	9.994	< 0.0001
Familiarity with New Technology	0.284	0.036	7.994	< 0.0001
Expect Assist	1.231	0.140	8.819	< 0.0001
Expect Ease	-1.071	0.095	-11.289	< 0.0001
Extraversion	0.665	0.038	17.342	< 0.0001
Emotional Stability	-0.778	0.073	-10.720	< 0.0001
Openness	-0.415	0.074	-5.630	< 0.0001

Only statistically significant results ($p < .05$) are shown. Estimate indicates the regression coefficient; S.E. is the standard error.

Table 4.5: LMM Results for Middle Eastern Model – S-

Variable	**Estimate**	**S.E.**	**T**	**p-value**
Height	-0.031	0.007	-4.553	< 0.0001
Familiarity with Robotic Arms	-0.238	0.061	-3.903	0.0005
Ease Level Task	-0.256	0.064	-4.003	0.0004
Extraversion	-0.323	0.072	-4.486	< 0.0001
Agreeableness	-0.183	0.074	-2.466	0.0201
Conscientiousness	-0.335	0.058	-5.727	< 0.0001
Openness	0.676	0.111	6.071	< 0.0001

Only statistically significant results ($p < .05$) are shown. Estimate indicates the regression coefficient; S.E. is the standard error.

negatively correlated with societal-level negative attitudes, suggesting these factors decrease societal concerns about robots. Openness ($p < 0.001$) was positively correlated with higher societal negative attitudes, indicating that individuals with higher openness had increased societal concerns about robots.

4.8.2 Western Participants

Personal-Level Positive Attitudes (P+) The analysis identified significant predictors for personal-level positive attitudes toward robots among Western participants. Expectations regarding robot safety (Expect Safe), expected assistance (Expect Assist), expected ease (Expect Ease), understanding of the task, and agreeableness positively influenced these attitudes. Conversely, emotional stability ($p < 0.001$) and the selection of human-controlled mode ($p = 0.005$) negatively impacted positive attitudes, indicating a preference for autonomous mode.

Table 4.6: LMM Results for Western Model – P+

Variable	**Estimate**	**S.E.**	**T**	**p-value**
Mode Selected (Human Controlled Mode)	-0.324	0.105	-3.084	0.0048
Expect Assist	0.471	0.094	5.004	< 0.0001
Expect Ease	0.415	0.127	3.284	0.0029
Expect Safe	1.085	0.106	10.272	< 0.0001
Understanding Level Task	0.323	0.052	6.259	< 0.0001
Agreeableness	0.379	0.066	5.724	< 0.0001
Emotional Stability	-0.449	0.058	-7.744	< 0.0001

Only statistically significant results ($p < .05$) are shown. Estimate indicates the regression coefficient; S.E. is the standard error.

Personal-Level Negative Attitudes (P-) Gender (males) ($p < 0.001$), familiarity with robotic arms, and expectations of robot safety (Expect Safe, $p = 0.002$) positively impacted personal-level negative attitudes, indicating increased discomfort or concern. Conversely, comfort level during tasks, expectations of assistance (Expect Assist), familiarity with new technology, height, agreeableness, openness, emotional stability, and conscientiousness negatively impacted negative attitudes, thus reducing discomfort or concern.

Societal-Level Positive Attitudes (S+) Expectations regarding robot safety, expected assistance, understanding level of the task, comfort during tasks, and familiarity with new technology positively impacted societal-level positive attitudes. Conversely, openness, conscientiousness ($\beta = -0.356$, $p < 0.001$), extraversion, Gender (Male) ($p < 0.001$), and selection of human-controlled mode ($p = 0.023$) negatively impacted societal-level positive attitudes, suggesting participants preferred autonomous modes.

Table 4.7: LMM Results for Western Model – P-

Variable	**Estimate**	**S.E.**	**T**	**p-value**
Height	-0.044	0.009	-5.007	< 0.0001
Gender (Male)	1.474	0.155	9.494	< 0.0001
Familiarity with Robotic Arms	0.521	0.089	5.838	< 0.0001
Familiarity with New Technology	-0.327	0.058	-5.647	< 0.0001
Expect Assist	-0.815	0.103	-7.892	< 0.0001
Expect Safe	0.419	0.119	3.524	0.0018
Comfort Level Task	-0.615	0.055	-11.120	< 0.0001
Agreeableness	-0.458	0.063	-7.302	< 0.0001
Conscientiousness	0.573	0.058	9.850	< 0.0001
Emotional Stability	0.552	0.059	-9.404	< 0.0001
Openness	-0.314	0.082	-3.833	0.0009

Only statistically significant results ($p < .05$) are shown. Estimate indicates the regression coefficient; S.E. is the standard error.

Table 4.8: LMM Results for Western Model – S+

Variable	**Estimate**	**S.E.**	**T**	**p-value**
Gender (Male)	-0.535	0.094	-5.705	< 0.0001
Mode Selected (Human Controlled)	-0.184	0.075	-2.453	0.0230
Familiarity with New Technology	0.249	0.057	4.341	0.0003
Expect Assist	0.858	0.124	6.928	< 0.0001
Expect Safe	1.251	0.100	12.479	< 0.0001
Comfort Level Task	0.297	0.076	3.911	0.0008
Understanding Level Task	0.328	0.050	6.562	< 0.0001
Extraversion	-0.323	0.065	-4.952	< 0.0001
Conscientiousness	-0.356	0.047	-7.533	< 0.0001
Openness	-1.010	0.084	-12.038	< 0.0001

Only statistically significant results ($p < .05$) are shown. Estimate indicates the regression coefficient; S.E. is the standard error.

Societal-Level Negative Attitudes (S-) Expectations of ease, openness, familiarity with robotic arms, comfort level during tasks, emotional stability, and selecting human-controlled mode ($p = 0.002$) positively impacted societal-level negative attitudes, thus

increasing societal concerns about robots. Factors such as familiarity with new technology, ease of task, understanding of the task, extraversion, and Gender (Male) negatively impacted societal-level negative attitudes, thus decreasing societal concerns.

Table 4.9: LMM Results for Western Model – S+

Variable	**Estimate**	**S.E.**	**T**	**p-value**
Gender (Male)	-0.535	0.094	-5.705	< 0.0001
Mode Selected (Human Controlled)	-0.184	0.075	-2.453	0.0230
Familiarity with New Technology	0.249	0.057	4.341	0.0003
Expect Assist	0.858	0.124	6.928	< 0.0001
Expect Safe	1.251	0.100	12.479	< 0.0001
Comfort Level Task	0.297	0.076	3.911	0.0008
Understanding Level Task	0.328	0.050	6.562	< 0.0001
Extraversion	-0.323	0.065	-4.952	< 0.0001
Conscientiousness	-0.356	0.047	-7.533	< 0.0001
Openness	-1.010	0.084	-12.038	< 0.0001

Only statistically significant results ($p < .05$) are shown. Estimate indicates the regression coefficient; S.E. is the standard error.

4.8.3 Chinese Participants

Personal-Level Positive Attitudes (P+) Height, percentage of life spent in Western cultures ($p = 0.002$), familiarity with new technology, expectations of robot assistance (Expect Assist), ease of task, understanding of the task, extraversion, and emotional stability positively impacted personal-level positive attitudes among Chinese participants. Conversely, expectations regarding robot safety (Expect Safe), agreeableness, conscientiousness, and openness negatively impacted these attitudes.

Personal-Level Negative Attitudes (P-) Percentage of life spent in Western cultures ($p < 0.001$), emotional stability, and perceived safety positively impacted personal-level negative attitudes, increasing discomfort or concern. Counterwise, height, human-controlled mode (Mode selected) ($p = 0.009$), familiarity with robotic arms , familiarity with new technology, expectations of assistance (Expect Assist), extraversion, and agreeableness negatively impacted negative attitudes, thus reducing discomfort or concern.

Table 4.10: LMM Results for Chinese Model – P+ (Positive Perceptions)

Variable	**Estimate**	**S.E.**	**T**	**p-value**
Height	0.044	0.011	4.107	0.0005
Percentage Western Life	0.010	0.003	3.629	0.0016
Familiarity with New Technology	0.352	0.052	6.763	< 0.0001
Expect Assist	1.455	0.209	6.974	< 0.0001
Expect Safe	-0.322	0.110	-2.929	0.0080
Ease Level Task	0.461	0.096	4.817	< 0.0001
Understanding Level Task	0.211	0.077	2.720	0.0128
Extraversion	0.322	0.056	5.736	< 0.0001
Agreeableness	-0.353	0.102	-3.472	0.0023
Conscientiousness	-0.955	0.126	-7.582	< 0.0001
Emotional Stability	0.284	0.083	3.430	0.0025
Openness	-0.279	0.079	-3.518	0.0020

Only statistically significant results ($p < .05$) are shown. Estimate indicates the regression coefficient; S.E. is the standard error.

Table 4.11: LMM Results for Chinese Model – P+

Variable	**Estimate**	**S.E.**	**T**	**p-value**
Height	0.044	0.011	4.107	0.0005
Percentage Western Life	0.010	0.003	3.629	0.0016
Familiarity with New Technology	0.352	0.052	6.763	< 0.0001
Expect Assist	1.455	0.209	6.974	< 0.0001
Expect Safe	-0.322	0.110	-2.929	0.0080
Ease Level Task	0.461	0.096	4.817	< 0.0001
Understanding Level Task	0.211	0.077	2.720	0.0128
Extraversion	0.322	0.056	5.736	< 0.0001
Agreeableness	-0.353	0.102	-3.472	0.0023
Conscientiousness	-0.955	0.126	-7.582	< 0.0001
Emotional Stability	0.284	0.083	3.430	0.0025
Openness	-0.279	0.079	-3.518	0.0020

Only statistically significant results ($p < .05$) are shown. Estimate indicates the regression coefficient; S.E. is the standard error.

Societal-Level Positive Attitudes (S+) Expectations of ease, expectations regarding robot safety, ease during tasks, understanding of the task, conscientiousness, and emotional stability positively impacted societal-level positive attitudes. On the other hand, human-controlled mode ($p = 0.013$), familiarity with robotic arms, and familiarity with new technology negatively impacted societal-level positive attitudes, indicating a preference for autonomous modes.

Table 4.12: LMM Results for Chinese Model – S+

Variable	**Estimate**	**S.E.**	**T**	**p-value**
Mode Selected (Human Controlled)	-0.332	0.124	-2.681	0.0128
Familiarity with Robotic Arms	-0.174	0.077	-2.266	0.0323
Familiarity with New Technology	-0.217	0.059	-3.691	0.0011
Expect Ease	0.552	0.134	4.124	0.0004
Expect Safe	0.357	0.109	3.282	0.0030
Ease Level Task	0.554	0.098	5.661	< 0.0001
Understanding Level Task	0.350	0.083	4.191	0.0003
Conscientiousness	0.903	0.137	6.608	< 0.0001
Emotional Stability	0.332	0.092	3.627	0.0013

Only statistically significant results ($p < .05$) are shown. Estimate indicates the regression coefficient; S.E. is the standard error.

Societal-Level Negative Attitudes (S-) Familiarity with robotic arms, familiarity with new technology, expectations of assistance (Expect Assist), and extraversion positively impacted societal-level negative attitudes, thus increasing societal concerns about robots. Conversely, the percentage of life spent in Western cultures ($p < 0.001$), expectations of ease, expectations regarding safety, understanding of the task, agreeableness, conscientiousness, emotional stability, and openness negatively impacted societal-level negative attitudes, thus reducing societal concerns.

4.8.4 SAM Post and Pre Task Comparisons

The SAM data was tested for normality using the Shapiro-Wilk test and Q-Q plots and was found not normally distributed across the participants. Hence, the Wilcoxon signed-rank was used for the statistical analysis. Participants rated their emotional states after each

Table 4.13: LMM Results for Chinese Model – S)

Variable	**Estimate**	**S.E.**	**T**	**p-value**
Percentage Western Life	-0.037	0.006	-6.723	< 0.0001
Familiarity with Robotic Arms	0.559	0.149	3.737	0.0011
Familiarity with New Technology	0.294	0.081	3.643	0.0014
Expect Assist	1.761	0.364	4.841	< 0.0001
Expect Ease	-0.818	0.169	-4.829	< 0.0001
Expect Safe	-0.930	0.211	-4.418	< 0.0001
Understanding Level Task	-0.389	0.120	-3.247	0.0037
Extraversion	0.667	0.106	6.303	< 0.0001
Agreeableness	-1.337	0.203	-6.582	< 0.0001
Conscientiousness	-0.458	0.185	-2.481	0.0212
Emotional Stability	-0.426	0.126	-3.390	0.0026
Openness	-0.334	0.156	-2.139	0.0438

Only statistically significant results ($p < .05$) are shown. Estimate indicates the regression coefficient; S.E. is the standard error.

task iteration, with Task 2 always containing an error in both autonomous and human-controlled modes. The purpose of this analysis was to examine how errors in Task 2 influenced participants' affective states and to compare differences across cultural categories. In the results, we refer to the pre-experiment SAM as (Pre_ SAM) , and for each Task for example, will be SAM_ 1, which represents SAM questionnaire answered after Task 1, the same for Task 2 and Task 3. Finally, for Task 2, we will refer to it as (Task 2 with error) because this is where we implemented the intentional error

The box plots in Figures 4.8 and 4.9 illustrate the variations in valence, arousal, and dominance scores across cultural groups in both conditions.

