

Development of a Social Robot for the Study of Robot Roles in Robot-Assisted Language Learning

by

Victoria (Ya-Ting) Yang

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, 2024

© Victoria (Ya-Ting) Yang 2024

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

The objective of this thesis encompasses both the development of a novel social robot, named "Maple," and the examination of its effectiveness in various roles within the context of education and Human-Robot Interaction (HRI). A study was conducted to compare the efficacy of two common roles that social robots can assume in HRI research concerning learning: as a tutor and as a peer-learner. Individual participants, including university students and staff, engage in a vocabulary learning task in Esperanto with Maple assuming either a tutor or a peer-learner role. The research findings indicate participants' positive attitudes and perceptions towards both roles enacted by Maple. Usability assessments reveal high scores across both conditions, with participants reporting predominantly positive user experiences. While no significant difference in learning outcomes is observed between the two robot roles, participants exhibit a 30% increase in quiz completion speed when interacting with Maple in the peer-learner role, suggesting the peer-learner robot promotes and fosters meaningful learning. Thematic analysis of participants' responses underscores their preferences for increased interaction with Maple as a peer-learner, acknowledgment of Maple's capacity for companionship and interactivity, and envisioning practical applications of such learning systems in real-life settings. The insights obtained from this study not only contribute to understanding the effectiveness of social robots in educational settings but also offer valuable directions for future research and technology development in Human-Robot Interaction and Robot-Assisted Language Learning.

Acknowledgements

I extend my sincere appreciation to my supervisors, Professor Yue Hu and Professor Kerstin Dautenhahn, for their advice and guidance throughout the completion of this thesis. Furthermore, I would also like to express my gratitude to my thesis review committee members, Professor Cayley MacArthur and Professor Sanjeev Bedi, for their time and valuable feedback. In addition, I extend my gratitude to my colleagues for their support and camaraderie. I am also thankful to the Waterloo community for providing such a conducive environment for academic growth and collaboration.

Dedication

This is dedicated to my family and friends, who taught me the importance of diligently pursuing my passions and aspirations in life.

Table of Contents

Author’s Declaration	ii
Abstract	iii
Acknowledgements	iv
Dedication	v
List of Figures	x
List of Tables	xii
1 Introduction	1
1.1 Motivation	1
1.1.1 Immigration in Canada	1
1.1.2 Quality Education, Digital Technologies, and Robots	2
1.1.3 In-field Investigation with Local Immigration Services	3
1.2 Research Goal and Scope	4
1.3 Thesis Overview	5
2 Background: Literature Review	7
2.1 HRI and Social Robotics	7

2.1.1	Overview of HRI	7
2.1.2	Overview of Social Robotics	8
2.1.3	Applications of Social Robots	9
2.2	Social Robots in Education	10
2.2.1	Robot Roles in Education	10
2.2.2	Robot-Assisted Language Learning	12
3	Research Question	15
4	Robot Platform Development	17
4.1	Robot Selection Process	17
4.1.1	Robot Comparison and Selection	17
4.2	About the Maple Robot	22
4.3	Robot Framework	22
4.3.1	Robot System Architecture	22
4.3.2	Robot Actuator Control API Architecture	24
4.4	Web Applications for Robot Face, Content Display and User Interface . . .	25
4.5	Documentation for Maple Robot	25
5	User Study: Language Crash Course with a Robot	26
5.1	Hypotheses	26
5.2	Methodology and Design	27
5.2.1	Study Recruitment	27
5.2.2	Experiment Procedure, Design and Materials	27
5.2.3	Robot Behaviour Design	30
5.2.4	Learning Activity - Esperanto Lesson	33
5.3	Measures	38
5.3.1	Pre-experiment measures: Participant Demographics	39

5.3.2	Post-experiment measures: Participant Experience and Performance Assessment	39
5.3.3	Summary of Measures Used	39
5.3.4	Details of Measures Used	40
5.4	Data Analysis Methods	44
5.4.1	Statistical Analysis	45
5.4.2	Thematic Analysis	45
5.5	Results	46
5.5.1	Quantitative Results	47
5.5.2	Qualitative Results	53
5.6	Discussion	58
5.6.1	Revisiting Hypotheses	58
5.6.2	Implications of Robot's Role	60
5.7	Conclusion	61
6	Thesis Conclusion	62
6.1	Summary of Contributions	62
6.2	Follow-up with Local Immigration Services	63
6.3	Limitations	67
6.4	Literature on Using Robots or Digital Technologies to Assist Immigrants and Refugees	68
6.5	Future Work	71
	References	73
	APPENDICES	85
A	Robot Development Supplementary Documentation	86
A.1	Motor Properties	86
A.2	Maple Robot SOP	88

B	User Study Supplementary Data	97
B.1	Participant Demographics	97
B.2	Thematic Analysis - Additional Notable Participant Quotes	99
C	Ethics Clearance Certificate	102

List of Figures

2.1	Kismet the First Social Robot	9
4.1	Appearance of robots considered during selection process	19
4.2	Robot System Architecture Diagram	23
4.3	Robot API Architecture Diagram	24
5.1	Experiment Space Layout and Participant Movement	28
5.2	Robot Facial Expressions	32
5.3	Esperanto Lesson Flow Chart	35
5.4	Learning Interface of Esperanto Lesson: (a) Learning interface when teaching a word during the Esperanto lesson. The Esperanto word to be learned is goja (happy), displayed above a picture illustrating ‘being happy’. (b) Word recognition practice problem during Esperanto Lesson. (c) Pronunciation/Listening practice problem during the Esperanto lesson	36
5.5	Setups of Robot and Tablet in the two experimental conditions.	37
5.6	Ten-Item Personality Measure Items	40
5.7	User Experience Questionnaire (UEQ) Survey Items	41
5.8	Robot Social Attribute (RoSAS) Scale Survey Items	43
5.9	General Language Familiarity of Participants (Counts = Number of People)	49
5.10	Source Language Familiarity of Participants (Counts = Number of People)	49
5.11	Big Five Personality Traits of Participants Comparing Experimental Conditions, with * representing $p \leq 0.05$. Dots represent outliers.	49
5.12	Social attributes of the Maple robot for each condition. *** indicates $p \leq 0.001$	50

5.13	Boxplots of UEQ Scales for System	51
5.14	Usability Scores of Robot in Each Condition Rated by Participants	52
5.15	(a) Quiz scores and (b) Quiz completion time for each condition	53
B.1	Age Distribution	97
B.2	Gender Demographics	97
B.3	Faculty Demographics	98
B.4	Occupation Demographics	98
B.5	Familiarity with Other Languages	98

List of Tables

4.1	Summary of Specifications for Each Robot Considered (Part 1)	20
4.2	Summary of Specifications for Each Robot Considered (Part 2)	21
5.1	Summary of Robot Gestures Created For Esperanto Lesson	31
5.2	Vocabulary taught in Esperanto lesson	34
5.3	Esperanto Vocabulary Quiz	44
5.4	Resulting scales for numerical data obtained from post-processing	44
5.5	Tutor Robot Condition Demographic Characteristics of Participants	48
5.6	Peer-Learner Robot Condition Demographic Characteristics of Participants	48
5.7	Median and IQR (Inter-Quartile Range) Results of UEQ Scales for System	51
5.8	Summary of Themes Identified in Participant Response to Open-Ended Questions	54
6.1	Summary of Literature on Learning Support for Immigrants and Refugees - Purpose, Target Population	69
6.2	Summary of Literature on Learning Support for Immigrants and Refugees - Design, Results, Limitations, Future Work	70
A.1	Motor Position Properties and Joint Mapping Summary for Maple Robot	87

Chapter 1

Introduction

This chapter details the sources of motivation for the research conducted in this thesis, the research goal and scope, and a thesis overview.

1.1 Motivation

The initial discussions surrounding the thesis topic centered on researching the integration of social robots with immigrant children to facilitate language acquisition and provide social support. Within this context, the subsequent exploration of topics such as immigration in Canada, educational challenges, and the use of digital technologies laid the foundation for identifying pressing social issues and led to in-field investigations with local communities. The information collected led to the formulation of the research question and scope outlined in Section [1.2](#).

1.1.1 Immigration in Canada

Canada's population continued to grow older [\[114\]](#), and the fertility rate was below the population replacement level [\[112\]](#). Canada faced depletion in the labor force because the number of seniors aged 65 and older grew six times faster than children aged 0-14, leading to a lack of replenishment [\[116\]](#). The shortage in the labor force was also highlighted by the COVID-19 Pandemic, where job vacancies increased to 80% higher than pre-pandemic levels in late 2021 [\[115\]](#).

Immigration has been the main cause of population growth in Canada. From 2016 to 2021, immigrants made up around 80% of labor force growth, where more than one-third (34.5%) of new economic immigrants were selected under one of the skilled worker programs [114]. This is because close to two-thirds of recent immigrants were of core working age, and amongst these, an increasing number had pre-admission experience in Canada, and a large portion of recent immigrants were selected for their ability to contribute to Canada’s economy [114].

Governments at all levels and many other non-governmental and community organizations utilized this census information to develop and evaluate immigration policies and programs, and to plan and implement education, health care, housing, and other services [114]. This led to the discussion of the motivation for this research.

Due to the surge of new and young immigrants into Canada, the 2016 and 2021 censuses both showed the highest proportion of recent immigrants aged less than 15 years with respect to the total recent immigrant population, which was around 20% for both censuses [113]. From the age pyramid chart of non-Canadian citizens in the 2021 census, those aged 0-4 and 5-9 made up around 1.5% and 4% of total non-Canadian citizens, respectively [117]. UNICEF defines young children as being up to 8 years old, and calls this time period “Early Childhood,” stating it is critical for cognitive, social, emotional, and physical development [121]. This shows a potential need to provide better support for immigrant children, such as the need for learning tools that are more accessible, which social robots can provide, as they can be used anywhere, at any time, and have updated content.

1.1.2 Quality Education, Digital Technologies, and Robots

Digital technologies are becoming a prevalent learning tool [54]. In an era of increasingly fast technological development and integration into all industries, the education sector also witnessed a significant adoption of digital technologies, which created a paradigm shift in the entire education system [54]. Diversifying and enhancing educational tools became essential to meet the evolving curriculum and learning environment. Digital technology also offered the potential to address educational challenges in resource-poor settings [74].

Quality education is a cornerstone of the United Nations (UN) Sustainable Development Goals (SDGs), aiming to ensure inclusive and equitable quality education and promote lifelong learning opportunities for all [82]. Variations in the quality of education mainly between urban and rural areas were prevalent [71, 68, 127], with schools in urban areas generally receiving more government funding and having more resources, while rural schools

faced a lack of resources due to lesser population density and socio-economic challenges [21, 83].

The rise of digital technologies as learning tools, such as mobile devices, smartboards, massive open online courses (MOOCs), tablets, laptops, and online discussion forums, has transformed the education system, providing efficiency in terms of immediate learning environments, faster evaluations, and more dynamic learning [54]. However, the increasing time spent learning online, especially during mandatory online learning periods like the COVID-19 pandemic, brought forth concerns.

Social robots, as a novel addition, began to be recognized for their unique ability to engage and personalize learning in one-on-one interactions [80, 24, 25]. They emerged as a promising solution, particularly in contexts where human resources were limited or where they could offer unique interactive opportunities in education.

1.1.3 In-field Investigation with Local Immigration Services

Understanding the specific situation of the local immigrant population was crucial as this population was anticipated to be involved as informants and/or subjects of the research. Establishing connections with local immigrant communities and professionals working directly with immigrants was important to gain firsthand insight into their challenges and needs.

Investigation and community engagement involved communication and interviews with local municipality and immigration services, aiming to understand the needs and difficulties immigrant families faced, with a focus on the learning and transition challenges experienced by young immigrant children.

A meeting with professionals in the community outreach team at Waterloo Region Immigration Partnership was conducted in July 2022 to gain insight from those working closely with newly arrived immigrant families. This meeting provided an understanding of the challenges immigrants faced and the features a social robot that provides learning support should have.

The main challenges identified for immigrant families were:

- The COVID-19 pandemic created a delay in the admittance of immigrants, causing a backup and resulting in prolonged stays in temporary housing, leading to social isolation for children and difficulties in developing social skills.

- Immigrant children faced challenges in adapting to new school environments and often felt estranged.
- Immigrant parents struggled to help and connect with their children due to cultural differences.
- The diverse backgrounds of immigrants, including varying fluency and skills in their native languages, complicated the learning of a new language.
- Trauma and dislocation experienced by immigrants prolonged their integration and learning of new skills.

Professionals suggested that a social robot should be developed to help immigrant children prepare for a new environment and provide social companionship. They also recommended allowing parents to participate in human-robot interactions to improve the connection and understanding between parents and children.

1.2 Research Goal and Scope

Building upon the insights gained from the motivational exploration and recognizing the complexity of Human-Robot Interaction (HRI), it became clear that successful integration of social robots requires consideration of various components tailored to specific application needs. Despite inherent limitations, robots can be designed to engage effectively with humans across diverse interaction contexts.

The original research goal was to develop small social robots to facilitate engaging interactions with young immigrant children newly arrived in Canada. The aim was to investigate whether social, interactive robots could support these children in transitioning to a new environment and becoming more comfortable with the English language, thereby aiding in overcoming language and cultural barriers. Initially intended as a temporary tool for early adaptation, these robot interactions were designed to be casual and game-like, enabling children to learn commonly used English words and phrases essential for daily life.

However, challenges in recruiting newly arrived young immigrant children and arranging regular interactions with the Waterloo Region Immigration Partnership arose, mainly due to their overwhelming backlog and other responsibilities. Additionally, the development of a working platform for the new social robot was hindered by the absence of a working

simulation, 3D model, or complete example of robot movements in the original product. Consequently, the user study was conducted with participants recruited from within the university.

This thesis aims to:

- Develop and test a social robot for use as a learning and companionship tool specifically focusing on language learning activities.
- Investigate the impact of social robots assuming different social roles, comparing a less investigated role (peer-learner role) to a more common role (tutor role).

The experiment in the user-study aimed to evaluate the robot by gathering participants' perceptions using standardized measures on user experience, usability, and social attributes, along with open-ended questions for feedback and assessments for measuring learning outcomes. Participant backgrounds were also collected to understand the effects of personal differences on perceptions of the robot and the learning experience. The findings are expected to provide insights into how social robotics can be integrated into educational frameworks to support and enhance learning across diverse demographic and geographical contexts.

1.3 Thesis Overview

The remainder of this thesis is structured as follows:

- Chapter 2: Discusses a literature review on relevant topics/research fields, details information gathered from background investigation and community engagement about the target population of this research, and shows the formulation of the research question of this thesis. Relevant topics/research fields include robot roles in education, robot-assisted language learning (RALL).
- Chapter 3: Identifies the knowledge gaps and research question this thesis aims to address.
- Chapter 4: Presents the design and process of the robot platform development, including the robot selection, architecture design of the Maple robot, and of the web application used with the robot.

- Chapter 5: Presents the experimental user study which tests the Maple robot system and explores and compares the influence of the social role the robot assumes on participants' perception of Maple, their learning experience, and learning outcomes.
- Chapter 6: Provides a conclusion of the research and recommended future work.

Supplementary information is included in the Appendices as follows:

- Appendix A: Includes additional specifications and documentation created for the Maple robot.
- Appendix B: Includes supplementary data from the user study conducted.
- Appendix C: Includes the ethics clearance certificate for the user study conducted.

Chapter 2

Background: Literature Review

2.1 HRI and Social Robotics

This section presents an overview of the research fields of human-robot interaction (HRI) and social robotics, and summarizes applications of social robots in HRI scenarios.

2.1.1 Overview of HRI

Human-Robot Interaction (HRI) is a multidimensional challenge at the intersection of multiple disciplines, including psychology, cognitive science, social sciences, artificial intelligence, computer science, robotics, engineering, and human-computer interaction [41]. HRI encompasses scenarios where humans and robots coexist, borrowing methodologies and approaches from other disciplines. The primary objective of HRI is to equip robots with the necessary skills to enable them to interact with humans. To achieve this, HRI investigates humans in interaction with robots to enhance the effectiveness and intuitiveness of robots in HRI [27].

The theoretical background of Human-Robot Interaction (HRI) came from the belief that as robots become more capable, they will transition from highly structured environments like factories or hospitals to more unstructured ones such as homes and outdoors [27]. The importance of HRI has grown due to increased prevalence of robots in everyday life. When robots were uncommon, ineffective or inadequate interactions could be compensated for by more user training. However, with increased likeliness to encounter robots

in everyday life, effective and adequate interactions between robots and untrained user has become more important, as these users may never receive specific training[27].

In HRI research, the emphasis lies on the interaction between humans and robots, rather than solely on autonomous robot behavior [40]. Often in research studies as discussed in [97], interactions may be created with different ways of robot operation. Commonly in studies interactions were simulated was using a Wizard of Oz (WoZ) scenario, where a researcher controlled the robot to achieve desired interactions with participants. Alternatively, experiments may have used prerecorded videos of human-robot interactions for external observation by participants. In the case where autonomous robot was available, direct interactions of the robot acting autonomously with users were studied. Sometimes, a combination of autonomous robot behavior and WoZ was used in HRI studies.

The foundational research in HRI came primarily from Human-Human and occasionally Human-Animal Interaction research in the fields of psychology and cognitive science. For example, the study of imitation learning in robots drew upon developmental psychology, using concepts such as tutelage (where a human being guided the attention of a robot or agent in order to teach it a given concept) [70, 111], and scaffolding (a process in which an experienced individual assisted a novice to solve a problem or carry out a task) [131, 19] to enable effective learning from human guidance. Similarly, the development of verbal communication skills in robots was influenced by psycholinguistics, leveraging elements like verbal deixis and prosody in speech production and comprehension. Robot controllers were made to emulate humans' capacity to identify social cues, including tracking gaze and interpreting facial expressions of emotion [31]. Experiment analysis in HRI adhered to established protocols in these fields, and used methods such as questionnaires or objective coding to evaluate interaction scenarios, including instances of physical contact.

2.1.2 Overview of Social Robotics

In the field of HRI, social robots hold a special role, and are categorized under 'proximate interaction', where humans and robots interacted as peers or companions [55]. Social robots are designed to interact with humans and operate in human environments with people [32].

While there is no universally agreed upon definition of social robots and no consensus on what makes a robot social [55], Amongst various definitions analyzed by Sarrica et. al., [102], it was concluded that social robots can be defined as "(...) physically embodied agents that have some (or full) autonomy and engage in social interactions with humans, by communicating, cooperating, and making decisions. These behaviours are then interpreted by human onlookers as 'social', according to current norms and conventions." [102]. For

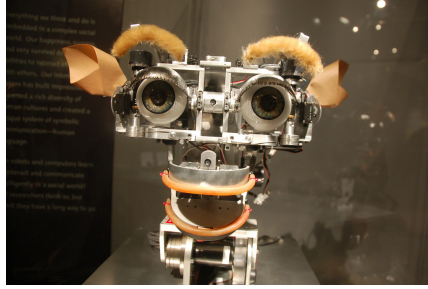


Figure 2.1: Kismet the First Social Robot

example, Figure 2.1 shows Kismet, a robot head made in the 1990s at Massachusetts Institute of Technology (MIT) by Dr. Cynthia Breazeal. Kismet was designed to recognize and simulation emotions for natural and expressive face-to-face interaction with humans. The Kismet robot is widely recognized as the first social robot in the field [30].

As discussed in Henschel [55], Mejia and Kajikawa [76] identified key clusters that represent the core knowledge established in social robotics research. The major clusters in social robotics research are 'robots as social companions' and 'human factors and ergonomics in human-robot interaction'. Research trends show most common practical uses of social robots which are as companions, educators for children, and helpers for the elderly. This trend was also observed in [132], where interviews conducted with robotics researchers in both the USA and Japan saw a common perception of social robots as technological solutions to solve urgent societal problems.

2.1.3 Applications of Social Robots

As explained in reviews by Broadbent [34] and by Maija and Kajikawa [76], the current landscape of social robotics encompasses a wide array of applications, showcasing the development and integration of various robot prototypes into real-world settings. These applications span across healthcare, education, and public services, showing versatility and potential impact of social robots in different domains.

In healthcare, robots are being developed to address the challenges posed by an aging population and increasing healthcare demands (Robinson et al., 2014). One notable example is Paro, a companion robot designed to provide emotional support and stimulation to older adults, particularly those with dementia (Robinson et al., 2013). Research has shown that Paro can reduce loneliness and increase social interactions among users, highlighting its potential as a valuable tool in dementia care (Robinson et al., 2013).

Additionally, healthcare robots such as iRobi and Cafero are being explored for tasks such as medication management, physical assistance, and companionship (Broadbent et al., 2014; Orejana et al., 2015). These robots aim to enhance the quality of life for individuals with chronic conditions or disabilities by providing personalized assistance and support.

In education, robots are utilized for a variety of purposes, ranging from STEM (science, technology, engineering, and mathematics) education to language teaching and cognitive development (Mubin et al., 2013). Humanoid robots like Nao are employed as educational tools to engage students and facilitate interactive learning experiences (Han et al., 2015). Research suggests that robot-assisted education can enhance motivation, learning outcomes, and social interaction among students, particularly those from disadvantaged backgrounds (Han et al., 2015).

Additionally, robots have seen increased deployment as guide robots in public spaces such as shopping malls and museums for visitor assistance and experience enhancement (Sabelli & Kanda, 2015). These guide robots provided information, directions, and entertained users, which made for a more interactive and engaging environment (Sabelli & Kanda, 2015).

Overall, the applications of social robots in healthcare, education, and public services demonstrate potential of robots to revolutionize various aspects of human life. From providing companionship and support to older adults to facilitating interactive learning experiences for students, social robots show promise in addressing societal challenges and improving quality of life for individuals across different age groups and contexts.

2.2 Social Robots in Education

2.2.1 Robot Roles in Education

When used for education, robots can take on social roles such as tutor/teacher (the most common role), peer, novice to the human learner. As a tutor/teacher, robots provide direct curriculum support through hints, tutorials, and supervision (e.g., [62, 44]). As a peer, robots are usually portrayed as the more knowledgeable peer and guides the learner (e.g., [44]). As a novice, the learner teaches the robot (e.g., [118]) [25].

Feil-Seifer [47] provided an early insight into the role of robots in education, highlighting their potential in specific instructional domains. It found that robots could be more effective than humans in certain teaching contexts. Robots have been utilized not only

for regular academic instruction but also for teaching social skills, such as imitation, self-initiation of behavior, and were found to be effective tools in special education, such as provoking proactive social behavior in children with Autism spectrum disorder [47].

Building upon these earlier findings, Belpame et al. conducted a comprehensive review that further explores the evolving role of robots in educational settings. Their meta-analysis reveals that robots predominantly functioned as tutors (48%) and teachers (38%), with a small proportion (9%) serving as peers or novices [25]. This data indicates that there is a lot less research into robots assuming peer or novice roles, compared to tutor and teacher roles.