In the autonomous mode (Figure 4.8), we observe significant fluctuations in participants' valence, arousal, and dominance scores, particularly between SAM_1 and SAM_2 (pre- and post-task with an error) across all cultural groups. For example, middle eastern participants show a notable drop in valence and dominance scores after Task 2, which then increases again in SAM_3, indicating a potential recovery in emotional response. The Wilcoxon test results in Table 4.14 confirm that these differences are statistically significant ($p < 0.05$), particularly in valence and dominance.

Similarly, in the human-controlled mode (Figure 4.9), the introduction of an error in Task 2 significantly impacted valence, arousal, and dominance scores across all cultural

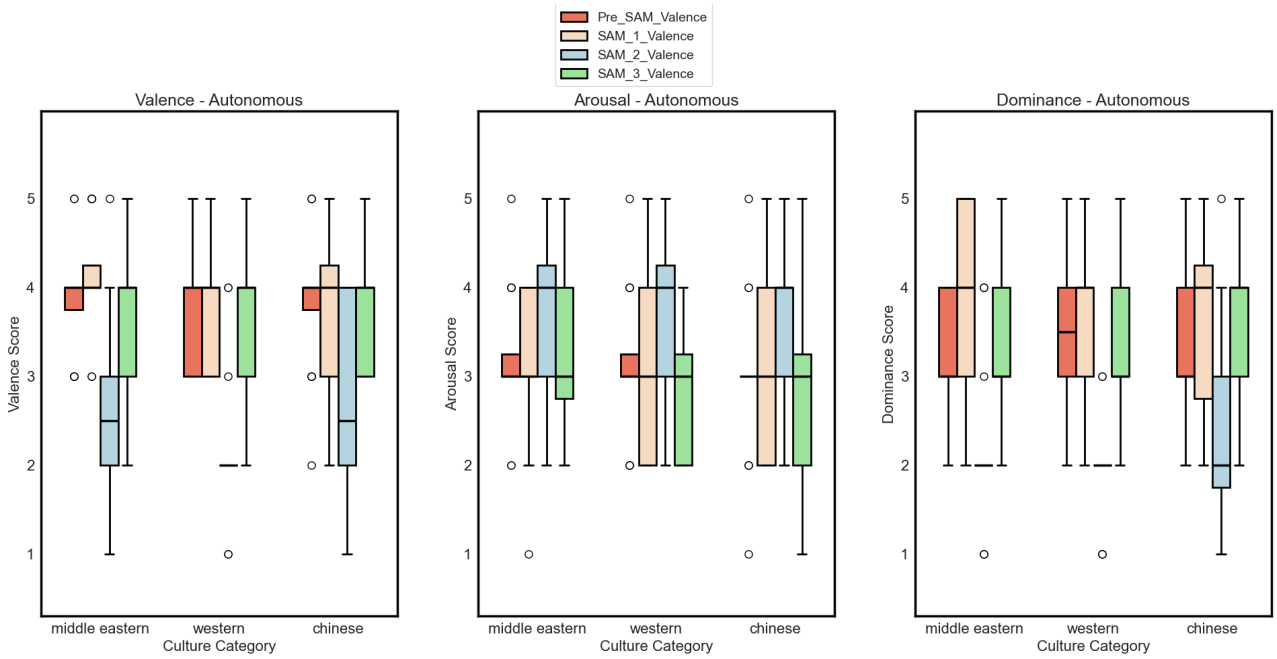


Figure 4.8: The SAM scales data distribution among culture groups for Pre Task, Task 1, 2, and 3

groups. Western participants exhibited a substantial decrease in valence from SAM_1 to SAM_2, which is statistically confirmed by the Wilcoxon test (Table 4.15). The arousal scores in the Chinese group also dropped significantly after Task 2 but showed an increasing trend in SAM_3, indicating an adjustment to the error.

Overall, the results suggest that Task 2 error, where there was an intentional error in the robot movements, consistently induced negative affective shifts across all cultural groups, as seen in lower valence and dominance scores. However, the recovery observed in SAM_3 implies that participants adapted to the error over time.

To further explore cross-cultural differences, we performed a Mann-Whitney test to compare emotional responses between Cultures. The results indicate that in the autonomous mode, there is a statistically significant difference in valence scores between these two cultural groups at SAM_2 (Mann-Whitney test statistic = 129.5, $p = 0.0314$). This suggests that the introduction of the error in Task 2 influenced valence perception differently across these cultures. Similarly, in the human-controlled mode, dominance scores at SAM_2 also showed a significant difference between Western and Chinese participants (Mann-Whitney test statistic = 125.5, $p = 0.0206$).

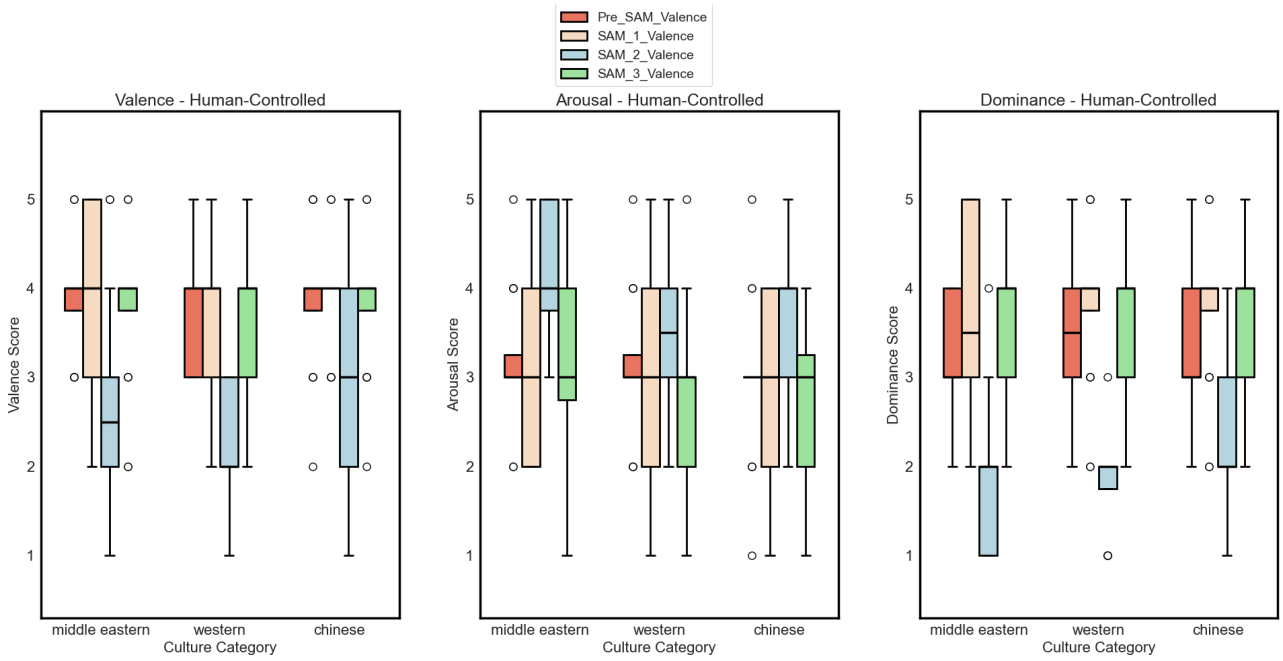


Figure 4.9: The SAM scales data distribution among culture groups for Pre Task, Task 1 , 2, and 3

4.8.5 Physiological data results

The HR and GSR data was tested for normality using the Shapiro-Wilk test and Q-Q plots and was found not normally distributed across the participants. Hence, Wilcoxon signed rank was used for the statistical analysis to test differences within each culture, and Mann-Whitney was used to test for cross-cultural differences.

Heart Rate Data Analysis

The HR responses of participants were analyzed across three tasks, each performed under two conditions: **Autonomous mode** and **Human-Controlled mode**. Figure 4.10 illustrates the distribution of HR changes across cultural groups (Middle Eastern, Chinese, and Western) in each task. Table 4.16 provides the statistical significance of HR differences between the two modes

Table 4.14: Wilcoxon Test Results for Autonomous Mode – SAM Comparisons

Culture Category	Comparison	Dimension	Statistic	p-value
Middle Eastern	SAM_1 vs. SAM_2	Valence	0.0	0.0004
Middle Eastern	SAM_2 vs. SAM_3	Valence	13.5	0.0021
Middle Eastern	SAM_1 vs. SAM_2	Dominance	13.5	0.0009
Middle Eastern	SAM_2 vs. SAM_3	Dominance	15.0	0.0031
Western	SAM_1 vs. SAM_2	Valence	0.0	< 0.0001
Western	SAM_2 vs. SAM_3	Valence	0.0	0.0003
Western	SAM_1 vs. SAM_2	Arousal	17.0	0.0119
Western	SAM_2 vs. SAM_3	Arousal	5.5	0.0013
Western	SAM_1 vs. SAM_2	Dominance	0.0	0.0002
Western	SAM_2 vs. SAM_3	Dominance	0.0	0.0002
Chinese	SAM_1 vs. SAM_2	Valence	4.0	0.0012
Chinese	SAM_2 vs. SAM_3	Valence	4.5	0.0020
Chinese	SAM_1 vs. SAM_2	Arousal	25.5	0.0437
Chinese	SAM_2 vs. SAM_3	Arousal	15.0	0.0150
Chinese	SAM_1 vs. SAM_2	Dominance	4.5	0.0013
Chinese	SAM_2 vs. SAM_3	Dominance	8.5	0.0015

All results are statistically significant at $p < 0.05$.

Task 1: The analysis of HR responses in Task 1, as shown in Figure 4.10, indicates that participants across all cultural categories exhibited significant differences between the *Autonomous* and *Human-Controlled* modes. The statistical comparison in Table 4.16 confirms this, with notable differences for Middle Eastern participants ($p = 0.0002$) and Western participants ($p = 0.0028$).

Task 2: In Task 2, where errors were intentionally introduced, heart rate variations were more pronounced. As seen in Figure 4.10, the spread of HR values increased across all cultural groups. The statistical results in Table 4.16 confirm significant differences for Middle Eastern ($p = 2.67 \times 10^{-5}$) and Western ($p = 0.0039$) participants.

Task 3: For Task 3, heart rate differences remained statistically significant across cultural groups. The results in Table 4.16 show significant HR differences for Middle Eastern ($p = 0.0005$), Chinese ($p = 0.0108$), and Western ($p = 7.25 \times 10^{-5}$) participants. The higher significance levels in this task suggest that prolonged exposure to robotic interactions or

Table 4.15: Wilcoxon Test Results for Human-Controlled Mode – SAM Comparisons

Culture Category	Comparison	Dimension	Statistic	p-value
Middle Eastern	SAM_1 vs. SAM_2	Valence	5.5	0.0004
Middle Eastern	SAM_2 vs. SAM_3	Valence	6.0	0.0009
Middle Eastern	SAM_1 vs. SAM_2	Arousal	10.0	0.0021
Middle Eastern	SAM_2 vs. SAM_3	Arousal	3.5	0.0027
Middle Eastern	SAM_1 vs. SAM_2	Dominance	3.5	0.0002
Middle Eastern	SAM_2 vs. SAM_3	Dominance	0.0	0.0003
Western	SAM_1 vs. SAM_2	Valence	0.0	0.0003
Western	SAM_2 vs. SAM_3	Valence	0.0	0.0003
Western	SAM_1 vs. SAM_2	Arousal	12.0	0.0093
Western	SAM_2 vs. SAM_3	Arousal	0.0	0.0015
Western	SAM_1 vs. SAM_2	Dominance	0.0	< 0.0001
Western	SAM_2 vs. SAM_3	Dominance	3.0	0.0002
Chinese	SAM_1 vs. SAM_2	Valence	3.0	0.0015
Chinese	SAM_2 vs. SAM_3	Valence	11.0	0.0083
Chinese	SAM_1 vs. SAM_2	Arousal	13.0	0.0033
Chinese	SAM_2 vs. SAM_3	Arousal	7.0	0.0035
Chinese	SAM_1 vs. SAM_2	Dominance	26.0	0.0235
Chinese	SAM_2 vs. SAM_3	Dominance	3.5	0.0007

All results are statistically significant at $p < 0.05$.

increased task complexity leads to noticeable physiological differences between *Autonomous* and *Human-Controlled* modes.