One of the more notable examples of a peer robot is Robovie, the first fully autonomous robot introduced in an elementary school, which represents the peer-to-peer interaction approach in education. Deployed in a Japanese elementary school to assist with English language learning for first and sixth graders, Robovie demonstrated that some children improved their language skills significantly. Interestingly, when compared to an otherwise identical tutor robot, a peer robot like Robovie elicited longer attention spans, faster, and more accurate responses in learning tasks [130].

However, Belpame et al. noted that the effectiveness of these robot-led teaching strategies for adult learners is yet to be confirmed, suggesting that methods successful with young children might not directly apply to older learners [25]. Furthermore, Belpame et al. saw social robots as an integral part of future educational infrastructure, similar to traditional tools like paper, whiteboards, and tablets. They emphasized that these robots offered unique personal and social dimensions to learning, potentially transforming and enhancing the educational experience by tailoring it to individual learner needs. The authors also state that use of social robots in education curricula could not only address current resource limitations in education but may also offload work from human teachers and thus allow them to focus more on their strengths in providing comprehensive, empathetic, and rewarding educational experiences [25].

Out of the smaller portion of robots that assumed peer roles, many studies discovered benefits of peer-based interactions with robots for children, and indicated children preferred this mode over tutor-based interactions [23, 130, 90].

A notable study by Zaga et al. specifically investigated the effects of a robot’s social character either as a peer or a tutor on children’s engagement with tasks [130]. In this study, children of the same age and gender were grouped together and assigned tasks under two conditions: the peer character or the tutor character. The distinction between conditions was in the robot’s behavior, which varied in gestures, speech, and postures to align with either peer-like or tutor-like characteristics. The study used the NAO robot, oper-

ated remotely via a Python script in a Wizard of Oz setup, where a researcher controlled the robot. The findings were highly significant: children in the peer character condition not only maintained their focus on the robot and the task for a longer duration but also completed puzzles more quickly and effectively than those in the tutor character condition. This research underscores the significance of a robot’s role and interaction style in influencing children’s engagement and performance in learning tasks.

For adults, investigations into whether peer-based interactions have the same benefits as tutor-based interactions are very limited, and translating findings from studies with children to adult learners is unsuitable. Adults may have different expectations from a robotic peer, potentially due to their more extensive experience in interacting with human peers and a higher level of maturity and cognitive development than children. These factors could lead to adults being less receptive to a robot peer.

A notable work that contributed to this knowledge gap is a study conducted by Salomons et al. which focused on how adults with varying levels of prior domain knowledge respond to peer versus tutor robots [101]. In this experiment, participants built electronic circuits with the assistance of an autonomous robot (UR5 robot) that either took on a tutor role or a peer role. The robot’s role was manipulated through its communication style. In the peer condition, the robot used inclusive pronouns like “we” and “us,” positioning itself as an equal and collaborative partner. Conversely, in the tutor condition, it used “you,” establishing itself as an authority figure. Participants underwent pre-tests and post-tests to measure the skills acquired and completed questionnaires to share their perceptions of the robot. The study’s outcomes revealed that all participants improved their skills from pre-test to post-test. However, a significant difference was found among participants with lower initial knowledge, who learned more effectively with the peer robot than with the tutor robot. Interestingly, the peer robot was perceived as more intelligent, sociable, and friendly across all skill levels. Furthermore, participants felt more respected by the peer robot compared to the tutor robot.

These findings highlight the potential of peer-like robots in adult education, especially for those with limited prior knowledge in a given domain. They also suggest that the way robots present themselves and communicate can significantly impact learning effectiveness and the overall perception of the robotic instructor.

2.2.2 Robot-Assisted Language Learning

Robots have also increasingly been implemented as tutors in both first- and second-language education, where more common digital technologies such as instructive virtual

games, chat programs, tablets, and animated books are being used widely [123]. Robots have advantages over the aforementioned technologies such as allowing the learner to interact with the real-life physical environment, such as through manipulation of real-life objects or body movement and gestures, which have been found to benefit children’s vocabulary learning. Another advantage is robots can create more natural interaction than other forms of technology because of their appearance as humans tend to anthropomorphize robots, thus making robots be more likely to be seen as a tutor/teacher, peer over rather than a machine over traditional digital-technologies such as computers or tablets. Since the use of a robot would allow for more opportunities for interaction with physical environment and they are more capable of promoting natural, enjoyable interaction, they present a significant advantage over traditional computer-aided methods of learning [123].

In a survey conducted by Randall [95], the author looks into research area of Robot-Assisted Language Learning (RALL), a sub-domain of Robot-Assisted Learning (RAL), which is defined as the use of robots to teach people language expression or comprehension skills such as speaking, writing, reading, or listening in native and non-native language instruction. RAL is an area in social robotics and human-robot interaction (HRI) that focuses on using robots to assist humans in learning. RAL and RALL focuses on aspects such as the importance of acceptability by users, the desire for interpretable and safe interactions, and relationship development. In addition, for RAL and RALL, robots in teaching or tutoring roles are also to be judged by their efficacy in information transfer. Therefore, there need to be RAL domain-specific considerations when designing a robot’s form, speech, and behaviors. Still, RALL incorporates findings from HRI more generally and can use previous work to guide future areas of study [95]. Findings in RALL can also be applied to HRI more broadly—such as the novelty effect, which was found by Kanda et al., one of the first studies on RALL [58]. Thus, findings from HRI and RALL can inform each other for better research design.

In a summary by Engwall & Lopes [46] on methods in RALL, the author concluded that robot types, teaching strategies and robot roles should be combined differently to take advantage of intrinsic characteristics of the robot for the application, and to suit different learner age groups. The author also concluded that RALL designed for adult learners may benefit from a combination of 1) a teaching strategy directly related to human–human situations to increase realism of communication, 2) a robot role allowing learners to be involved in defining the practice (i.e., mainly using peer, learner, opponent or social companion roles), and 3) a robot type that allows for the use of human-like verbal and non-verbal interaction signals and an interaction promoting the learners’ own intrinsic motivation (focusing on social exchange) rather than extrinsic motivation (focusing on explicit rewards in the practice). [46].

Existing teaching strategy methods in human-human situations that may be used with robots as detailed by [46] included methods such as the practice of specific learning material which involves audio-visual methods like multimedia-based education (MBE) where images and words are shown on a screen, and audio-lingual method (ALM) where the learner mimics audio utterances; physical interaction methods such as total physical response (TPR) method where body gestures which are connected to verbal instructions are performed to enhance interaction; communication based methods such as communication learning teaching (CLT) where emphasis is put on encouraging information transfer, fluency and confidence, rather than structured conversation in practice; role playing strategies based on task-based language teaching (TBLT) which focuses on establishing a social relationship between the robot and learner in a presented scenario (e.g., a shop scenario where people play the customer or shop keeper role); and collaborative language learning (CLL) where there is peer support in the learning process.

Robot role may be chosen based on the level of learner initiative. Engwall & Lopes [46] identified five different roles: 1) teaching assistant role which acts as a motivational enhancement by introducing a novel element in traditional classroom scenarios where a human teacher is present, 2) tutor role where a robot teaches and interacts with the human learner without a human teacher, 3) peer, partner, opponent, or tool role where the robot is presented as having similar knowledge level as the human learner, 4) tutee role which is defined by Engwall & Lopes [46] as the human learner having to teach the robot as part of the learning process for the human, and 5) social companion role, where robot and human learner engages in social interaction to practice. In the case of a teaching assistant role, Engwall & Lopes [46] rationalized that the robot's appearance may be more robot- or toy-like, as the interaction with a robot is expected to be different than with a human teacher and thus less anthropomorphic features may be an advantage to enhance the robot's motivational effect on the human learner. For a tutor robot, the robot must convince learners that it has a good mastery of the practice material. This introduces learner age dependent requirements on credibility in appearance and interaction. A similar logic applies to peer robot as tutor robots, where Engwall & Lopes [46] recommended that the more realistic the task and more advanced the learners are, the more anthropomorphic the robot needs to be in appearance and interaction behaviour. For learner robots, the robot should behave and interact like a less advanced peer, and in terms of appearance, the more advanced the learning task and human learners are, the more humanoid the robot should be. For social companion robots, a higher degree of anthropomorphic interaction is required because the role is based on creating a social relationship from human-human interaction.

Chapter 3

Research Question

From the related literature, we found that the peer-learner robot role has been investigated less than the tutor robot role [25]. In Robot Assisted Language Learning, there are also few previous works in RALL with adult learners, which means a lack of design guidelines, which made the design of the robot behavior and tasks very experimental for RALL [46]. It was concluded by Engwall & Lopes [46] who formulated general guidelines for RALL design, that combining existing teaching strategy methods that are used in human-human situations, different robot roles, and different robot types (to take advantage of the appearance and functions of the robot type) can be done to suit different learner age groups.

Inspired by the study conducted by Salomons et al. [101], [130] and taking into recommendations for RALL design from [46], this research aims to compare the results and findings from a user study where a humanoid robot assumes either a tutor role or a peer-learner role in its interaction with participants in a beginner level language learning activity that applies teaching strategies from human-human situations. The peer-learner role in this case is a robot presented as a peer who is also learning from the activity alongside the human (the human does not teach the robot). The teaching strategy used is one that combined elements of MBE and ALM to show learning material and have visual and audio to follow and during learning and practice, along with CLL method, where the robot provides support regardless of its role; using robot roles of tutor and peer which represent the different conditions being investigated; and a humanoid-like robot to match the tutor and peer role for adult learners.

Consequently, the aims of this research are to compare the results and findings of a humanoid robot taking different roles in its interaction with participants in a beginner level language learning activity. The following research question (RQ) was formulated:

RQ: "How does the use of a peer-learner humanoid robot in educational settings influence perceptions related to Human-Robot Interaction (HRI) and affect learning outcomes?"

This research question was addressed through by the following things done: developed software that added social interaction functionality to a robot, designed and conducted a HRI user study and analyzed the results of the study. These are described in the following chapters.

Chapter 4

Robot Platform Development

This chapter details how the robot platform used in this research was developed. The process included the robot selection process, development of the robot control framework and its integration with the webpage applications which served as a complete platform for learning applications being investigated in this thesis. Documentation for the robot was also developed to allow future development, continuous usage and maintenance.

4.1 Robot Selection Process

Robot selection criteria focused on robot specifications such as price, size, appearance, degrees of freedom (range of movements possible), battery life (and charge time), connectivity and network, computational power, sensors, the ability for third-party development such as software development kit (SDK), application programming interface (API), and expandability of hardware and software. A list of robots was compiled and a comparison was conducted to make a final decision. The following sections describe the selection process and final choice.

4.1.1 Robot Comparison and Selection

One of the most important specifications was the size of the robot. Initially, only palm-sized robots were considered as the child HRI research originally envisioned was to have a portable robot that children can carry with them, which would allow for more opportunities for child-robot interaction. At the start of this research (Spring 2022), palm sized

robots commercialized on the market were the EMO by Living AI [3], the Vector 2.0 by Digital Dream Labs (DDL) [9], and Cozmo 2.0 also by Digital Dream Labs [2], which were marketed as pet-companion, a companion and assistant, and a companion learning toy/tool, respectively. Notably, the vector and cozmo robots have been used in research, such as [91, 56, 39, 129]. Other larger sized more feature-packed social robots commonly used in HRI research with education applications were also considered. This included the NAO by Aldebaran [1], Misty 2 by Misty Robotics [6], MiRo-E by Consequential Robotics [5], and QTrobot by LuxAI [4], their sizes approximately close to a baby or young toddler (35 cm to 60 cm height range). Another robot, the Dr. R configuration of the Engineer Kit 1 by Robotis [12] was also considered, it is not palm-sized but around the size of a stuffed toy (23 cm in height). The appearance of each robot is shown in Figure 4.1. The specifications and features of each robot that was considered are summarized in Table 4.1 and Table 4.2.

After a comparison of robot properties, the EMO robot was eliminated due to a lack of third-party development software and framework. The Cozmo robot was eliminated due to a lack of onboard memory and its reliance on storing data in the cloud in the Cozmo mobile application, which would have added unnecessary complications in data retrieval and data privacy concerns. While the Cozmo SDK is advertised as being readily available for download, its documentation and downloads for the SDK were unavailable from its official site, making it unreliable for long-term development. The Vector robot was eliminated due to production delays [8]. Had there been no delay, it would have been considered more thoroughly as the research team had prior experience using it for HRI research, and initial review of its specifications seemed sufficient and suitable. The NAO robot was also eliminated due to inaccessibility. The MiRo-E was eliminated due to its animal and pet-like appearance. The MiRo-E resembled too close to an actual pet animal such that it would be unsuitable for use in any HRI study where the robot is doing human activities such as being part of a language lesson as a tutor or as a peer-learner. The QTrobot was eliminated due to limited sensor features, which rely on extracting information from its 3D depth camera. This posed a significant disadvantage because in order to have more sensors, a lot of additional hardware needs to be mounted externally, which can be a negative influence on the appearance and making it a less attractive choice as there are other comparable robots with more sensors built in. The Misty robot was eliminated for its low DoF (6 total), which creates a constraint for the types of movements Misty can perform and thus also limit future development. The size of the NAO, MiRo-E, QTrobot, and Misty were also still too big for the purpose of this research and was another reason for elimination. A good middle ground between small size and technical features was found in the Dr.R robot in the Engineer Kit 1 by Robotis. Dr. R robot has several similarities to QTrobot.



(a) Emo by Living AI [17]



(b) Vector 2.0 by DDL [125]



(c) Cozmo 2.0 by DDL [78]



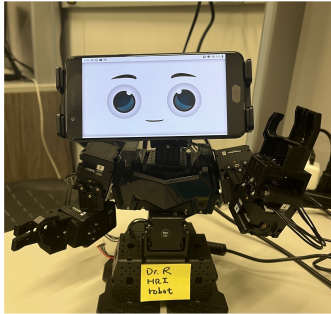
(d) Misty II by Misty Robotics



(e) MiRo-E by Consequential Robotics [57]



(f) QTrobot by LuxAI



(g) Dr. R - Engineer Kit 1 by Robotis



(h) NAO by Aldebaran [38]

Figure 4.1: Appearance of robots considered during selection process

Dr.R is nearly 1/3 the size of QTrobot, also has 12 DoF, has third-party development and expandability, and same battery life. In addition, Dr. R is equipped with more sensors than QTrobot, such as touch/force, acceleration, gyroscope, and actuator temperature and voltage sensors. Thus, the robot chosen for use in this research was the Dr. R.

Robot	EMO	Vector 2.0	Cozmo 2.0	Engineer Kit 1 - Dr. R
Company	Living AI	Digital Dream Labs	Digital Dream Labs	Robotis
Appearance	Abstract	Vehicle-like (Forklift)	Vehicle-like (Forklift)	Humanoid
Size & Weight	Palm-sized, Weight: 250 g	Palm-sized, Weight: 250 g	Palm-sized, Weight: 250 g	Height: 23 cm Width: 23 cm Depth: 10 cm Weight: 0.8 kg
Product type	AI Desktop Pet	AI Robot Companion	Educational Toy Robot	Smart AI-Based Multi-Joint Educational Robot Kit
Recommended use(s)	Pet companion	Home robot (Assistant and companion)	Learn STEM subjects using Cozmo apps	Education, hobby, entertainment, competition,
DoF	2	2	2	12
Base Mobility	Leg base bipedal-like	Wheeled base	Wheeled base	Immobile base
Upper Mobility	Head movement	1 DOF raiseable arm	Raiseable arm Kinetics K02	Arms, Torso, Head
Computation	CPU	1 x Qualcomm 200	microcontroller, (no onboard memory)	1 ARM MCU
Wifi, Bluetooth	Yes	Yes	Yes	Yes (part of controller)
Microphone	Yes	Yes	Yes	Yes (part of controller)
Speaker	Yes	Yes	Yes	Yes (part of controller)
Camera	1 x 2 MP, 15 fps	1 x 2 MP, 15 fps	1 x 2 MP, 15 fps	No (Expandable)
Display	Built-in screen (non-touch)	Built-in screen (non-touch)	Built-in screen (non-touch)	Attach any external device with screen
Sensors	Distance, Drop, Touch, Acceleration, Light	Distance, Drop, Touch, Acceleration, Light	Distance, Drop, Touch, Acceleration, Light, Gyroscope	Force, Acceleration, Gyroscope, Voltage, Temperature
Hardware extensibility	None	None	None	Yes (UART, USB Ports)
Development Platform / Architecture	None	Vector SDK (requires purchase), information not available	Cozmo SDK, Cozmo Code Lab (Android/iOS app), information not available	ROS SDK (C++, Python, JS, etc)
Available Simulation	None	None	None	None
Battery life (on full charge)	up to 3 hours	30-45 minutes	1-2 hours	4 hours active
Price (USD)	\$399.99	\$399.99	\$499.99	\$1,089.00

Table 4.1: Summary of Specifications for Each Robot Considered (Part 1)

Robot	Misty II	MiRo-E	QTrobot	NAO
Company	Misty Robotics	Consequential Robotics	LuxAI	Aldebaran
Appearance	Humanoid	Pet-like	Humanoid	Humanoid
Size & Weight	Height: 35.56 cm Width: 20.23 cm Depth: 25.4 cm Weight: 5 kg	Height: 36cm Width 34cm Weight: 3.3 kg	Height: 57cm Width: 31 cm Depth: 27 cm Weight: 5 Kg	Height: 58 cm Width: 31 cm Depth: 28 cm Weight: 5.5 kg
Product type	Robotics platform	Robotic device for education	AI-Powered Social Robot	Interactive companion, open programmable platform
Recommended use(s)	Teaching tool, Research, Development	Academic research and development, Education, Healthcare	Human-AI Teaching and Research	Research, Education, Healthcare
DoF	6	11	12	25
Base Mobility	Wheeled base	Wheeled base	Immobile base	Bipedal
Upper Mobility	Raiseable arm	Head Movement	Arms, Torso, Head	Arms, Torso, Head, Hand, Fingers, Eyes
Computation	2 x ARM MCUs	GPU, 1 x Raspberry Pi	1 intel (QTPC), 1 Raspberry Pi (QTRP), GPU & accelerator	ATOM Z530 1.6 GHz CPU
Wifi, Bluetooth	Yes	Yes	Yes	Yes
Microphone	Yes	Yes	Yes	Yes
Speaker	Yes	Yes	Yes	Yes
Camera	1 x 13 MP fisheye camera	2 x HD cameras	1 x 3D camera	2 x 2D cameras
Display	Built-in screen (non-touch)	Built-in screen	Built-in screen (non-touch)	None
Sensors	Touch, Distance, Bump,	Touch, Position, Detection	Depth camera,	Position, Temperature, Current, Stiffness, Touch, Gyroscope, Acceleration, Sonar, Force, Center of pressure
Hardware extensibility	Yes (UART, USB ports)	Yes (UART, USB ports)	Yes (UART, USB ports)	Yes
Development Platform / Architecture	HTTP APIs, Misty Studio	MDK and API, MiRoCODE (Blockly, Python),	ROS API (Python, C++, JS), Graphical Studio (online)	NAOqi API, NAOqi Framework,
Available Simulation	ROS Gazebo package	ROS Gazebo package	None	Webots project
Battery life (on full charge)	2+ hours	6+ hours active, 12+ hours standby	4 hours	1 hour active, 1.5 hours normal
Price (USD)	\$3,995 to \$4,295	~\$3,100 (£2,450)	\$10,000 to \$16,000	\$18,990

Table 4.2: Summary of Specifications for Each Robot Considered (Part 2)

4.2 About the Maple Robot

The hardware in the Engineer Kit 1 includes one controller, 6 two degree-of-freedom Dynamixel X-series actuators exclusive to the kit, one lithium polymer battery, and parts for assembling a robot in either one of the recommended configurations (Dr.R, SPI, Max-EI) by Robotis [12], or any custom configuration. This research uses the Dr.R configuration, which resembles a humanoid robot and allows for a smartphone to be attached as its face and also can be a medium for displaying content. The robot was given the name "Maple", inspired by the name of Canada's national tree called the Sugar maple.

4.3 Robot Framework

4.3.1 Robot System Architecture

Figure 4.2 shows the robot system architecture.

In the development of the robot system architecture, the primary goal was to establish a foundation that not only met current requirements but also allowed for further development and integration for more sophisticated high-performance robotic applications. The robot and system was designed with pre-programmed movements and sequence of events which ran automatically without a human in the loop.

For development, the Robot Operating System (ROS) [86] middleware was used for its essential tools and libraries for robot software development, which included hardware control, common functionality, process communication, and cross-platform compatibility. C++ was chosen for its high performance and extensive library support, which aligned with our goal of creating a dynamic and scalable system architecture for sophisticated robotic applications. This combination of ROS and C++ enabled adaptability and scalability, and ensured the architecture remained relevant for future advancements if needed.

The ROS Noetic distribution was used, it was the latest long-term support version of the first version of ROS [87], and was chosen for its compatibility with Robotis's Dynamixel software development kit [11, 99] and with Ubuntu 20.04 [87].

A Raspberry Pi 4B, was used to house an Ubuntu Server 20.04 LTS [119] with ROS Noetic, Dynamixel SDK installed. This handled the computational demands and facilitated wireless control of the robot through SSH for direct command-line access via ROS, and

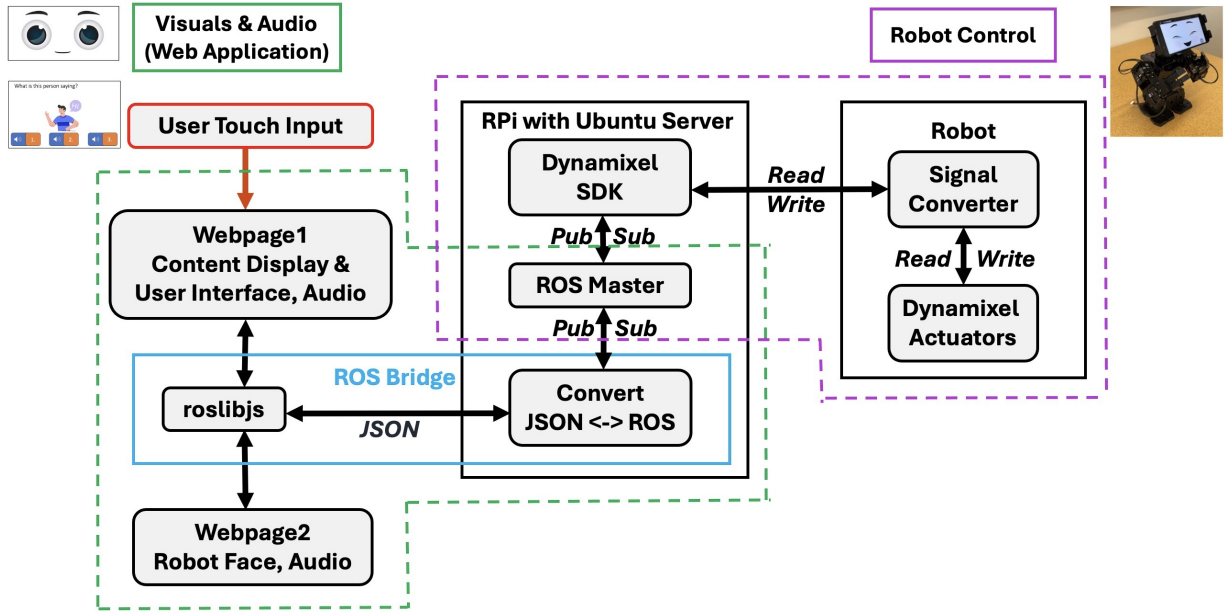


Figure 4.2: Robot System Architecture Diagram

wireless control of Esperanto lesson web application interface using ROS library which communicated between the robot and web application via ROS also.