GSR Data Analysis

The relative GSR distributions across cultural categories reveal some patterns, as seen in Figure 4.11. The Middle Eastern group exhibits a wider spread of GSR values with a higher median compared to the Chinese group. Across all three tasks, the distributions for autonomous and human-controlled modes remain similar, with slightly increased variability observed in Task 2, which includes an intentional error. The Chinese group consistently shows the lowest median GSR values among the three cultural groups, with a more compact distribution and a smaller interquartile range, indicating lower variability

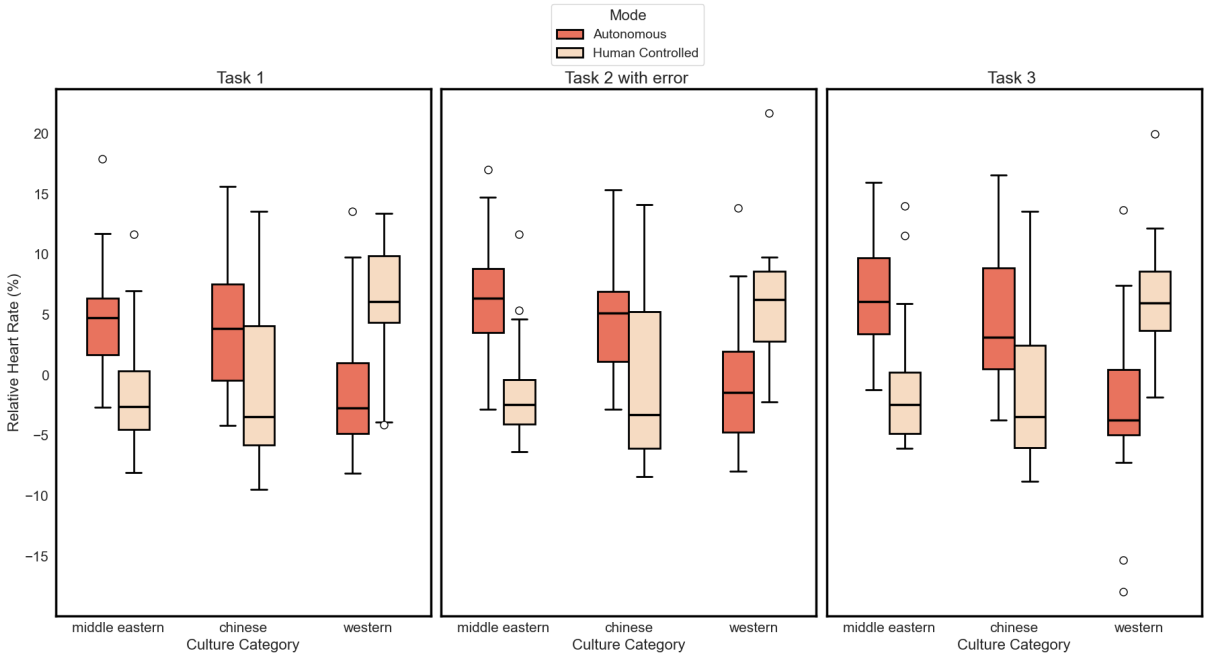


Figure 4.10: Realtive Heart Rate(%) across the Culture categories

Table 4.16: Heart Rate Differences Between Autonomous and Human-Controlled Sessions

Culture Category	Session	Statistic	p-value
Autonomous vs Human-Controlled			
Middle Eastern	1	14	0.0002
Middle Eastern	2	6	< 0.0001
Middle Eastern	3	15	0.0005
Chinese	3	33	0.0108
Western	1	24	0.0028
Western	2	26	0.0039
Western	3	7	< 0.0001

Results indicate statistically significant differences ($p < 0.05$) in heart rate responses between the Autonomous and Human-Controlled conditions.

and fewer extreme outliers. The Western group presents slightly higher variability than

the Chinese group, with a spread similar to that of the Middle Eastern group. The median values remain relatively stable across tasks, with no significant distinction between the autonomous and human-controlled modes. When comparing control modes, the overall distributions largely overlap across all cultural groups, suggesting that whether the robot operated autonomously or was human-controlled did not substantially influence gsr. To further examine these differences, statistical tests were conducted to compare relative GSR levels across cultural categories and control modes. However, no statistically significant differences were found across any of the three tasks, unlike the Heart rate.

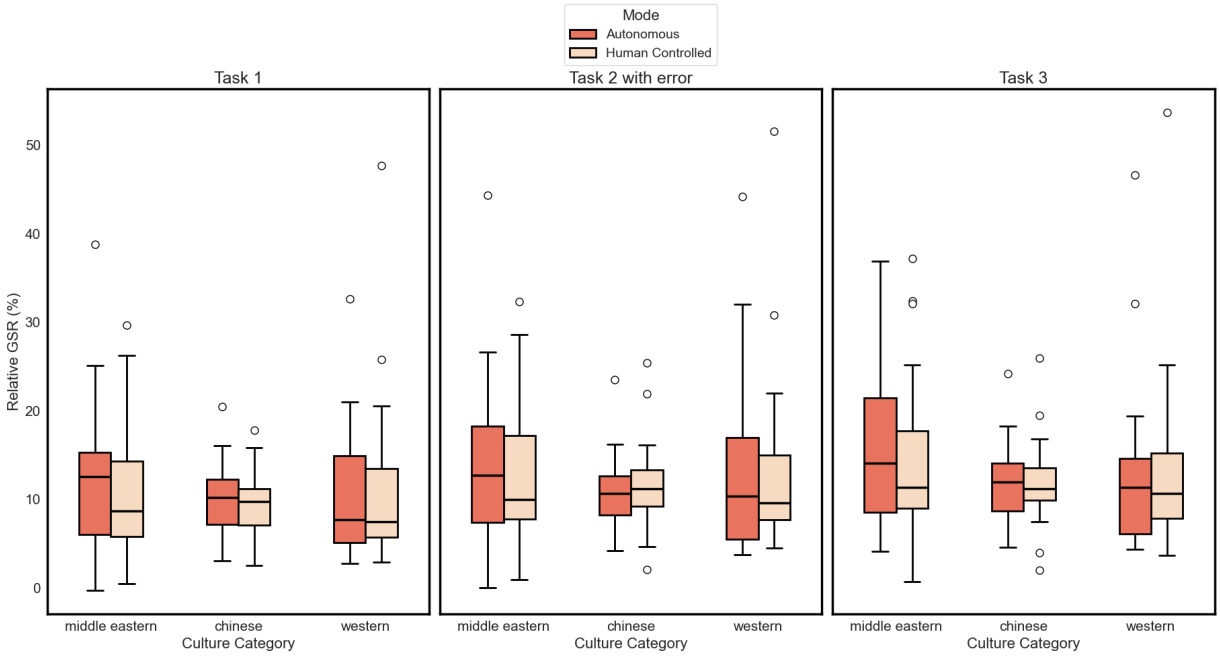


Figure 4.11: Realtive GSR(%) across the Culture categories

4.8.6 Body Pose Data

The body pose data was tested for normality using the Shapiro-Wilk test and Q-Q plots and was found not normally distributed across the participants. Hence, Wilcoxon signed rank was used for the statistical analysis to test differences within each culture, and Mann-Whitney was used to test for cross-cultural differences

The results of the body pose analysis across cultural categories and robot interaction modes are summarized in Figures 4.12, 4.13, and 4.14. Across all tasks, User Position re-

mained relatively stable between autonomous and human-controlled modes, with a notable difference observed in the Western group during Session 1 vs. Session 3 ($p = 0.0311$) and Session 2 vs. Session 3 ($p = 0.0526$), suggesting minor variations in movement adaptation.

Table 4.17: Comparison of Body Pose Variables Across Mode, Culture, and Sessions

Culture	Session Comparison	Variable	Statistic	p-value
Autonomous Mode				
Chinese	Session 2 vs Session 3	Right Arm Extension	45	0.0446
Chinese	Session 1 vs Session 3	User Position	23	0.0093
Western	Session 1 vs Session 3	Torso Tilt Angle	65	0.0459
Western	Session 1 vs Session 3	Torso Roll Angle	64	0.0424
Human-Controlled Mode				
Chinese	Session 1 vs Session 2	Right Arm Extension	38	0.0385
Western	Session 1 vs Session 3	User Position	36	0.0311
Western	Session 2 vs Session 3	User Position	41	0.0426

Results show statistically significant comparisons ($p < 0.05$) of body pose variables across sessions.

Torso movements, including tilt and roll angles, showed a comparable distribution across cultural groups, with significant differences in Western participants for Session 1 vs. Session 3 in torso tilt ($p = 0.0459$) and torso roll ($p = 0.0424$), indicating slight changes in posture across sessions. Right arm extension exhibited the most variation, particularly in the Chinese group for Session 2 vs. Session 3 ($p = 0.0446$) and in the human-controlled mode for Session 1 vs. Session 2 ($p = 0.0385$), suggesting a response to varying task conditions. As illustrated in Figures 4.12–4.14, Task 2 (with error) introduced slightly higher movement variability, particularly in torso movements and arm extension, but overall distributions between autonomous and human-controlled modes remained largely overlapping. The statistical results reinforce that while some minor session-based differences exist, cultural background and robot interaction mode did not have a strong or consistent impact on body movement patterns.

For User Position(Z), middle eastern displayed a higher significance in session 3 with ($p = 0.0431$),when comparing autonomous vs human controlled mode. They stood further away in the Autonomous mode after the error session.

Mann-Whitney Test for Autonomous Mode The Mann-Whitney test results for User Position in autonomous mode (Table 4.18) reveal that Western participants positioned

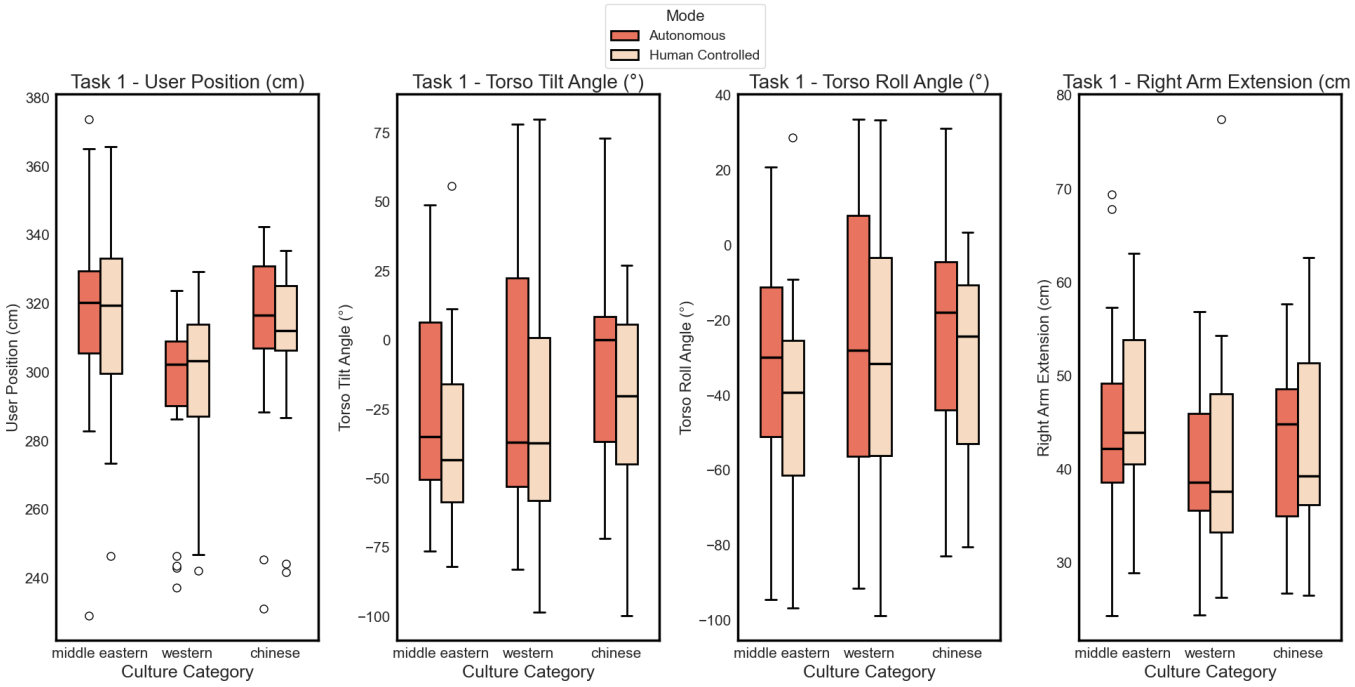


Figure 4.12: Boxplots of body pose variables across cultural categories for Task 1 .

themselves closer to the robot compared to both Middle Eastern and Chinese participants. In Session 1, Western participants showed significantly lower User Position compared to Middle Eastern ($p = 0.0130$) and Chinese participants ($p = 0.0058$), suggesting that Western participants were more comfortable maintaining a closer proximity to an autonomous robot.