The mechanical movement of the robot was driven by Dynamixel X-series Motors/Actuators from Robotis, which were manipulated using the control functions in the Dynamixel Software Development Kit (SDK).

On the software front, the Dynamixel SDK for ROS was used, it is optimized for C++ integration, and provided a library for controlling and interfacing with the Dynamixel motors within the ROS ecosystem. Additionally, a GitHub repository maintained by Robotis offered documentation which were useful for customizing motor controls and programming specific robotic functions.

USB communication of the robot to other devices was enabled by the U2D2 USB communication converter from Robotis [13], a key interface component between the Raspberry Pi and the Dynamixel motors. The system's flexibility in power sourcing is further enhanced by the U2D2 Power Hub from Robotis [14], which is equipped with multiple power connectors and can be connected to either a power supply or an external battery.

This robot system architecture, developed with C++ and ROS, resulted in a scalable and easy-to-use platform that meets the current current needs and is well-suited for the

ongoing evolution and integration of more advanced robotic hardware and software frameworks for research.

4.3.2 Robot Actuator Control API Architecture

Figure 4.3 shows the API architecture for controlling the robot. To control the Dynamixel actuators of the robot, the ROS was used in conjunction with the C++ library of the Dynamixel SDK. The SDK functions SyncSetPosition, SyncGetPosition, GetPosition, SetPosition were used to read and write actuator positions. The actuators can assume positions in the range 0-4095 and need to be converted to degrees for easier understanding during robot movement and gesture design and development, equation 4.1 shows the conversion that was found after rounds of testing and documenting joint positions. Joint limits of each actuator also needed to be tested in the current robot configuration to prevent collisions between different parts of the robot hardware. Table A.1 in the appendices summarize the zero position, rotation plane, position range limits and positive orientation of each motor, and corresponding joint number of each actuator in the Unified Robotics Description Format (URDF) file used for simulation, which were found from several rounds of testing.

$$(Pose_{motor}) = (Pose_{deg}) \times \frac{4095}{360} \quad (4.1)$$

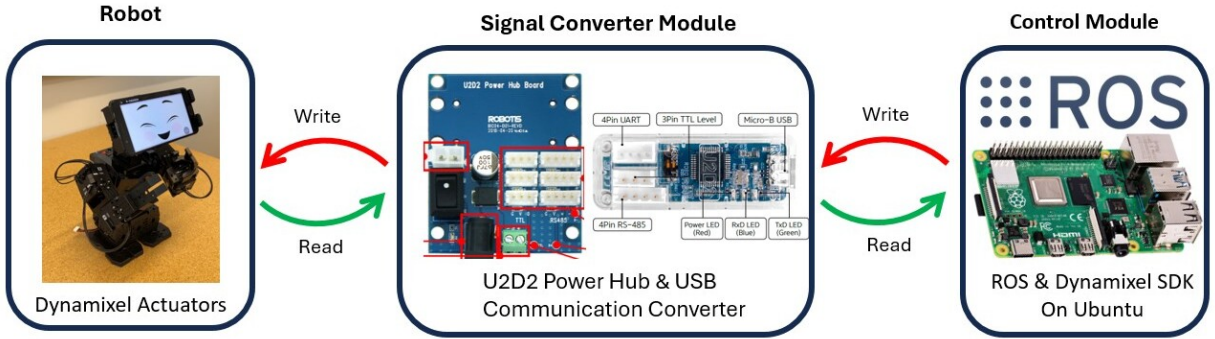


Figure 4.3: Robot API Architecture Diagram

4.4 Web Applications for Robot Face, Content Display and User Interface

Two web applications were made, one which displayed the Maple robot's face, and one which acted as a user interface and also displayed content. The web applications were made using React [7] which is a Javascript library that served as a framework for building user interfaces, and the ROS Javascript library which allowed networking with the physical robot. The React webpage application included the following components: app framework and ROS client libraries, resources (ex. images, audio), networking protocols for ROS, state and time variables that control if a resource is being displayed or played and coordinates what happens in the app. The integration of the two web applications with the Maple robot is shown in Figure 4.2.

4.5 Documentation for Maple Robot

Documentation was developed for servicing and operating the robot (turning on the robot and bringing it online (i.e., connecting network and to ROS and starting scripts); the start-up procedure; operating procedure; shutdown procedure; all related maintenance information; and safety information. All this documentation can be found in the safety operation procedure (SOP) document for the Maple robot, the full document can be found in Appendix A.

Chapter 5

User Study: Language Crash Course with a Robot

This chapter is a publication submitted to International Journal of Social Robotics (SORO) in February 2024.

5.1 Hypotheses

The following hypotheses are formulated for the study done in this research to address the research question (see Chapter [3](#)):

H1: Participants will have higher quiz scores when robot assumes a peer learner role.

H2: Participants will take less time to complete the quiz when the robot assumes peer learner role.

H3: Participants will perceive the robot in a peer learner role more positively (e.g., higher warmth, lower discomfort, higher competence).

H4: The robot assuming a peer learner role will yield higher participant ratings in terms of usability and user experience.

5.2 Methodology and Design

This study took a between-participant design where one of two conditions, Tutor robot condition (T) and Peer-learner robot condition (P) was assigned to each participant. The rest of this text may refer to the conditions as Tutor condition and Peer condition, or T and P, respectively. The experiment consists of each participant learning 12 vocabulary words in Esperanto in the presence of Maple the robot (which assumed a different role depending on the condition), using a tablet to display the learning material and answering practice problems.

5.2.1 Study Recruitment

Participants were recruited from within the university which can include students, staff, and faculty. Recruitment methods included putting up posters across the university campus and advertising in a university wide digital newsletter that is delivered via email. A remuneration of 15 Canadian dollars was given to participants upon completion of the study, which took up to one hour for each participant. Inclusion criteria included that the participant's age must be 18 years of age or older. In post-processing of the data, participants were either excluded due to technical issues during the experiment or because they had prior experience with Esperanto.

5.2.2 Experiment Procedure, Design and Materials

Across both experimental conditions (Tutor (T), or Peer (P)), we ensured that the robot had the same voice in terms of tone, pace, and pitch. The behavior of the robot varied across conditions as it assumed different roles. The lesson content and teaching content were the same across both conditions. The entire study for each participant took 30 minutes and up to 1 hour. The following sections explain the experiment design in detail.

Study Procedure and Setup

The study procedure is as follows and figure 5.1 shows the study space, equipment layout and flow. The procedure of the study and experiment is as follows. For each item of the procedure, the experimenter told the participant what component was being completed and helped the participant navigate to the corresponding location within the study space. While

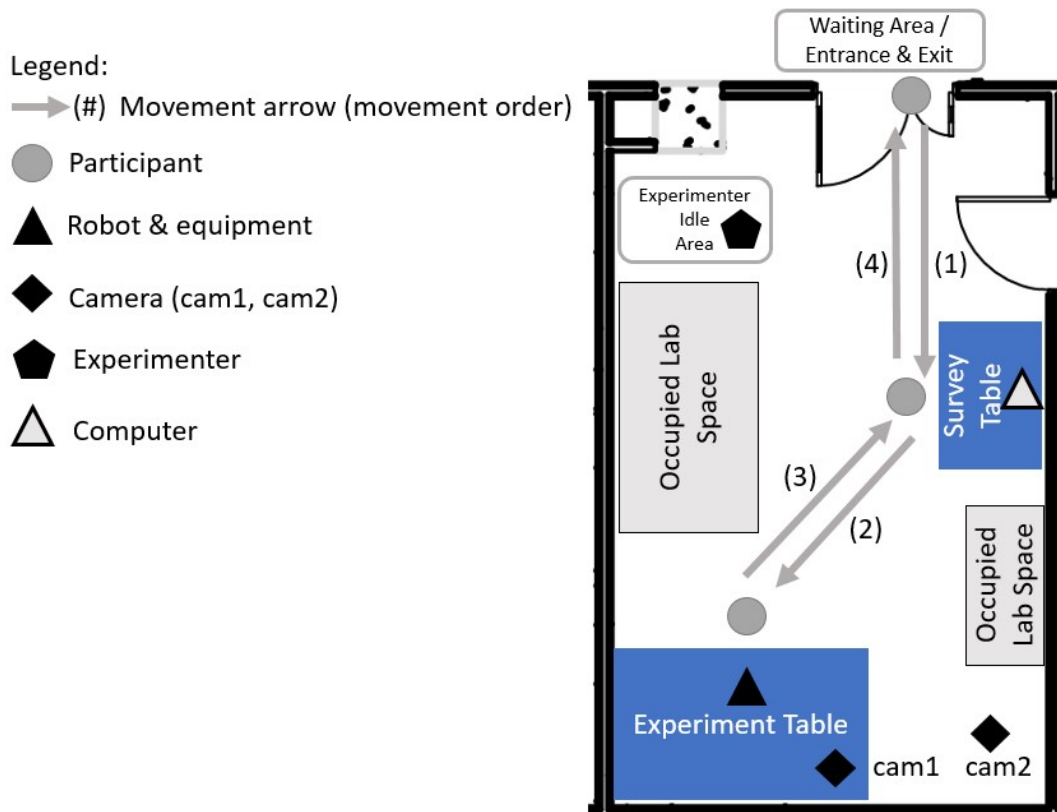


Figure 5.1: Experiment Space Layout and Participant Movement

the participant completed each component, the experimenter waited in the "experimenter idle area" shown in Figure 5.1.

1. Briefing: The participant is welcomed and is given a run-down of the general steps in this study by the experimenter.
2. Pre-experiment component: The participant signs a consent form and completes pre-experiment questionnaires on a computer at the survey table.
3. Start of the experiment: Participant moves to the experiment table, sits down, and is prompted by the experimenter to press "Start Lesson" to start
4. Robot introduction: Once the "Start Lesson" is pressed, the robot "activates" (goes from resting to awake mode) and greets and introduces itself and its role to the participant (robot dialogue for role identification in T: "Hello, I am Maple, your tutor for this lesson...", in P: "Hello, my name is Maple... today I am also here to attend this lesson.")
5. Background on Esperanto: depending on the condition either the robot or tablet provides a brief background on the origins of Esperanto
6. Esperanto Lesson: Teaching occurs (depending on the condition either the robot or tablet is the teacher and the robot is either a peer or tutor) and practice problems are administered. The robot says goodbye to the participant at the end of the lesson.
7. End of the experiment: After the lesson ends, the experimenter tells the participant that the experiment has ended and asks the participants to move to a computer at the survey table to complete the post-experiment components.
8. Post-experiment component: Participant moves back to the table with the computer and completes questionnaires.
9. Quiz: After completing the post-experiment questionnaires, the participant completes the quiz and is given a maximum time limit of 15 minutes.
10. Debrief: The Participant is provided a detailed explanation of the motivation and significance of this experiment, and the applications and potential uses of the results of this experiment and future development.
11. Remuneration: The participant is issued remuneration (15 Canadian dollars) for completing the study.

5.2.3 Robot Behaviour Design

Movements Design

In Belpaeme et al. (2020), the focus on gesture and movement speed highlights that full body posture and movement are key in communication, just like facial expressions. Postures help in deciphering emotional states; slow movements, drooping shoulders, and lethargic gestures typically indicate a downcast mood, while rapid movements and upright posture signal a positive attitude [22]. In light of this, the design of robot gestures in the study was intentionally brisk, complementing an always upright resting posture. This approach was adopted to reflect a positive demeanor, aligning with the understanding that physical expressions are integral to effective non-verbal communication.

Robot Behavior - Gestures

A total of 8 gestures were created for the robot and were used in both conditions. The gestures are waving with a single arm, raising both arms over the head, raising both arms to shoulder height and flipping forearms out to indicate a question, bringing both hands together (i.e., clapping), gesturing/pointing with a single arm (right or left), nodding and tilting forward of torso, head tilt (left or right). Due to the robot's placement on the table in the peer-learner robot condition, a rotation of the robot's torso was added and paired with the gestures because the interaction between the robot and participant required the robot to orient itself such that it was facing the participant during an interaction. Table 5.1 shows the gestures implemented for the robot in the experiment.

Robot Behavior - Facial Expressions

4 Facial expressions of the robot were used and include neutral, smiling, sad, and resting face shown in Figure 5.2a, 5.2b 5.2c, 5.2d. All animated and the fully animated expressions can be seen in the demonstration videos. In all facial expressions except the resting face, the robot's eyes are blinking, whereas the resting face shows the robot's mouth opening and closing to indicate it is sleeping (i.e., resting). Only four expressions were chosen because each can represent a range of emotions depending on the immediate context. Different contexts are made by coupling different gestures and dialogue, and also on the type of activity in the lesson (such as teaching, practice pronunciation, or practice problems).

Category	Movement	Meanings and Uses
Arm	Bring both hands together	Clapping, indicate happiness, celebration, beat gesture, illustrator gesture
	Pointing with one arm (left or right)	Draw attention to something, deictic gesture
	Single-arm wave	Greeting
	Raise both arms overhead	Indicate happiness, celebration, general communication gesture
	Raise both arms to shoulder and flipping both forearms out	Indicate a question
Head	Tilt head (left or right)	Indicate question, attention, curiosity, idle movement, beat gesture
	Nod	Indicate agreement, acknowledgement
Torso	Tilt forward	Tutor condition only: used in conjunction with gesture to show robot is looking at participant's response to practice problem
	Rotation (left or right)	Peer-learner condition only: used to move robot such that robot faces participant where there is an interaction

Table 5.1: Summary of Robot Gestures Created For Esperanto Lesson

The neutral face is used to show the attention of the robot, e.g. during ‘listening’, ‘thinking’, or generally speaking. It is the default facial expression used during the experiment. It is used by the robot in the following cases: 1) When the tutor robot is teaching or providing instructions during the lesson, 2) When the robot is not moving or speaking during learning, pronunciation, or practice problem activity in the lesson (e.g. listening or waiting for the participant or a new lesson).

The smiling face is used to represent happiness and excitement. It is used by the robot in the following cases: 1) when teaching the word *goja* (happy) to illustrate the word, 2) to encourage the participant during pronunciation, 3) as a reaction when a practice problem is answered correctly by the participant, or 4) when the robot is making remarks during the lesson.

The sad face can represent sadness and feeling apologetic. It is used by the robot in the following cases: 1) When teaching the word *malgoja* (sad) to illustrate the word, 2)

as a reaction when a practice problem is answered incorrectly by the participant, or 4) when the robot is making remarks saying that it does not know the answer to a practice problem.

The resting face is used in two specific cases: 1) It is used to show the robot is in a resting state which occurs at the start of the lesson where the robot activates in response to the participant clicking on the start button, and at the end of the lesson when it returns to resting mode after saying goodbye to the participant. 2) It is used when teaching the word *dormu* (sleep): the tutor robot would display this face when it teaches the word, and the peer-learner robot would show this face when it pronounces the word.

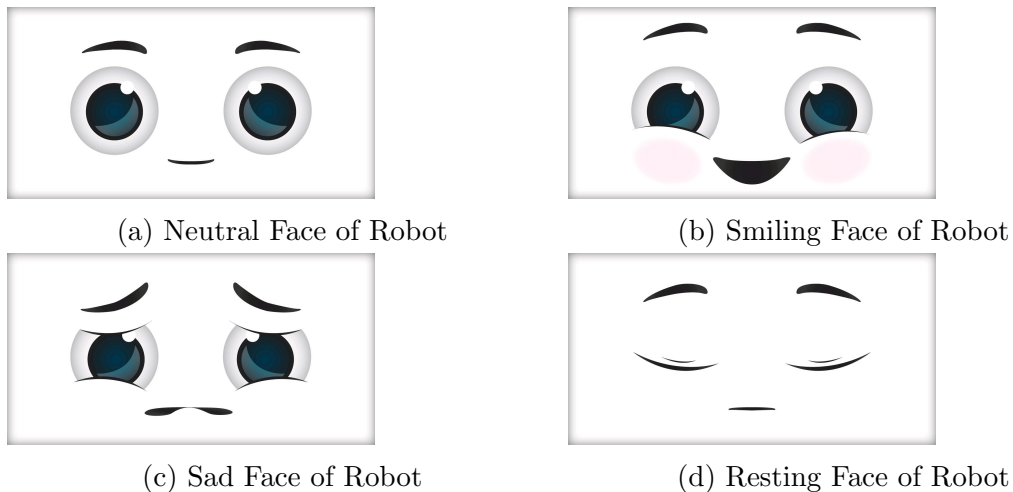


Figure 5.2: Robot Facial Expressions

Robot Behavior - Voice and Dialogue

The voice of the robot was generated using the Justin voice from Amazon Polly Text-to-Speech standard voices [110]. The Justin voice is US English and sounded semi-robotic, with human-like intonations, which we thought matched the mechanical appearance of the robot. The voice also sounded mature and adult-like which we thought matched the age of participants that we expected since we recruited from the university. A US English accent voice was chosen because the study is taking place in a Canadian university and there was no suitable voice with Canadian English that was found. Matching the age of the robot's voice to the participants' expected age range was important as the robot would assume either a tutor or peer-learner role, both of which require a voice that is not perceived

as juvenile (details about the conditions will be discussed in Section 5.2.4). The robot used the same voice in both conditions. A pre-determined set of dialogue was written and generated, and the generated dialogue was saved as audio files and played back as part of the web application.

Designing the robot’s dialogue that distinguished between robot roles involved using different tones of voice and different pronouns. The full dialogue can be seen in the demonstration videos included in the supplementary files. The tutor robot role uses an informative tone of voice, which is instructive and more formal, and its dialogue involves the delivery of lesson content, guidance on how to complete the practice problems, and statements of encouragement after the participant answers each problem. The peer-learner robot role uses a casual tone of voice and its dialogue involves making comments during the lesson, sharing its thoughts or inquiring the participant what the answer should be for practice problems, and making statements of encouragement or acknowledgment that it was wrong. The peer-learner robot’s comments during practice problems include suggesting the correct answer, suggesting the wrong answer, or saying it does not know what the answer is.

5.2.4 Learning Activity - Esperanto Lesson

Language Selection and About

To create as much uniformity between participants proficiency of the learning subject at the start of experiment, the language Esperanto was chosen, as there is very little use and adoption of it (the number of Esperanto speakers in the world is estimated at around 100,000), and thus highly likely that participants have never heard of or have had any experience learning this language.

Esperanto is artificial language constructed in 1887 by L.L. Zamenhof, a Polish oculist, and intended for use as an international second language. Its words were derived from roots commonly found in the European languages, particularly in the Romance languages. The source languages that Esperanto vocabulary was derived from includes Dutch, German, French, and Italian [33].

Learning Content

The lesson plan (which is the experiment) consisted of two different components, a vocabulary teaching component and a multiple-choice practice problem component. The lesson was divided into sub-sections based on topic, where each sub-section consisted of teaching

vocabulary words of a certain topic, followed by practice problems for the words that had just been taught. The vocabulary teaching component had the following structure: an introduction of a word in Esperanto, an explanation in English of its definition and origins, an example pronunciation of the word in Esperanto, and finally a prompt for the participant to pronounce the word. The multiple choice practice problem consisted of questions with three choices as potential answers to choose from. There were two types of multiple choice questions, one where the answer choices showed different words, and the second type where the answer choices played short audio clips with the pronunciation of the word.

The lesson was split into 4 topics: 1) greetings, 2) objects and locations, 3) emotions, and 4) activities. A total of 12 vocabulary words were taught. Table 5.2 below illustrates the categories and words taught.

The Esperanto lesson content was displayed to the participant using a touchscreen tablet. In the tutor robot condition, the tablet only showed the lesson content because the robot was teaching, whereas in the peer-learner condition, the tablet delivered the teaching portion of the lesson since the robot was a peer-learner. Consequently, the tablet also had a voice, which was generated using David voice from Microsoft Speech API [124]. The David voice is a US English voice and is semi-robotic and adult-like, it was chosen because these qualities were suitable for instructing the Esperanto lesson with the expected age range of the participants. Figure 5.3 shows the structure of the Esperanto lesson for both conditions.

Vocabulary Category	Esperanto Vocabulary	English meaning
Greeting	saluton adiaŭ bonan matenon	hello good-bye good-morning
Objects and Locations	aŭtomobilo biciklo domo gimnazio	car bicycle house gym
Emotions	ĝoja malĝoja	happy sad
Activities	studi manĝi dormu	study eat sleep
Prefix	mal-	opposite of

Table 5.2: Vocabulary taught in Esperanto lesson

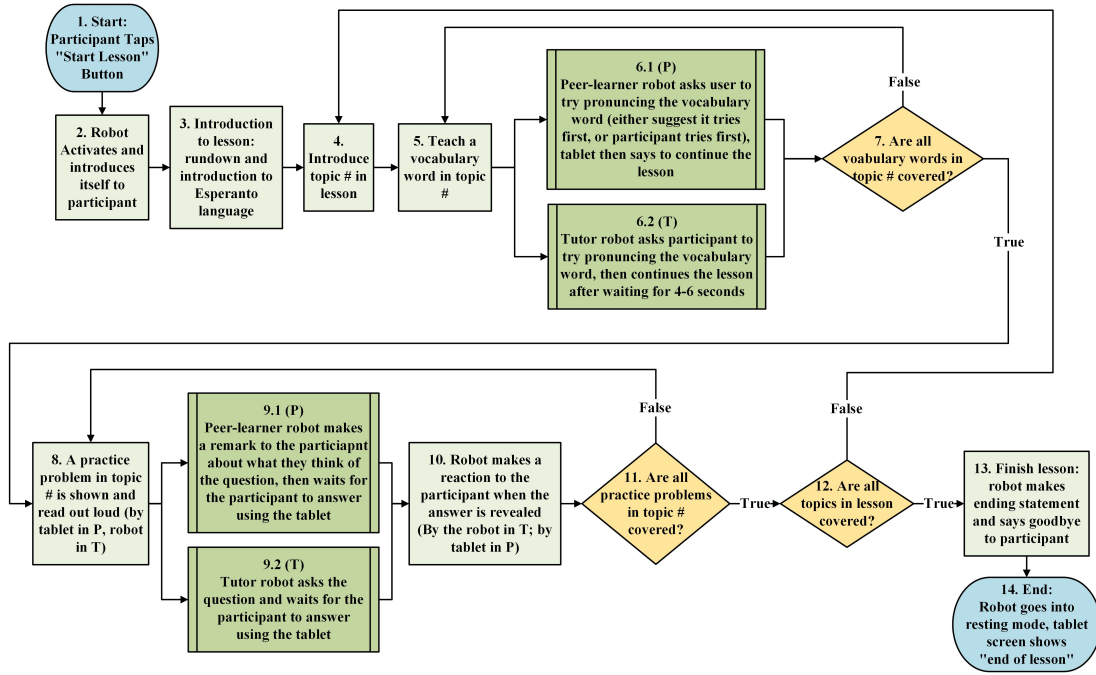


Figure 5.3: Esperanto Lesson Flow Chart

Learning Interface

The learning interface contains all learning content for the Esperanto lesson and is shown on a separate touch-screen device from the robot in the form of a webpage application. The web application uses React and JSX language for the application, and ROS libraries to synchronize and communicate with the robot actuators through ROS bridge server.