Table 4.18: Mann-Whitney Test Results for Cultural Comparisons in Autonomous Mode

Session	**Cultural Comparison**	**Variable**	**Statistic**	**p-value**
Session 1	Middle Eastern vs Western	User Position	241	0.0130
Session 1	Chinese vs Western	User Position	237	0.0058
Session 2	Middle Eastern vs Western	User Position	239	0.0147
Session 2	Chinese vs Western	User Position	228	0.0365

Results indicate significant differences ($p < 0.05$) in user position along the Z-axis across cultural groups in the Autonomous mode.

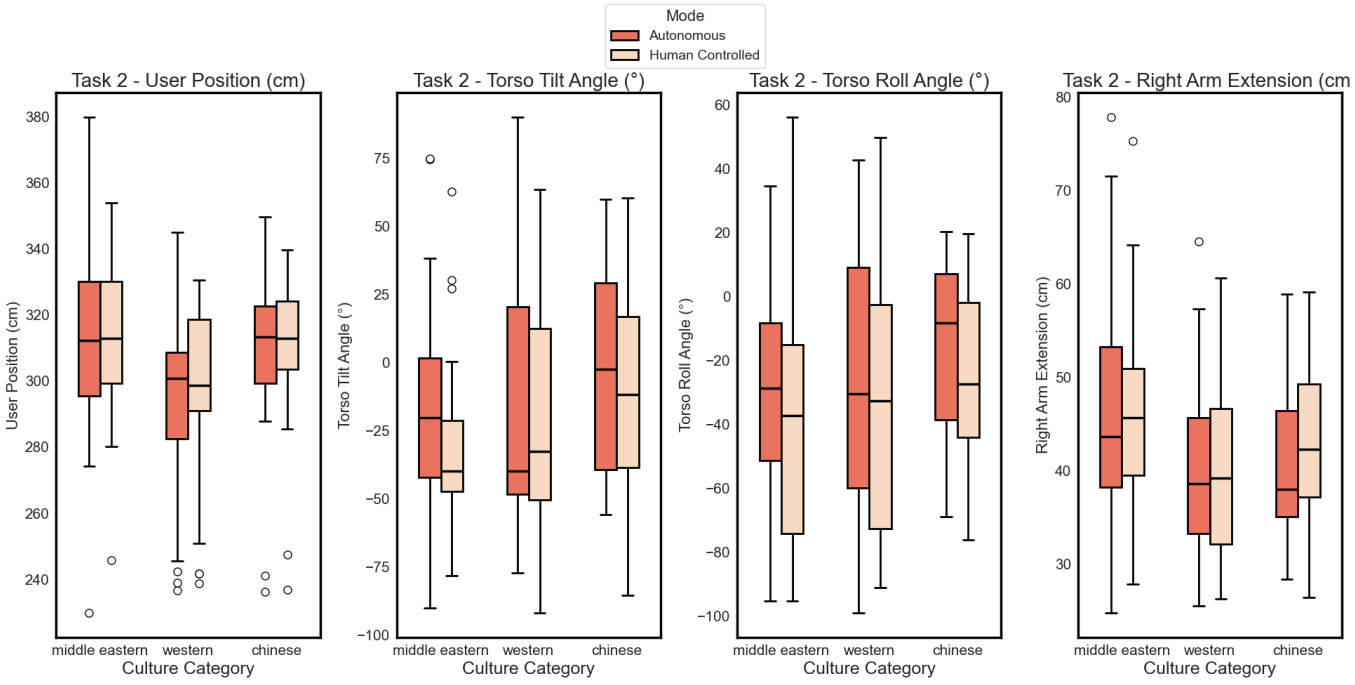


Figure 4.13: Boxplots of body pose variables across cultural categories for Task 2 (Task with Induced Error).

This pattern persisted in Session 2, where Chinese participants also exhibited a significantly higher User Position compared to Western participants ($p = 0.0365$), reinforcing the trend that Western participants interacted from a shorter distance, while Chinese participants maintained a greater distance. Similarly, in Head Position (Z), Middle Eastern participants showed a significantly higher positioning compared to Western participants ($p = 0.0147$), indicating that Western participants tended to adjust their head position lower, potentially engaging more actively with the robot's workspace. These findings suggest that Western participants were more likely to position themselves closer to the robot in an autonomous interaction, while Chinese and Middle Eastern participants preferred maintaining a slightly greater distance.

Mann-Whitney Test for Human-Controlled In the human-controlled condition (Table 4.19), significant cultural differences were observed in User Position , Right Arm Extension, and Torso Tilt Angle. In Session 1, Middle Eastern participants positioned themselves significantly further from the robot compared to Western participants ($p = 0.0198$),

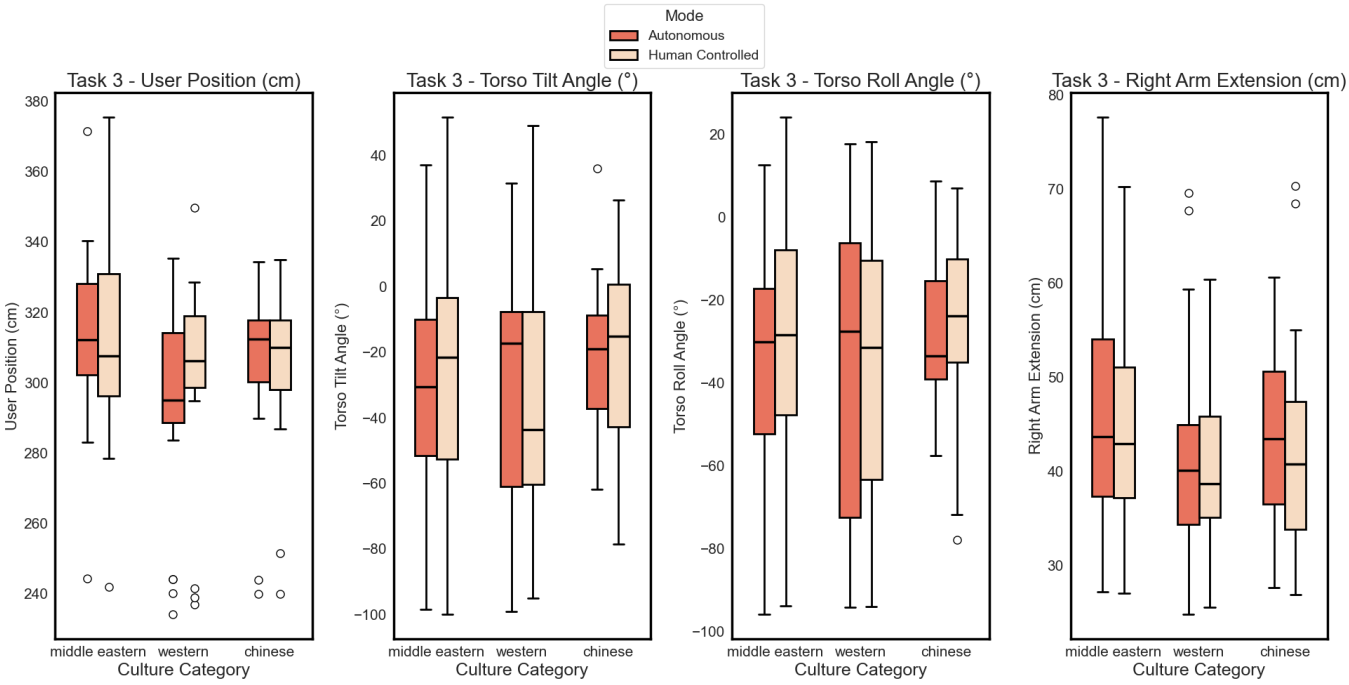


Figure 4.14: Boxplots of body pose variables across cultural categories for Task 3.

Table 4.19: Mann-Whitney Test Results for Cultural Comparisons in Human-Controlled Mode

Session	**Cultural Comparison**	**Variable**	**Statistic**	**p-value**
Session 1	Middle Eastern vs Chinese	Left Wrist Extension	264	0.0044
Session 1	Middle Eastern vs Western	Right Arm Extension	292	0.0335
Session 1	Middle Eastern vs Western	User Position	286	0.0198
Session 1	Chinese vs Western	User Position	260	0.0174
Session 2	Middle Eastern vs Chinese	Torso Tilt Angle	103	0.0246

Results indicate statistically significant differences ($p < 0.05$) in selected body pose variables across cultural groups in the Human-Controlled mode.

while Chinese participants also maintained a greater User Position compared to Western participants ($p = 0.0174$). This suggests that Western participants, consistent with the autonomous mode findings, preferred closer interaction distances regardless of whether the robot was controlled autonomously or by a human. Additionally, Right Arm Extension

was significantly higher in Middle Eastern participants compared to Western participants ($p = 0.0335$), suggesting that Middle Eastern participants used more extended reaching movements during human-controlled interactions, potentially indicating a preference for greater physical engagement with the robot from a slightly farther distance. In Session 2, Torso Tilt Angle was significantly greater in Middle Eastern participants compared to Chinese participants ($p = 0.0246$), suggesting that Middle Eastern participants exhibited more pronounced torso movements while engaging with the human-controlled robot, possibly as a strategy to compensate for the increased distance they maintained.

4.8.7 Mode Preference and Trust by Culture

This subsection explores whether participants' cultural background influenced their preferences and trust toward the robot's control mode—autonomous, human-controlled, or neither—during physical human-robot interaction (pHRI).

Figures 4.15 and 4.16 illustrate participants' preferences and trust ratings across the three cultural groups: Chinese, Middle Eastern, and Western.

Qualitative Overview

Mode Preference As shown in Figure 4.15, Western participants strongly preferred the autonomous mode. Fourteen out of 20 chose autonomy, while only six preferred the human-controlled mode. None selected "neither." This reflects a clear cultural tendency among Western participants to trust technological autonomy.

Middle Eastern participants showed the opposite pattern: 16 out of 20 preferred the human-controlled mode, only three preferred autonomy, and one participant selected "neither." This suggests a cultural disposition toward maintaining human oversight and responsibility.

Chinese participants exhibited more balanced preferences: 12 favored the human-controlled mode, six selected autonomy, and two chose "neither." These results reflect a flexible stance toward robot autonomy, indicating openness to both control paradigms.

Trust in Mode Figure 4.16 presents participants' trust ratings. Again, Western participants predominantly trusted the autonomous system, with 16 out of 20 reporting trust in this condition, compared to four for human control. This trust alignment further supports their preference for autonomous modes.

In contrast, Middle Eastern participants primarily trusted the human-controlled mode (14 out of 20), while six trusted autonomy. This mirrors their preference distribution and reinforces the finding that cultural orientation influences trust.

Chinese participants again displayed a mixed profile: 11 trusted human control, 7 trusted autonomy, and 2 chose "neither." Their trust pattern aligns closely with their mode preferences and reflects moderate flexibility.

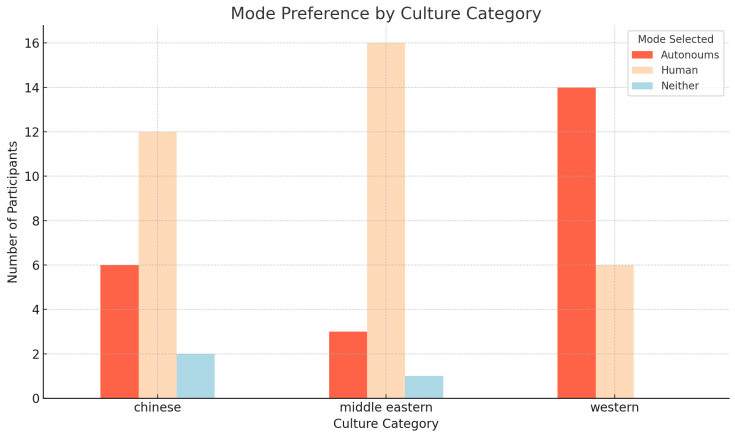


Figure 4.15: Mode preference by cultural category: Number of participants selecting autonomous, human-controlled, or neither mode.

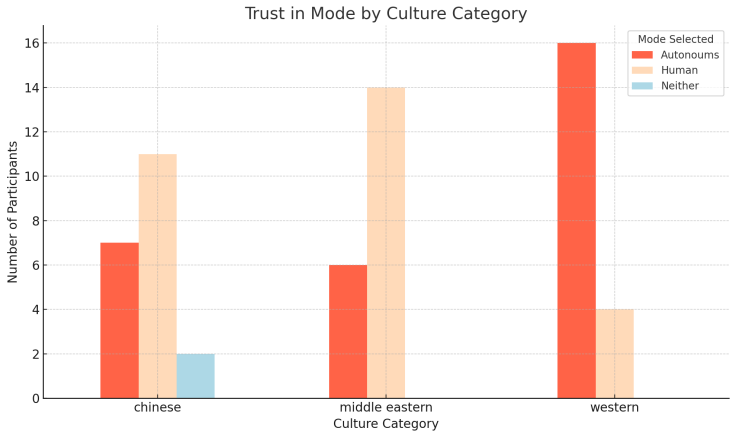


Figure 4.16: Trust in mode by cultural category: Number of participants who expressed trust in each control mode.