The purpose of the learning interface is to display Esperanto vocabulary words with graphics to represent each word, and for participants to complete practice problems with. Thus the interface is designed to take touch inputs. It indicates where it can take input from the participant by an explicit statement on the screen "Click on one of the blue boxes to submit an answer." (see subfigure (a) in Figure 5.4 or "To hear the choices, click on one of the blue boxes, to submit an answer, click on one of the orange boxes." (see subfigure (b) in Figure 5.4), and verbally communicating this information via either the robot (in the tutor condition) or via both tablet and robot (in the peer-learner condition) on different occasions. The interface acknowledges that it received touch input by changing the color of the buttons.

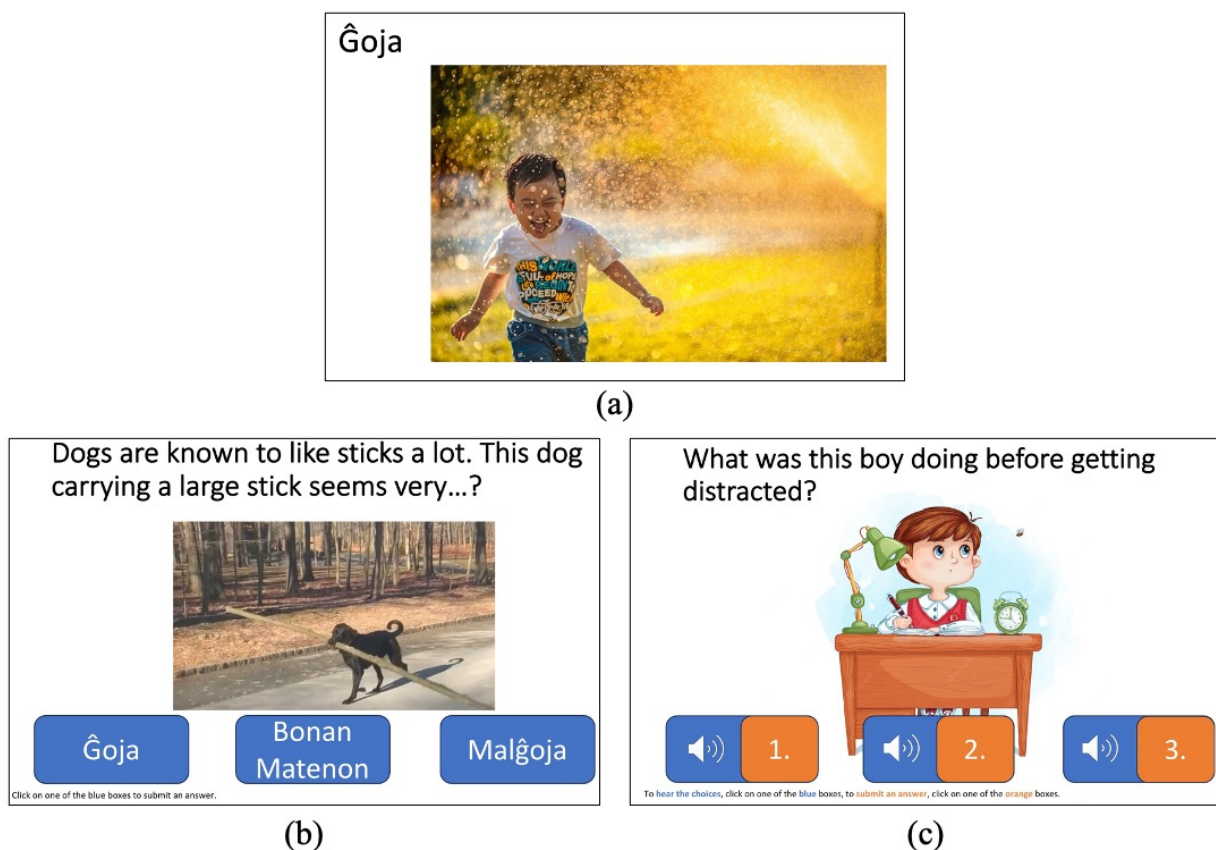


Figure 5.4: Learning Interface of Esperanto Lesson: (a) Learning interface when teaching a word during the Esperanto lesson. The Esperanto word to be learned is goja (happy), displayed above a picture illustrating ‘being happy’. (b) Word recognition practice problem during Esperanto Lesson. (c) Pronunciation/Listening practice problem during the Esperanto lesson

Figure 5.4 is an example of how a vocabulary word is presented on screen when it is being taught (by either the robot or tablet itself, depending on the condition), where the screen shows the Esperanto word being taught and a graphic representing the word. Audio dialogue accompanied the visuals on screen so it was decided to keep the visuals simple and not include other text or additional graphics. Subfigures (a) and (b) in Figure 5.4 show examples of a practice problem with buttons showing different possible answers and also for the user to tap to submit their answer. At the end of the lesson, an ending page

is displayed to visually indicate to the participant that the experiment has ended. The demonstration videos show how users use the interface.

Experiment Conditions

The experiment consists of two different conditions, the tutor robot condition (T) and the peer learner robot condition (P). Figures 5.5a and 5.5b show the experiment setups for T and P conditions respectively. Given that in each condition the robot assumed a different role, the conditions differed in the following aspects: 1) Who/what delivered the lesson to the participant, 2) Occasions where the robot interacted with the participant, 3) How the robot introduced itself at the start of the experiment 4) How the robot said goodbye to the participant. Robot's behavior differed according to the mentioned differences.

In the tutor condition, the robot delivered the lesson and interacted with the participant on occasions when teaching or instructions were given, the behavior of the robot was to convey and suit the tutor role.

In the peer-learner condition, the tablet (that also functioned as the learning interface) delivered the lesson, and the robot interacted with the participant at practice components in the lesson.

Videos of the tutor robot's behavior and interaction with participants can be found in the supplementary material.



(a) Tutor Robot Condition Setup



(b) Peer-Learner Robot Condition Setup

Figure 5.5: Setups of Robot and Tablet in the two experimental conditions.

Tutor Robot Condition

In the Tutor Robot condition (T) the robot assumes the role of a tutor and delivers the lesson to the participant while the tablet shows the lesson material, as shown in Figure 5.5a. The robot introduces itself indicates its tutoring role and uses dialogue that teaches and guides the participant on the lesson content. The placement of the robot is behind and elevated higher than the tablet to mimic a typical classroom environment where the teacher is usually in front of the student.

Peer Learner Robot Condition

In the Peer Learner Robot condition (P) the robot assumes the role of a peer learner who is also learning, alongside the participant, as a novice. Figure 5.5b shows the experiment setup where the tablet in this condition is showing the lesson material and is also delivering the lesson (in condition T, the robot is delivering the lesson). The tablet speaks with a unique voice (different from the voice of the robot), to help participants differentiate between the two voices. When the robot introduced itself at the start of the experiment, it explained its role to the participant by stating that it was also attending the lesson. Throughout the lesson, the robot initiated interaction by turning toward the participant and directly addressing the participant during the pronunciation of vocabulary words and practice problem activities. During pronunciation practice, the peer-learner robot is pre-programmed to either: **1)** suggest the participant try pronouncing first, wait for 4-6 seconds (depending on the length of the vocabulary word), and then say that it is now its turn, and then proceeds to pronounce the word, or **2)** tell the participant that it will try first, proceeds to pronounce the word, then tells the participant it is their turn, and then waits for 4-6 seconds (depending on length of the vocabulary word). Afterwards, the lesson will automatically continue.

5.3 Measures

Both quantitative and qualitative measures are collected in this study through online forms using the Qualtrics platform [94] which the participants completed on the ‘survey table’ with a computer, shown in Figure 5.1. This section describes the measures and metrics used, which were administered using online forms before and after the experiment.

5.3.1 Pre-experiment measures: Participant Demographics

Participant demographics data was collected including age, gender, occupation (student, faculty or staff), department (of study or work), language familiarity, and personality traits. For age and gender, a text box was provided for participants to complete. For the occupation, department, and language familiarity fields, multiple choice options were provided for participants to pick from and also included an 'other' choice for them to write down an alternative option that was not listed. The language familiarity options included languages that were borrowed and used in the creation of Esperanto (i.e., source languages including Dutch, French, German, and Italian) and other languages which will be categorized by its language family [81] in post-processing and data analysis. To measure personality traits, the Ten-Item Personality Inventory (TIPI) is used for measuring the Big-Five Personality Traits of participants [52]. The demographic data was collected to investigate if the two participant groups in the two experimental conditions were comparable.

5.3.2 Post-experiment measures: Participant Experience and Performance Assessment

Section 5.3.3 provides a summary of all the measures used post-experiment, and Section 5.3.4 provides details of each measure.

5.3.3 Summary of Measures Used

Standardized surveys were used to measure participant's experience. The Robot Social Attribute Scale (RoSAS) [37] to obtain three attributes (competence, warmth, and discomfort); the User Experience Questionnaire (UEQ) [67, 107] was administered, which is commonly used to measure the User Experience of interactive products where the resulting scales of the questionnaire cover pragmatic (goal-directed) or hedonic (non goal-directed) aspects of the experience. These aspects are represented by six scales rated on a scale of -3 to +3: Attractiveness, Perspicuity, Efficiency, Dependability, Stimulation, and Novelty; the System Usability Scale (SUS) [35] was administered to measure the usability aspects of our learning system developed for this study. Based on the literature, a SUS score above 68 would be considered above average and anything below 68 is below average [35]. Custom open-ended questions requesting written feedback was also administered to participants.

5.3.4 Details of Measures Used

TUPI Survey

The TIPI is a commonly used in social robotics and HRI research [22] and serves as a very brief measure of the Big-Five personality domains when there are time constraints, it is a self-reported survey that asks the participant to rate themselves on the level they agree with ten statements applied to themselves [52]. Figure 5.6 shows the TIPI survey in its entirety.

Here are a number of personality traits that may or may not apply to you. Please write a number next to each statement to indicate the extent to which *you* agree or disagree with *that* statement. You should rate the extent to which the pair of traits applies to you, even if one characteristic applies more strongly than the other.

Disagree strongly	Disagree moderately	Disagree a little	Neither agree nor disagree	Agree a little	Agree moderately	Agree strongly
1	2	3	4	5	6	7
<i>I see myself as:</i>						
1. Extraverted, enthusiastic.						
2. Critical, quarrelsome.						
3. Dependable, self-disciplined.						
4. Anxious, easily upset.						
5. Open to new experiences, complex.						
6. Reserved, quiet.						
7. Sympathetic, warm.						
8. Disorganized, careless.						
9. Calm, emotionally stable.						
10. Conventional, uncreative.						