Statistical Analysis

To quantify the observed cultural differences, two chi-square tests of independence were conducted.

1. Culture Category × Mode Preference

- Chi-square statistic: $\chi^2(4) = 17.63$
- p-value = 0.00145
- Cramér's $V = 0.38$ (Medium to large effect)

Interpretation: There is a statistically significant association between participants' cultural background and their preferred control mode ($p < 0.05$). The effect size ($V = 0.38$) suggests a medium to large relationship, confirming that cultural orientation influences preference toward autonomy or human control.

2. Culture Category × Trust in Mode

- Chi-square statistic: $\chi^2(4) = 13.62$
- p-value = 0.0086
- Cramér's $V = 0.34$ (Medium to large effect)

Interpretation: A significant association was also found between culture and the mode participants reported trusting most. With a Cramér's V of 0.34, the effect size again indicates a medium to large impact. This result supports the conclusion that trust in autonomous vs. human-controlled mode was affected by culture background.

4.9 Discussion

This section presents the findings from the study investigating the influence of cultural differences on user preferences for robot control modes (autonomous vs. human-controlled) and their reactions to robot errors during pHRI. The study was guided by three primary

research questions (RQ1, RQ2, and RQ3), which explored the interplay between robot control modes, intentional errors, and cultural background. The results provide valuable insights into how these factors influence user perception, acceptance, and comfort during physical interaction with robots..

RQ1: How does the control mode (autonomous vs. human-controlled) influence participants' acceptance and comfort during physical interaction tasks? Subjective ratings such as comfort, safety, and task ease showed no significant differences between autonomous and human-controlled modes across all cultural groups. Participants reported similar experiences regardless of who was controlling the robot. However, preferences expressed in the mode selection task and trust ratings varied by culture.

As shown in Figure 4.15, most Western participants (14 out of 20) preferred the autonomous mode, and 16 out of 20 reported higher trust in it (Figure 4.16). This aligns with LMM results, where selecting the human-controlled mode was negatively associated with both personal (P+) and societal (S+) positive attitudes.

In contrast, Middle Eastern participants strongly favored the human-controlled mode (16 out of 20) and trusted it more (14 out of 20), suggesting a strong preference for human oversight. This aligns with physiological data, where they showed increased heart rate and stood farther away from the robot in the autonomous condition.

Chinese participants showed a more balanced distribution. Although 12 out of 20 preferred the human-controlled mode and 11 trusted it more, their attitudes were more mixed. LMM results showed that selecting the human-controlled mode reduced negative attitudes (P-), but also lowered societal-level positive attitudes (S+). Open-ended feedback supported this trend: many Chinese participants expressed reassurance when a human was in control, even if they were open to robot autonomy in principle.

Overall, these results suggest that while direct subjective ratings did not vary by mode, underlying cultural expectations, trust, and comfort were reflected in preferences, trust data, and behavioral responses. Western participants preferred autonomous systems, Middle Eastern participants preferred human control, and Chinese participants leaned slightly toward human control but remained flexible.

RQ2: Does the introduction of an intentional error in the robot's behavior affect participants' preference for one control mode over the other? The introduction of an error in Task 2 significantly influenced participants' emotional responses. sam scores showed consistent drops in valence and dominance across all cultural groups

and control modes, confirming increased discomfort during the error. These changes were statistically significant and observed in autonomous and human-controlled conditions.

Participants' emotional scores improved in Task 3, suggesting they adapted to the error over time. This recovery was evident across cultural groups and control modes.

Heart rate data showed increased arousal during Task 2 for Middle Eastern and Western participants, confirming the stress-inducing nature of the error. GSR data, while not statistically significant across control modes, showed greater variability during the error task, especially among Middle Eastern participants.

There was no direct shift in mode preference following the error, but subjective and physiological responses suggested different emotional attributions. Participants appeared more forgiving of errors in the human-controlled mode, potentially attributing mistakes to the human operator rather than the robot. In some groups (mainly Middle Eastern and Chinese), this may have strengthened trust in the human-controlled option.

RQ3: How does cultural background moderate the effects of control mode and error introduction on user perception and acceptance of the robot?

Western participants showed a clear preference for autonomous control and trust in autonomous systems. They rated the human-controlled mode less favorably in the LMM analysis and consistently selected and trusted the autonomous mode. Their body posture indicated greater proximity to the robot, suggesting higher physical comfort. However, they were highly sensitive to robot errors, showing sharp drops in sam scores and elevated heart rates following Task 2.

Middle Eastern participants were more cautious overall. They strongly preferred and trusted the human-controlled mode and stood farther from the robot in the autonomous condition, especially after errors. Their physiological data showed consistently higher arousal in autonomous interactions. LMM results revealed stronger societal concerns (S-) and discomfort with automation, especially when errors occurred.

Chinese participants showed mixed attitudes. Mode preference and trust charts revealed a stronger lean toward human-controlled interaction. LMM results showed that selecting the human-controlled mode reduced personal-level negative attitudes (P-), suggesting that participants felt safer or more in control with a human involved. However,

societal positive attitudes (S+) were also lower when this mode was selected, indicating a concern about the broader implications of relying on human control. Their physiological responses were milder than other groups, and their SAM scores showed consistent recovery after errors, suggesting adaptability and emotional resilience.

Cultural differences also influenced movement patterns. Western participants displayed more significant torso changes across tasks, while Chinese participants varied in arm extension. Middle Eastern participants adjusted their position significantly after the error, maintaining greater distance in later tasks.

4.10 Limitations

Several limitations were identified in this study that should be considered when interpreting the findings. Firstly, although intentional error conditions were implemented, the nature and frequency of these errors were not varied, limiting insights into how different types of errors might differently impact user trust and perceptions across cultures. Furthermore, the controlled laboratory-based setting might not entirely reflect naturalistic reactions and behaviors in real-world contexts. Additionally, the interaction duration was relatively brief, which may limit the understanding of how trust and comfort evolve over more extended periods of robot interaction. Moreover, employing only one robotic arm (Sawyer) constrains the generalizability of findings to robots with varying designs or functionalities. Lastly, participants' prior experiences or biases regarding automation and robotic technologies were not explicitly measured or controlled, which could have influenced their responses. Future research could address these limitations by employing diverse robot models, varying error types and frequencies, and extending interaction durations.

Chapter 5

Conclusion

5.1 Summary of Findings

This subsection revisits the research questions (RQI, RQII, RQIII, and RQIV) outlined in Chapter 2, summarizing key results from the two empirical studies presented in Chapters 3 and 4. The findings are concisely tied back to the original thesis aims, providing a clear overview of the results.

RQI: *How do users from different cultural backgrounds behave and feel during physical interactions with robots?*

The first study (Chapter 3) demonstrated that cultural background significantly shapes user behavior and emotions during pHRI, Subjective measures from the General Attitudes Toward Robots Scale GAToRS revealed distinct cultural patterns: Western and Chinese participants reported more positive personal and societal attitudes (higher P+ and S+ scores) toward both robots aligning with previous studies [8, 74], with a stronger preference for the humanoid TALOS. In contrast, Middle Eastern participants exhibited higher negative attitudes (elevated P- and S- scores), particularly during collaborative tasks, suggesting greater discomfort and mistrust.

Behavioral observations further highlighted these differences. Analysis of body pose data indicated that Western and Chinese participants maintained closer proximity to the robots (lower User Position values) and adopted more open, relaxed postures, reflecting greater comfort and engagement (Mann-Whitney U test, $p < 0.05$). Middle Eastern participants, however, stood farther away and displayed more defensive postures, especially with MOVO, indicating heightened caution. Qualitative feedback reinforced these findings: Western participants often described TALOS as "friendly" and "intuitive," while

Middle Eastern participants expressed concerns about the robots' "unpredictability" and "potential for harm." These results demonstrate that cultural background influences not only emotional responses but also physical interaction styles in pHRI, with Western and Chinese participants showing greater ease compared to Middle Eastern participants' wariness.

RQII: *In a collaborative physical task, how does the robot's form factor—such as being humanoid vs. non-humanoid, its size, and limb design—influence the efficiency of the collaboration, the human's trust and comfort, and overall task safety?*

The humanoid robot TALOS was consistently perceived as more approachable and comfortable than the non-humanoid MOVO across all cultural groups, particularly during independent tasks such as pick-and-place. Western and Chinese participants demonstrated higher trust and comfort with TALOS, while Middle Eastern participants maintained cautious attitudes regardless of robot form. Although task efficiency was not explicitly quantified, qualitative feedback and physiological data suggest TALOS's anthropomorphic features positively influenced perceptions of safety and interaction quality, aligning with existing literature on anthropomorphism in HRI [30]. A two-way ANOVA confirmed a significant main effect between the pre- and post-experiment questionnaire and robot type, with an interaction effect ($p < 0.05$).Spearman correlation analysis further revealed a positive relationship between perceived movement predictability and comfort levels ($p < 0.01$), particularly with TALOS, indicating that its form facilitated more intuitive interactions.

RQIII: *How do physiological responses (e.g., HR, GSR) during physical human-robot interactions vary across different cultural backgrounds, and how can these differences help shape robotic systems to suit varied cultural preferences?*

The first study explored how participants from Western, Middle Eastern, and Chinese cultural backgrounds interacted with two robots—the humanoid TALOS and the non-humanoid MOVO—during collaborative tasks like handovers and independent tasks like pick-and-place. Western and Chinese participants reported higher comfort levels with TALOS, supported by physiological data showing lower HR and galvanic skin response GSR levels, indicating reduced stress compared to interactions with MOVO. For example, Western participants' average HR was significantly lower with TALOS ($p < 0.05$), suggesting that its human-like form enhanced perceived safety and trust. In contrast, Middle Eastern participants exhibited greater caution and higher physiological stress with both robots, particularly during collaborative tasks. Their HR and GSR levels were elevated, reflecting discomfort or wariness, possibly tied to cultural attitudes toward technology or personal space. These findings address how robot appearance influences trust and comfort, with TALOS being more favorably received by Western and Chinese participants, while Middle

Eastern participants remained hesitant regardless of form.

The second study investigated preferences for robot control modes (autonomous vs. human-controlled) and responses to intentional robot errors among the same cultural groups, using the Sawyer robot. Western participants strongly preferred autonomous modes, with 14 out of 20 selecting them and 16 out of 20 reporting higher trust, recovering quickly from errors as shown by their Self-Assessment Manikin (SAM) scores and physiological data. Chinese participants displayed mixed preferences, with 12 out of 20 favoring human-controlled modes but showing moderate trust in both; their physiological responses were milder, with quicker recovery post-errors, and GSR dropped faster than Middle Eastern participants. Middle Eastern participants, however, favored human-controlled modes (16 out of 20 preferred, 14 out of 20 trusted) and exhibited prolonged stress after errors in autonomous conditions, with significantly higher HR and GSR levels persisting, HR remained elevated by 8-12 bpm post-error ($p < 0.05$). This underscores clear cultural differences: Western and Chinese participants experienced lower stress and adapted more readily to autonomy and errors. In contrast, Middle Eastern participants' heightened physiological stress highlighted a preference for human oversight and sensitivity to robot mistakes.

RQIV: *How do cultural differences influence user preferences for autonomous versus human-controlled robotic systems in the presence of errors?*

The second study (Chapter 4) revealed that cultural background significantly moderates preferences for robot control modes, particularly in the presence of errors. Western participants preferred and trusted autonomous modes, with statistical analysis indicating significant effects on personal positive attitudes ($p < 0.01$) and societal positive attitudes ($p < 0.05$). In contrast, Middle Eastern participants favored human-controlled modes, with preferences associated with reductions in negative personal attitudes ($p < 0.01$) and negative societal attitudes ($p < 0.05$). A significant interaction effect between culture and control mode was also observed ($p < 0.05$), suggesting heightened sensitivity to errors among Middle Eastern participants.

Chinese participants exhibited mixed preferences and moderate trust levels. Selection of the human-controlled mode was linked to a decrease in personal negative attitudes ($p < 0.05$), alongside a reduction in societal positive attitudes ($p < 0.05$). These findings, supported by qualitative data, underscore the importance of cultural background in shaping control mode preferences and attitudinal responses to robot errors.

Collectively, these findings illustrate that cultural backgrounds profoundly shape user experiences in physical human-robot interactions, with robot appearance and control modes serving as significant moderators.

5.2 Contribution to Knowledge

This subsection outlines the novel insights contributed by this research to the field of HRI, particularly within the context of cultural influences on pHRI. It highlights how this work addresses the research gaps identified in Chapter 2 and discusses its implications for robot design.

Cultural Influences in pHRI: This thesis offers empirical evidence that cultural background significantly shapes user perceptions, trust levels, and physiological responses during physical collaboration with robots—an area less explored than sHRI. By including participants from Western, Middle Eastern, and Chinese backgrounds, the study uncovers distinct interaction patterns, such as the cautious approach of Middle Eastern participants compared to the greater openness to autonomy observed among Western participants.

Robot Appearance: The comparison between humanoid (TALOS) and non-humanoid (MOVO) robots reveals that robot form influences comfort and trust in culturally dependent ways. This finding extends prior research on anthropomorphism into the realm of pHRI, suggesting that humanoid designs may be particularly effective in enhancing acceptance within certain cultural groups.

Control Modes and Errors: The investigation into autonomous versus human-controlled modes—especially under error conditions—demonstrates clear cultural preferences: Western participants favored autonomy, whereas Middle Eastern and some Chinese participants preferred human oversight. Observing that errors are more tolerated in human-controlled modes provides a nuanced understanding of trust dynamics and error resilience across cultures.

Physiological Measures: The integration of heart rate and galvanic skin response data alongside subjective reports offers a more broad view of emotional responses to robot interaction. This methodological contribution highlights the importance of the need for multi-modal, objective assessments in HRI research.

Together, these contributions inform the design of robot systems that are sensitive to cultural differences, aiming to improve user trust, comfort, and interaction effectiveness across diverse application domains such as healthcare and industry.

5.3 Limitations

This section addresses key limitations of the research, particularly those related to participant representation and methodological constraints, and situates them within the broader

context of HRI studies.

5.3.1 Participant Selection and Representation:

Participants in this study self-identified as belonging to Western, Middle Eastern, or Chinese cultural groups, without independent verification or assessment of the depth of their cultural affiliation. This approach may overlook intra-cultural variations and the complexity of cultural identity. All participants were residing and studying in Canada, representing a specific demographic—primarily international students or immigrants—which may not fully reflect the broader populations typically associated with the respective cultural categories. Exposure to Canadian culture could have influenced their responses, limiting the findings' generalizability to their home cultural contexts. Additionally, the sample was self-selecting, as participants volunteered based on interest in the study or in robotics (e.g., via recruitment posters). This introduces potential bias, which is common in HRI research [49], where participants may already possess favorable attitudes toward technology.

5.3.2 Methodological Constraints:

The use of specific robot platforms—TALOS, MOVO, and Sawyer—limits the extent to which findings can be generalized to other robot designs or form factors. For example, responses to larger industrial robots or highly stylized humanoids may differ significantly. The experiments were conducted in a controlled laboratory environment, which does not fully replicate real-world pHRI settings. In practical scenarios, environmental factors and prolonged interactions may shape user responses differently. Furthermore, the relatively small sample sizes per cultural group (e.g., 20 participants per group in Study 2) limit the statistical power to detect subtle or nuanced effects. Finally, the broad categorization into "Western," "Middle Eastern," and "Chinese" oversimplifies the rich diversity within each cultural group. This may mask important individual or regional differences..

5.4 Future Work

Building upon the findings and limitations of this thesis, this subsection outlines directions for future research aimed at advancing the understanding of cultural influences in pHRI and supporting more inclusive robot design.

Future studies should aim to recruit participants within their home countries or those with minimal exposure to Canadian culture to better capture authentic cultural influences. Expanding beyond the Western, Middle Eastern, and Chinese groups to include additional populations—such as South Asian or African participants—could uncover further cross-cultural variations in pHRI.To enhance the generalization of results, future research should explore a broader range of robot embodiments, including larger industrial robots, cartoon-like or stylized designs, and varying levels of autonomy, such as semi-autonomous control modes. This would allow a deeper understanding of how robot form and control influence user perceptions across cultures.

Investigating pHRI in naturalistic settings, such as homes, workplaces, or healthcare environments, over extended periods would provide insights into how trust, comfort, and collaboration evolve with repeated exposure and familiarity. This would move the field closer to ecological validity. Subsequent work should examine the effects of different error types, frequencies, and severities on user perceptions and preferences for control modes. Building on the single-error condition used in Study 2, such research could offer a more nuanced understanding of error tolerance across cultures.

Future development should explore robot systems that adjust their behavior, appearance, or autonomy level in response to users' cultural backgrounds. The insights from this thesis can inform the design of adaptive algorithms that aim to foster more personalized and effective interactions.To improve cultural assessment methods, collaborations with cultural anthropologists, psychologists, or sociologists are recommended. Going beyond self-identification, future research could incorporate validated measures capturing dimensions such as individualism vs. collectivism, uncertainty avoidance, or attitudes toward technology. Standardized methodologies for cross-cultural HRI research are needed to facilitate replicability and cross-study comparison. This includes protocols for participant recruitment, tools for cultural measurement, and frameworks for collecting and analyzing physiological signals.

References

[1] Henny Admoni and Brian Scassellati. Social eye gaze in human-robot interaction: a review. *Journal of Human-Robot Interaction*, 6(1):25–63, 2017.

[2] Muneeb Imtiaz Ahmad, Jasmin Bernotat, Katrin Lohan, and Friederike Eyssel. Trust and cognitive load during human-robot interaction. *arXiv preprint arXiv:1909.05160*, 2019.

[3] Muneeb Imtiaz Ahmad, Jasmin Bernotat, Katrin Lohan, and Friederike Eyssel. Trust and cognitive load during human-robot interaction. *arXiv preprint arXiv:1909.05160*, 2019.

[4] Ferdous Ahmed, A. S. M. Hossain Bari, and Marina L. Gavrilova. Emotion Recognition From Body Movement. *IEEE Access*, 8:11761–11781, 2020.

[5] Abdullah Alzahrani, Simon Robinson, and Muneeb Ahmad. Exploring factors affecting user trust across different human-robot interaction settings and cultures. In *Proceedings of the 10th international conference on human-agent interaction*, pages 123–131, 2022.

[6] Ammar Alzaydi, Kahtan Abedalrhman, Mohd Ismail, and Siti Nurhaliza. Human-robot interaction in saudi arabia's e-mobility transition-a literature review. *Social Science Journal for Advanced Research*, 4:74–96, 2024.

[7] Christoph Bartneck. Godspeed questionnaire series: Translations and usage. In *International handbook of behavioral health assessment*, pages 1–35. Springer, 2023.

[8] Christoph Bartneck, Dana Kulić, Elizabeth Croft, and Susana Zoghbi. Measurement Instruments for the Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of Robots. *International Journal of Social Robotics*, 1(1):71–81, January 2009.

[9] Christoph Bartneck, T. Nomura, T. Kanda, T. Suzuki, and K. Kennsuke. *A cross-cultural study on attitudes towards robots*. Lawrence Erlbaum Associates, 2005. Accepted: 2020-01-21T19:54:31Z.

[10] Christoph Bartneck, Tomohiro Suzuki, Takayuki Kanda, and Tatsuya Nomura. The influence of people's culture and prior experiences with Aibo on their attitude towards robots. *AI & SOCIETY*, 21(1):217–230, January 2007.

[11] Valentin Bazarevsky, Ivan Grishchenko, Karthik Raveendran, Tyler Zhu, Fan Zhang, and Matthias Grundmann. BlazePose: On-device Real-time Body Pose tracking. Technical Report arXiv:2006.10204, arXiv, 2020.

[12] Jenay M. Beer, Cory-Ann Smarr, Tiffany L. Chen, Akanksha Prakash, Tracy L. Mitzner, Charles C. Kemp, and Wendy A. Rogers. The domesticated robot: design guidelines for assisting older adults to age in place. In *International conference on Human-Robot Interaction*, page 335. ACM Press, 2012.

[13] Margaret M. Bradley and Peter J. Lang. Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(1):49–59, March 1994.

[14] C. Breazeal. Social interactions in HRI: the robot view. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 34(2):181–186, May 2004. Conference Name: IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews).

[15] Elizabeth Broadbent. Interactions with robots: The truths we reveal about ourselves. *Annual Review of Psychology*, 68:627–652, 2017. Place: US Publisher: Annual Reviews.

[16] Elizabeth Broadbent, Rebecca Stafford, and Bruce MacDonald. Acceptance of healthcare robots for the older population: Review and future directions. *International journal of social robotics*, 1:319–330, 2009.

[17] Maya Cakmak, Siddhartha S Srinivasa, Min Kyung Lee, Sara Kiesler, and Jodi Forlizzi. Using spatial and temporal contrast for fluent robot-human hand-overs. In *Proceedings of the 6th international conference on Human-robot interaction*, pages 489–496, 2011.

[18] Noah Castelo and Miklos Sarvary. Cross-cultural differences in comfort with human-like robots. *International Journal of Social Robotics*, 14(8):1865–1873, 2022.

[19] Elizabeth Cha, Anca D. Dragan, and Siddhartha S. Srinivasa. Effects of robot capability on user acceptance. In *International Conference on Human-Robot Interaction*, pages 97–98, Tokyo, Japan, March 2013. IEEE.

[20] Thierry Chaminade, Jessica Hodgins, and Mitsuo Kawato. Anthropomorphism influences perception of computer-animated characters' actions. *Social Cognitive and Affective Neuroscience*, 2(3):206–216, September 2007.

[21] Carlo Chiorri, Fabrizio Bracco, Tommaso Piccinno, Cinzia Modafferi, and Valeria Battini. Psychometric properties of a revised version of the ten item personality inventory. *European Journal of Psychological Assessment*, 2015.

[22] Sachin Chitta, Ioan Sucan, and Steve Cousins. Moveit! [ros topics]. *IEEE Robotics & Automation Magazine*, 19(1):18–19, 2012.

[23] Lara Christoforakos, Alessio Gallucci, Tinatini Surmava-Große, Daniel Ullrich, and Sarah Diefenbach. Can Robots Earn Our Trust the Same Way Humans Do? A Systematic Exploration of Competence, Warmth, and Anthropomorphism as Determinants of Trust Development in HRI. *Frontiers in Robotics and AI*, 8, 2021.

[24] Victoria Clarke and Virginia Braun. Thematic analysis. *The journal of positive psychology*, 12(3):297–298, 2017.

[25] Jacob Cohen. *Statistical power analysis for the behavioral sciences*. Routledge, 1988.

[26] Mark Coulson. Attributing Emotion to Static Body Postures: Recognition Accuracy, Confusions, and Viewpoint Dependence. *Journal of Nonverbal Behavior*, 28(2):117–139, 2004.

[27] K. Dautenhahn and I. Werry. A quantitative technique for analysing robot-human interactions. In *International Conference on Intelligent Robots and System*, volume 2, pages 1132–1138. IEEE, 2002.

[28] Kerry Dhakal. Nvivo. *Journal of the Medical Library Association: JMLA*, 110(2):270, 2022.

[29] Anca D. Dragan, Kenton C.T. Lee, and Siddhartha S. Srinivasa. Legibility and predictability of robot motion. In *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 301–308, 2013.

[30] Brian R Duffy. Anthropomorphism and the social robot. *Robotics and autonomous systems*, 42(3-4):177–190, 2003.

[31] Aaron Edelmann, Stefan Stümper, and Tibor Petzoldt. Cross-cultural differences in the acceptance of decisions of automated vehicles. *Applied ergonomics*, 92:103346, 2021.

[32] Ghadeer Eresha, Markus Häring, Birgit Endrass, Elisabeth André, and Mohammad Obaid. Investigating the influence of culture on proxemic behaviors for humanoid robots. In *2013 IEEE Ro-Man*, pages 430–435. IEEE, 2013.

[33] Vanessa Evers, Heidy C. Maldonado, Talia L. Brodecki, and Pamela J. Hinds. Relational vs. group self-construal: untangling the role of national culture in HRI. In *Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction*, pages 255–262, Amsterdam The Netherlands, March 2008. ACM.

[34] Friederike Eyssel and Dieta Kuchenbrandt. Social categorization of social robots: Anthropomorphism as a function of robot group membership: *Social categorization and social robots. British Journal of Social Psychology*, 51(4):724–731, December 2012.

[35] Rebecca Flook, Anas Shrinah, Luc Wijnen, Kerstin Eder, Chris Melhuish, and Séverin Lemaignan. On the impact of different types of errors on trust in human-robot interaction: Are laboratory-based hri experiments trustworthy? *Interaction Studies*, 20(3):455–486, 2019.

[36] Norina Gasteiger, Mehdi Hellou, and Ho Seok Ahn. Factors for Personalization and Localization to Optimize Human–Robot Interaction: A Literature Review. *International Journal of Social Robotics*, 15(4):689–701, April 2023.

[37] Riccardo Gervasi, Luca Mastrogiacomo, and Fiorenzo Franceschini. An experimental focus on learning effect and interaction quality in human–robot collaboration. *Production Engineering*, 17(3):355–380, 2023.

[38] J. Goetz, S. Kiesler, and A. Powers. Matching robot appearance and behavior to tasks to improve human-robot cooperation. In *International Workshop on Robot and Human Interactive Communication*, pages 55–60, Millbrae, CA, USA, 2003. IEEE.

[39] Steven Goodman. A dirty dozen: twelve p-value misconceptions. *Seminars in hematology*, 45(3):135–140, 2008.

[40] Fu Guo, Mingming Li, Jiahao Chen, and Vincent G Duffy. Evaluating usersâ preference for the appearance of humanoid robots via event-related potentials and spectral perturbations. *Behaviour & Information Technology*, 41(7):1381–1397, 2022.

[41] Kasper Hald, Matthias Rehmn, and Thomas B Moeslund. Human-robot trust assessment using motion tracking & galvanic skin response. In *International Conference on Intelligent Robots and Systems*, pages 6282–6287. IEEE, 2020.

[42] Peter A. Hancock, Deborah R. Billings, Kristin E. Schaefer, Jessie Y. C. Chen, Ewart J. De Visser, and Raja Parasuraman. A Meta-Analysis of Factors Affecting Trust in Human-Robot Interaction. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 53(5):517–527, October 2011.

[43] James M. Hanson and Kenneth E. Sinclair. Social constructivist teaching methods in Australian universities – reported uptake and perceived learning effects: a survey of lecturers. *Higher Education Research & Development*, 27(3):169–186, September 2008.

[44] Kerstin Sophie Haring, C^|^eacute;line Mougenot, Fuminori Ono, and Katsumi Watanabe. Cultural Differences in Perception and Attitude towards Robots. *International Journal of Affective Engineering*, 13(3):149–157, 2014.

[45] Kerstin Sophie Haring, David Silvera-Tawil, Katsumi Watanabe, and Mari Velonaki. The Influence of Robot Appearance and Interactive Ability in HRI: A Cross-Cultural Study. In Arvin Agah, John-John Cabibihan, Ayanna M. Howard, Miguel A. Salichs, and Hongsheng He, editors, *Social Robotics*, Lecture Notes in Computer Science, pages 392–401, Cham, 2016. Springer International Publishing.

[46] Marcel Heerink, Ben Krose, Vanessa Evers, and Bob Wielinga. Measuring acceptance of an assistive social robot: a suggested toolkit. In *International Symposium on Robot and Human Interactive Communication*, pages 528–533. IEEE, September 2009.

[47] Pamela J Hinds, Teresa L Roberts, and Hank Jones. Whose job is it anyway? a study of human-robot interaction in a collaborative task. *Human–Computer Interaction*, 19(1-2):151–181, 2004.

[48] Guy Hoffman and Wendy Ju. Designing robots with movement in mind. *Journal of Human-Robot Interaction*, 3(1):91–122, 2014.

[49] Guy Hoffman and Xuan Zhao. A Primer for Conducting Experiments in Human–Robot Interaction. *ACM Transactions on Human-Robot Interaction*, 10(1):1–31, February 2021.

[50] Geert Hofstede. *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations*. Sage publications, 2001.

[51] Sarah Hopko, Jingkun Wang, and Ranjana Mehta. Human Factors Considerations and Metrics in Shared Space Human-Robot Collaboration: A Systematic Review. *Frontiers in Robotics and AI*, 9:799522, February 2022.

[52] Chi-Ping Hsiung and Erin Chiou. Attribution biases and trust development in physical human-machine coordination: Blaming yourself, your partner or an unexpected event. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 63, pages 211–211. SAGE Publications Sage CA: Los Angeles, CA, 2019.

[53] Yue Hu, Naoko Abe, Mehdi Benallegue, Natsuki Yamanobe, Gentiane Venture, and Eiichi Yoshida. Toward Active Physical Human–Robot Interaction: Quantifying the Human State During Interactions. *IEEE Transactions on Human-Machine Systems*, pages 1–12, 2022.

[54] Saki Kato, Natsuki Yamanobe, Gentiane Venture, Eiichi Yoshida, and Gowrishankar Ganesh. The where of handovers by humans: Effect of partner characteristics, distance and visual feedback. *PLOS ONE*, 14(6):e0217129, June 2019. Publisher: Public Library of Science.

[55] Kheng Lee Koay, Dag Sverre Syrdal, Michael L. Walters, and Kerstin Dautenhahn. Five Weeks in the Robot House – Exploratory Human-Robot Interaction Trials in a Domestic Setting. In *International Conferences on Advances in Computer-Human Interactions*, pages 219–226. IEEE, February 2009.

[56] Manuel Koller. robustlmm: An R package for robust estimation of linear mixed-effects models. *Journal of Statistical Software*, 75(6):1–24, 2016.

[57] Oliver Korn, Neziha Akalin, and Ruben Gouveia. Understanding cultural preferences for social robots: a study in german and arab communities. *ACM Transactions on Human-Robot Interaction (THRI)*, 10(2):1–19, 2021.

[58] Mika Koverola, Anton Kunnari, Jukka Sundvall, and Michael Laakasuo. General attitudes towards robots scale (GAToRS): A new instrument for social surveys. *International Journal of Social Robotics*, 14(7):1559–1581, 2022.

[59] Mika Koverola, Anton Kunnari, Jukka Sundvall, and Michael Laakasuo. General attitudes towards robots scale (gators): A new instrument for social surveys. *International Journal of Social Robotics*, 14(7):1559–1581, 2022.

[60] Sabine Theresia Köszegi and Markus Vincze. *Trust in Robots*. TU Wien Academic Press, December 2022.

[61] Hee Rin Lee, JaYoung Sung, Selma Sabanovic, and Joenghye Han. Cultural design of domestic robots: A study of user expectations in Korea and the United States. In *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*, pages 803–808, Paris, France, September 2012. IEEE.

[62] Dingjun Li, P. L. Patrick Rau, and Ye Li. A Cross-cultural Study: Effect of Robot Appearance and Task. *International Journal of Social Robotics*, 2(2):175–186, June 2010.

[63] Shufei Li, Yingchao You, Pai Zheng, Xi Vincent Wang, and Lihui Wang. Mutual-cognition for proactive human–robot collaboration: A mixed reality-enabled visual reasoning-based method. *IISE Transactions*, pages 1–13, 2024.

[64] Velvetina Lim, Maki Rooksby, and Emily S. Cross. Social Robots on a Global Stage: Establishing a Role for Culture During Human–Robot Interaction. *International Journal of Social Robotics*, 13(6):1307–1333, September 2021.

[65] Karl F. MacDorman, Sandosh K. Vasudevan, and Chin-Chang Ho. Does Japan really have robot mania? Comparing attitudes by implicit and explicit measures. *AI & SOCIETY*, 23(4):485–510, July 2009.

[66] Henry B Mann and Donald R Whitney. On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics*, pages 50–60, 1947.

[67] Serena Marchesi, Abdulaziz Abubshait, Kyveli Kompatsiari, Yan Wu, and Agnieszka Wykowska. Cultural differences in joint attention and engagement in mutual gaze with a robot face. *Scientific Reports*, 13(1):11689, 2023.

[68] Chinmaya Mishra, Gabriel Skantze, Peter Hagoort, and Rinus Verdonschot. Perception of emotions in human and robot faces: Is the eye region enough? *arXiv preprint arXiv:2410.14337*, 2024.

[69] AJung Moon, Chris AC Parker, Elizabeth A Croft, and HF Machiel Van der Loos. Did you see it hesitate?-empirically grounded design of hesitation trajectories for collaborative robots. In *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1994–1999. IEEE, 2011.

[70] Masahiro Mori, Karl MacDorman, and Norri Kageki. The Uncanny Valley [From the Field]. *IEEE Robotics & Automation Magazine*, 19(2):98–100, June 2012.

[71] Jonathan Mumm and Bilge Mutlu. Human-robot proxemics: physical and psychological distancing in human-robot interaction. In *International conference on Human-robot interaction*, pages 331–338, 2011.

[72] Shinichi Nakagawa and Holger Schielzeth. A general and simple method for obtaining r2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 4(2):133–142, 2013.

[73] Stanislava Naneva, Marina Sarda Gou, Thomas L Webb, and Tony J Prescott. A systematic review of attitudes, anxiety, acceptance, and trust towards social robots. *International Journal of Social Robotics*, 12(6):1179–1201, 2020.

[74] Richard Nisbett. *The Geography of Thought: How Asians and Westerners Think Differently... and*. Simon and Schuster, 2010.

[75] Tatsuya Nomura, Takayuki Kanda, Tomohiro Suzuki, and Kensuke Kato. Prediction of Human Behavior in Human–Robot Interaction Using Psychological Scales for Anxiety and Negative Attitudes Toward Robots. *IEEE Transactions on Robotics*, 24(2):442–451, April 2008.

[76] Tatsuya Nomura, Tomohiro Suzuki, Takayuki Kanda, and Kensuke Kato. Measurement of negative attitudes toward robots. *Interaction Studies. Social Behaviour and Communication in Biological and Artificial Systems*, 7(3):437–454, November 2006.

[77] Mark L. Ornelas, Gary B. Smith, and Masoumeh Mansouri. Redefining culture in cultural robotics. *AI & SOCIETY*, 38(2):777–788, April 2023.

[78] Chris Papadopoulos, Tetiana Hill, Linda Battistuzzi, Nina Castro, Abiha Nigath, Gurch Randhawa, Len Merton, Sanjeev Kanoria, Hiroko Kamide, Nak-Young Chong, David Hewson, Rosemary Davidson, and Antonio Sgorbissa. The CARESSES study protocol: testing and evaluating culturally competent socially assistive robots among older adults residing in long term care homes through a controlled experimental trial. *Archives of Public Health*, 78(1):26, March 2020.

[79] Irena Papadopoulos and Christina Koulouglioti. The Influence of Culture on Attitudes Towards Humanoid and Animal-like Robots: An Integrative Review. *Journal of Nursing Scholarship: An Official Publication of Sigma Theta Tau International Honor Society of Nursing*, 50(6):653–665, November 2018.

[80] Irena Papadopoulos, Christina Koulouglioti, and Sheila Ali. Views of nurses and other health and social care workers on the use of assistive humanoid

and animal-like robots in health and social care: a scoping review. *Contemporary Nurse*, 54(4-5):425–442, July 2018. Publisher: Routledge _eprint: https://doi.org/10.1080/10376178.2018.1519374.

[81] Joanna Piasek and Katarzyna Wieczorowska-Tobis. Acceptance and Long-Term Use of a Social Robot by Elderly Users in a Domestic Environment. In *International Conference on Human System Interaction*, pages 478–482, Gdansk, 2018. IEEE.

[82] Joel Pinney, Fiona Carroll, and Paul Newbury. Human-robot interaction: the impact of robotic aesthetics on anticipated human trust. *PeerJ Computer Science*, 8:e837, 2022.

[83] Morgan Quigley, Ken Conley, Brian P Gerkey, Josh Faust, Tully Foote, Jeremy Leibs, Rob Wheeler, and Andrew Y Ng. Ros: an open-source robot operating system. In *ICRA workshop on open source software*, volume 3, page 5. Kobe, Japan, 2009.

[84] Vineeth Rajamohan, Connor Scully-Allison, Sergiu Dascalu, and David Feil-Seifer. Factors Influencing The Human Preferred Interaction Distance. In *International Conference on Robot and Human Interactive Communication*, pages 1–7, New Delhi, India, October 2019. IEEE.

[85] PL Patrick Rau, Ye Li, and Dingjun Li. Effects of communication style and culture on ability to accept recommendations from robots. *Computers in Human Behavior*, 25(2):587–595, 2009.

[86] Laurel D. Riek. Healthcare Robotics. *Communications of the ACM*, 60(11):68–78, October 2017. arXiv:1704.03931 [cs].

[87] Paul Robinette, Wenchen Li, Robert Allen, Ayanna M Howard, and Alan R Wagner. Overtrust of robots in emergency evacuation scenarios. In *2016 11th ACM/IEEE international conference on human-robot interaction (HRI)*, pages 101–108. IEEE, 2016.

[88] Michael J. Roszkowski and Margot Soven. Shifting gears: consequences of including two negatively worded items in the middle of a positively worded questionnaire. *Assessment & Evaluation in Higher Education*, 35(1):113–130, January 2010.

[89] Maha Salem, Gabriella Lakatos, Farshid Amirabdollahian, and Kerstin Dautenhahn. Would You Trust a (Faulty) Robot? Effects of Error, Task Type and Personality on Human-Robot Cooperation and Trust. In *Proceedings of the Tenth Annual*

ACM/IEEE International Conference on Human-Robot Interaction, HRI '15, pages 141–148, New York, NY, USA, March 2015. Association for Computing Machinery.

[90] Shane Saunderson and Goldie Nejat. How robots influence humans: A survey of nonverbal communication in social human–robot interaction. *International Journal of Social Robotics*, 11(4):575–608, 2019.

[91] Celal Savur and Ferat Sahin. Survey on physiological computing in human–robot collaboration. *Machines*, 11(5):536, 2023.

[92] Suleman Shahid, Emiel Krahmer, and Marc Swerts. Child–robot interaction across cultures: How does playing a game with a social robot compare to playing a game alone or with a friend? *Computers in Human Behavior*, 40:86–100, 2014.

[93] Ponkoj Chandra Shill, Md Azizul Hakim, Muhammad Jahanzeb Khan, and Bashira Akter Anima. Human reactions to incorrect answers from robots. *arXiv preprint arXiv:2403.14293*, 2024.

[94] Charles Spearman. The proof and measurement of association between two things. *The American journal of psychology*, 15(1):72–101, 1904.

[95] Steven J Stroessner and Jonathan Benitez. The social perception of humanoid and non-humanoid robots: Effects of gendered and machinelike features. *International Journal of Social Robotics*, 11:305–315, 2019.

[96] Jennifer G. Tichon, Guy Wallis, Stephan Riek, and Timothy Mavin. Physiological measurement of anxiety to evaluate performance in simulation training. *Cognition, Technology & Work*, 16(2):203–210, May 2014.

[97] Jennifer G Tichon and Guy M Wallis. Stress training and simulator complexity: why sometimes more is less. *Behaviour & Information Technology*, 29(5):459–466, 2010.

[98] Harry C Triandis. Individualism-collectivism and personality. *Journal of personality*, 69(6):907–924, 2001.

[99] Gabriele Trovato, Massimiliano Zecca, Salvatore Sessa, Lorenzo Jamone, Jaap Ham, Kenji Hashimoto, and Atsuo Takanishi. Cross-cultural study on human-robot greeting interaction: acceptance and discomfort by Egyptians and Japanese. *Paladyn, Journal of Behavioral Robotics*, 4(2):83–93, December 2013. Publisher: De Gruyter Open Access.

[100] Marieke van Dooren, J.J.G. (Gert-Jan) de Vries, and Joris H. Janssen. Emotional sweating across the body: Comparing 16 different skin conductance measurement locations. *Physiology & Behavior*, 106(2):298–304, 2012.

[101] Fabio Vannucci, Alessandra Sciutti, Hagen Lehman, Giulio Sandini, Yukie Nagai, and Francesco Rea. Cultural differences in speed adaptation in human-robot interaction tasks. *Paladyn, Journal of Behavioral Robotics*, 10(1):256–266, 2019.

[102] Lorenzo Vianello, Luigi Penco, Waldez Gomes, Yang You, Salvatore Maria Anzalone, Pauline Maurice, Vincent Thomas, and Serena Ivaldi. Human-humanoid interaction and cooperation: a review. *Current Robotics Reports*, 2(4):441–454, 2021.

[103] Lin Wang, Pei-Luen Patrick Rau, Vanessa Evers, Benjamin Krisper Robinson, and Pamela Hinds. When in Rome: The role of culture & context in adherence to robot recommendations. In *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 359–366, March 2010. ISSN: 2167-2148.

[104] Ronald L Wasserstein and Nicole A Lazar. The asa's statement on p-values: context, process, and purpose. *The American Statistician*, 70(2):129–133, 2016.

[105] Ronald L Wasserstein, Allen L Schirm, and Nicole A Lazar. Moving to a world beyond "p¡ 0.05". *The American Statistician*, 73:1–19, 2019.

[106] Rui Li Yunyi Jia Weitian Wang, Yi Chen. Learning and comfort in human–robot interaction: A review. *Applied Sciences*, 9(23):5152, Nov 2019.

[107] Frank Wilcoxon. Individual comparisons by ranking methods. *Biometrics bulletin*, 1(6):80–83, 1945.

[108] Kai Chi Yam, Tiffany Tan, Joshua Conrad Jackson, Azim Shariff, and Kurt Gray. Cultural differences in people's reactions and applications of robots, algorithms, and artificial intelligence. *Management and Organization Review*, 19(5):859–875, 2023.

[109] Masaki Yuki, William W Maddux, and Takahiko Masuda. Are the windows to the soul the same in the east and west? cultural differences in using the eyes and mouth as cues to recognize emotions in japan and the united states. *Journal of Experimental Social Psychology*, 43(2):303–311, 2007.

[110] Roberto Zangróniz, Arturo Martínez-Rodrigo, María T López, José Manuel Pastor, and Antonio Fernández-Caballero. Estimation of mental distress from photoplethysmography. *Applied Sciences*, 8(1):69, 2018.

[111] Roberto Zangróniz, Arturo Martínez-Rodrigo, María López, José Pastor, and Antonio Fernández-Caballero. Estimation of Mental Distress from Photoplethysmography. *Applied Sciences*, 8(1):69, January 2018.

[112] Pamely Zantou and David Vernon. Culturally sensitive human-robot interaction: A case study with the pepper humanoid robot. In *2023 IEEE AFRICON*, pages 1–6. IEEE, 2023.

[113] Mengxi Zhang. Older people's attitudes towards emerging technologies: A systematic literature review. *Public Understanding of Science*, 32(8):948–968, 2023. PMID: 37204075.

APPENDICES

Appendix A

Questionnaires

A.1 Screening Questionnaire for Study 1 and 2

Table A.1: Screening Questionnaire for Both Studies

Field	**Description**
Age	State your age.
Cultural Identity	Select from Western, Middle Eastern, Chinese, or specify another.
Cultural Familiarity	Rate your familiarity with cultural traditions.

A.2 Demographic Questions for study 1 and 2

Table A.2: Demographic Information Collected

Category	Statement
Personal Background	
	State your age.
	State your height (if not sure, ask the Research Assistant).
	State your Gender
	State the country of your origin.
	Current or former profession/occupation.
	Select the total number of months/years you have lived in Canada (If born here, select "Born in Canada").
Biological Info	
	Indicate whether you are right-handed or left-handed.
Technology and Lifestyle	
	Do you have a pet?
	Specify which technology items you own and use (e.g., computer, tablet, smartphone, smart home devices, robotic assistants, etc.).

A.3 Pre-Experiment Questionnaires for Study 1

Table A.3: Pre-Interaction Comfort and Expectation Survey for Study 1

Section	Statement
Expectations from Robot	
	I expect the robot to efficiently assist with household tasks.
	I believe the robot will be easy to use.
	I feel confident that interacting with the robot will be safe.
Familiarity with Robotics Arms	
	My initial comfort level with the idea of interacting with robots.
Familiarity with Humanoid Robots	
	How familiar are you with Humanoid Robots?
Familiarity with new Technology	
	How familiar are you with new technologies such as AI, smart homes, and robotic assistants?

A.4 Task Related Questionnaires for Study 1

Table A.4: Task Question for Study 1

Section	Statement
Task 1: Pick-and-Place	
	I felt comfortable performing the task alongside the robot.
	I found it easy to interact with the robot during the task.
	I understood the robot's actions and intentions during the task.
Task 2: Handover	
	I felt comfortable performing the cooperative task with the robot.
	The interaction with the robot was easy to manage during the cooperative task.
	I clearly understood the robot's actions and intentions during the cooperative task.

A.5 Open Questions for Study 1

1. Which robot among the two would you prefer having at home? Why?
2. Which robot would you the least want to have home with you? Why?
3. Were there any specific physical features of the robots that you disliked or made you uncomfortable? If so, which one(s)?
4. From the robot(s) you choose, if you have an older adult in your home environment, will this robot be helpful? If yes or no, please explain why.
5. If you had the ideal robotic assistant to work with you to help you with chores at home (so working side by side), describe what they would look like. We want to focus on how it looks and not what it can do.
6. In your opinion, how significantly does your cultural background influence your daily decisions and interactions?
    - Significantly
    - Moderately
    - Slightly
    - Not at all
7. How do you think your interaction with robots has been influenced by your cultural background?

A.6 Pre-Experiment and Familiarity Questions from Study 2

Table A.5: Familiarity and Expectations

Category	Statement
Cultural and Technological Familiarity	
	How familiar are you with your cultural background?
	How familiar are you with robotic arms/manipulators?
	How familiar are you with new technologies such as AI, smart homes, and robotic assistants?
Robot Interaction Expectations	
	I expect the robot to efficiently assist with the tasks.
	I believe the robot will be easy to use.
	I feel confident that interacting with the robot will be safe.
	My initial comfort level with the idea of interacting with the robot.

A.7 Open Questions for Study 2

- If you were to choose one mode for completing the task in the future, which one would you prefer to work with? Why? (Max 150 words)
- How would you describe your experience interacting with the robot in both modes? (Max 150 words)
- Which mode did you trust more to guide you accurately? What influenced this trust? (Max 150 words)
- How much control did you feel you had over your actions during each mode? Why? (Max 150 words)
- Were there any moments during the task where you felt unsure or had to adjust your approach? What caused those moments? (Max 150 words)
- Do you think your personal experiences and/or cultural background have an effect on which mode you chose to interact with? Did it influence how you perceived and interacted with the robot?

Appendix B

Ethics Approvals

UNIVERSITY OF WATERLOO

Notification of Ethics Clearance to Conduct Research with Human Participants

---

Principal Investigator: Yue Hu

Student investigator: Amr Hamdi

File #: 46088

Title: Evaluating collaborative robot arms for healthcare across different culture groups

---

The Human Research Ethics Board is pleased to inform you this study has been reviewed and given ethics clearance.

Initial Approval Date: 10/01/24 (m/d/y)

University of Waterloo Research Ethics Boards are composed in accordance with, and carry out their functions and operate in a manner consistent with, the institution's guidelines for research with human participants, the Tri-Council Policy Statement for the Ethical Conduct for Research Involving Humans (TCPS2 2022), the Ontario Personal Health Information Protection Act (PHIPA), and all laws and regulations of the province of Ontario (as applicable). Additionally, CREB operates in a manner consistent with the International Conference for Harmonization of Technical Requirements for Pharmaceuticals for Human Use (ICH) Guidance E6(R2): Good Clinical Practice , the International Organization for Standardization of Good Clinical Practices (GCP) as set out by ISO 14155 - Clinical investigation of medical devices for human subjects, Part C, Division 5 of the Food and Drug Regulations, Part 4 of the Natural Health Products Regulations, Part 3 of the Medical Devices Regulations. Both Boards are registered with the U.S. Department of Health and Human Services under the Federal Wide Assurance, FWA00021410, and IRB registration number IRB00002419 (HREB) and IRB00007409 (CREB).

Expiry Date: 10/02/25 (m/d/y)

Multi-year research must be renewed at least once every 12 months unless a more frequent review has otherwise been specified. Studies will only be renewed if the renewal report is received and approved before the expiry date. Failure to submit renewal reports will result in the investigators being notified ethics clearance has been suspended and Research Finance being notified the ethics clearance is no longer valid.

Level of review: Delegated Review

Signed on behalf of the Human Research Ethics Board



UNIVERSITY OF WATERLOO

Notification of Ethics Clearance to Conduct Research with Human Participants

---

Principal Investigator: Yue Hu (Mechanical and Mechatronics Engineering)

Principal Investigator: Wael Suleiman (University of Sherbrooke)

Co-Investigator: Christopher Yee Wong (Mechanical and Mechatronics Engineering)

Student investigator: Julie Pivin (Mechanical and Mechatronics Engineering)

Co-Investigator: Christopher Yee Wong (University of Sherbrooke)

Student investigator: Julie Pivin (University of Sherbrooke)

File #: 44264

Title: Influence of a robot's form on people during human-robot collaborative tasks

---

The Human Research Ethics Board is pleased to inform you this study has been reviewed and given ethics clearance.

Initial Approval Date: 06/10/22 (m/d/y)

University of Waterloo Research Ethics Boards are composed in accordance with, and carry out their functions and operate in a manner consistent with, the institution's guidelines for research with human participants, the Tri-Council Policy Statement for the Ethical Conduct for Research Involving Humans (TCPS, 2nd edition), International Conference on Harmonization: Good Clinical Practice (ICH-GCP), the Ontario Personal Health Information Protection Act (PHIPA), the applicable laws and regulations of the province of Ontario. Both Boards are registered with the U.S. Department of Health and Human Services under the Federal Wide Assurance, FWA00021410, and IRB registration number IRB00002419 (HREB) and IRB00007409 (CREB).

This study is to be conducted in accordance with the submitted application and the most recently approved versions of all supporting materials.

Expiry Date: 06/11/23 (m/d/y)

Multi-year research must be renewed at least once every 12 months unless a more frequent review has otherwise been specified. Studies will only be renewed if the renewal report is received and approved before the expiry date. Failure to submit renewal reports will result in the investigators being notified ethics clearance has been suspended and Research Finance being notified the ethics clearance is no longer valid.

Level of review: Delegated Review

Signed on behalf of the Human Research Ethics Board



Vanessa Buote, Ethics Advisor, vbuote@uwaterloo.ca, 519-888-4567, ext. 40321

This above named study is to be conducted in accordance with the submitted application and the most recently approved versions of all supporting materials.