Figure 5.6: Ten-Item Personality Measure Items

UEQ Questionnaire

The UEQ questionnaire [67, 107] shown in Figure 5.7 is used to measure the user experience of interactive products in terms of six scales: Attractiveness, Perspicuity, Efficiency, Dependability, Stimulation, and Novelty which describe either pragmatic (goal-directed) or hedonic (non goal-directed) aspects. The UEQ questionnaire consists of 26 items, each item is a pair of opposing attributes, where the user picks which attribute their experience aligns more with, and the degree of alignment (neutral alignment is also possible). Each pair of opposing attributes corresponds to one of six scales and are mapped to the six scales in post processing and analysis. The UEQ scales are scored on a range of -3 (most negative) to +3 (most positive). Thresholds to interpret the standard deviations of UEQ

scales was derived from the analysis of a large sample of UEQ studies. Three thresholds were defined, where high agreement corresponds to a standard deviation of scale below 0.83, medium agreement corresponds to a standard deviation of scale between 0.83 and 1.01, and low agreement: standard deviation of scale above 1.01 [106].

Participants are given the prompt: "For the assessment of the experience you just had with the robot, please fill out the following questionnaire. The questionnaire consists of pairs of contrasting attributes that may apply to the experience you just had with the robot. The circles between the attributes represent gradations between the opposites. You can express your agreement with the attributes by ticking the circle that most closely reflects your impression."

	1	2	3	4	5	6	7		
annoying	☐	☐	☐	☐	☐	☐	☐	enjoyable	1
not understandable	☐	☐	☐	☐	☐	☐	☐	understandable	2
creative	☐	☐	☐	☐	☐	☐	☐	dull	3
easy to learn	☐	☐	☐	☐	☐	☐	☐	difficult to learn	4
valuable	☐	☐	☐	☐	☐	☐	☐	inferior	5
boring	☐	☐	☐	☐	☐	☐	☐	exciting	6
not interesting	☐	☐	☐	☐	☐	☐	☐	interesting	7
unpredictable	☐	☐	☐	☐	☐	☐	☐	predictable	8
fast	☐	☐	☐	☐	☐	☐	☐	slow	9
inventive	☐	☐	☐	☐	☐	☐	☐	conventional	10
obstructive	☐	☐	☐	☐	☐	☐	☐	supportive	11
good	☐	☐	☐	☐	☐	☐	☐	bad	12
complicated	☐	☐	☐	☐	☐	☐	☐	easy	13
unlikable	☐	☐	☐	☐	☐	☐	☐	pleasing	14
usual	☐	☐	☐	☐	☐	☐	☐	leading edge	15
unpleasant	☐	☐	☐	☐	☐	☐	☐	pleasant	16
secure	☐	☐	☐	☐	☐	☐	☐	not secure	17
motivating	☐	☐	☐	☐	☐	☐	☐	demotivating	18
meets expectations	☐	☐	☐	☐	☐	☐	☐	does not meet expectations	19
inefficient	☐	☐	☐	☐	☐	☐	☐	efficient	20
clear	☐	☐	☐	☐	☐	☐	☐	confusing	21
impractical	☐	☐	☐	☐	☐	☐	☐	practical	22
organized	☐	☐	☐	☐	☐	☐	☐	cluttered	23
attractive	☐	☐	☐	☐	☐	☐	☐	unattractive	24
friendly	☐	☐	☐	☐	☐	☐	☐	unfriendly	25
conservative	☐	☐	☐	☐	☐	☐	☐	innovative	26

Figure 5.7: User Experience Questionnaire (UEQ) Survey Items

Descriptions of each scale are as follows, in each description, the word product is a general term and in the case of this research refers to the overall system participants used in the experiment (Maple robot and learning system). The attractiveness scale measures users' overall impression of the product (i.e., Do users like or dislike it?). The items belonging to attractiveness scale are: annoying / enjoyable, good / bad, unlikable / pleasing, unpleasant / pleasant, attractive / unattractive, friendly / unfriendly.

The perspicuity scale measures how easy it was for the user to get familiar with the product and learn how to use it. The items belonging to perspicuity scale are: not understandable / understandable, easy to learn / difficult to learn, complicated / easy, clear / confusing.

The efficiency scale measures the degree to which the user was able solve their tasks with the product without unnecessary effort. The items belonging to efficiency scale are: unpredictable / predictable, obstructive / supportive, secure / not secure, meets expectations / does not meet expectations.

The dependability scale measures the degree in which the user feel in control of the interaction? Items belonging to dependability scale are: unpredictable / predictable, obstructive / supportive, secure / not secure, meets expectations / does not meet expectations

The stimulation scale measures how exciting and motivating it was for the user to use the product? The items belonging to stimulation items: valuable / inferior, boring / exiting, not interesting / interesting, motivating / demotivating.

The novelty scales measure how innovative and creative the user thinks the product is. Items belonging to novelty scale are: creative / dull, inventive / conventional, usual / leading edge, conservative / innovative.

RoSAS Questionnaire

The Robot Social Attribute Scales (RoSAS) questionnaire, created by Carpinella et al. in 2017 [37], measures factors which influence the public's perception of social robots, which are "warmth," "competence," and "discomfort" also referred to as the social attributes of the robot. The items which constitute each RoSAS scale are listed in Figure 5.8. Participants evaluated the Maple robot on its social attributes after completing the Esperanto lesson by rating each item on a 7-point Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree). The questionnaire is presented as asking the participants this question "Using the scale provided, how closely are the following attributes associated with the robot in this study? 1 = not at all, 7 = very much so." with the items listed.

Quiz - Testing Recall and Recognition

After participants completed the post-experiment surveys, a quiz that combines recall and recognition tests is administered to objectively assess learning outcomes in participants.

Competence	Warmth	Discomfort
Reliable	Organic	Awkward
Competent	Sociable	Scary
Knowledgeable	Emotional	Strange
Interactive	Compassionate	Awful
Responsive	Happy	Dangerous
Capable	Feeling	Aggressive

Figure 5.8: Robot Social Attribute (RoSAS) Scale Survey Items

Recall involves producing an answer from memory (such as fill-in-the-blank), and recognition involves identifying the correct answer from a set of options or identifying correct statements (such as multiple-choice questions and true/false questions) [36]. Recall tests are often considered more challenging and are thought to encourage deeper learning and understanding, while recognition tests are more frequently used in standardized testing environments due to their objective nature and ease of scoring [36]. Three types of questions were used in the quiz: 1) fill in the blank, where the participant needs to fill the word or pre-fix being asked for in Esperanto; 2) multiple choice, where the correct Esperanto word must be chosen, and 3) True/False regarding a statement given. The quiz questions are shown in Figure 5.3.

Open-Ended Questions

Two open-ended questions were asked,

1. What were some things regarding anything in this study that you took notice of? It can be positive, neutral or negative.
2. What could be done to improve this experience with the robot?

The objective of asking these questions is to gain insight into participants' perceptions and expectations of their experience with the robot, which can not be captured in numeric ratings collected from standardized surveys.

Q1	----- means hello in Esperanto.
Q2	----- means goodbye in Esperanto. Choices: Ĥdieu / Adiau / Adiaŭ / Ĥdeau
Q3	"Mi amas -----", means "I love to sleep" in Esperanto. Choices: Dormio / Dormu / Domus / Domo
Q4	----- means "house" in Esperanto. Choices: Dormio / Dormu / Domus / Domo
Q5	"Mi volas -----", means "I want to eat" in Esperanto. Choices: Mangiारे / Eten / Manger / Mangi
Q6	----- means "car" in Esperanto. Choices: aŭtomobilo / automobile / automaton / aŭtobuso
Q7	ĝoja means sad in Esperanto. Choices: True / False
Q8	----- means "study" in Esperanto.
Q9	----- means "gym" in Esperanto. Choices: Gymnasium / Gymnase / Gimnasio / Gimnazio
Q10	----- means "bicycle" in Esperanto.
Q11	----- means "good morning" in Esperanto.
Q12	----- is a prefix used to indicate the opposite of a word in Esperanto.

Table 5.3: Esperanto Vocabulary Quiz

5.4 Data Analysis Methods

Prior to statistical analysis, post-processing of raw numerical data collected from surveys was done according to procedures specified by each standardized survey. The resulting numerical scales for each survey are summarized in Table 5.4.

Source of Numerical Data	Resulting Scale
TIPI survey	1 (low in trait) to 5 (high in trait)
RoSAS survey	1 (not at all in attribute) to 7 (very much so in attribute)
UEQ survey	-3 (most negative) to +3 (most positive)
SUS survey	0 to 100 (>68 is above average)
Esperanto Quiz	0 (zero marks) to 12 (full marks) points

Table 5.4: Resulting scales for numerical data obtained from post-processing

5.4.1 Statistical Analysis

Anderson-Darling test [10] was conducted to check that the data collected is normally distributed. After checking for normal distribution, two-sample t-Tests [79] are conducted on the data.

If no significant departure from normal distribution was found in the data (i.e., p values are greater than 0.05, which fails to reject the null hypothesis and concludes that the data is normal) [10], this means a t-Test can be used [128] to check for uniformity in participant demographics across all conditions, and to compare results from measures and assessments collected in the experiment.

The type of t-Test conducted is the two-sample t-Test [79] with p value of 0.05 (5% significance level) which is appropriate for this experiment design of two conditions. In addition, the two-sample t-Test was chosen for its greater robustness than the Pearson system over an extensive practical range of distributions, and it was found that for an equal-sample size, the two-sample t-Test is quite robust with respect to departures from normality [93]. Posten’s finding justifies the use the two-sample t-test as it decreases concerns about the normality of data of the measures that were used in this study.

5.4.2 Thematic Analysis

Participants’ written answers to the two-open ended questions were analyzed using the six-step Thematic Analysis (TA) method developed by Braun and Clarke (2006) [29]. Answers to both questions were analyzed together in the same TA because the questions are interrelated, and participants’ answers often informed and/or referenced each other.

A fine-grained analyses approach was taken to define a single data unit for coding, this means participant responses were segmented into individual lines which represents one data unit instead of treating the whole response from a participant as one data unit [88]. Codes (labels assigned to data units of text) were then generated from the segmented data. A theme was accepted if at least 20 codes related to it across all conditions. Exceptions to this rule were made if the number of related codes was close to 20, or the theme is significantly related to the research questions and hypotheses of this study [29].

After creating a coding frame, evaluating the intercoder reliability (ICR) of a coding frame is frequently recommended as good practice in qualitative analysis. ICR is a numerical measure of the agreement between different coders regarding how the same data should be coded and is appropriate when categorizing data at a nominal level (e.g., the presence or absence of an emotion) [88]. ICR is recommended as it promotes consistency

and transparency of the coding process, and provides confidence that specific efforts were made to ensure the final analytic framework represents a credible account of the data [88].

Evaluating the ICR first requires a second coder who will conduct the coding process of TA and create codes for 20% of all data, which is then used to evaluate the ICR. The second coder is usually another researcher who is independent from the research team to minimize potential bias in coding. Afterwards, the ICR can be calculated using one of the existing statistical tests developed. [88]. In this research, 20% of the data was taken by generating 10 random numbers from 1 to 50, each number corresponding to the response of an individual participant, the second coder was another researcher in our research group who was not involved with this research other than as an independent second coder.

Statistical tests developed for evaluating ICR currently include Cohen’s kappa [75], Krippendorff’s alpha [65], Scott’s pi [109], Fleiss’ K [48], Analysis of Variance binary ICC [96], and the Kuder-Richardson 20 [103]. This work uses the Cohen’s kappa to quantify ICR as it is the most commonly used and should be used when there are 2 coders. A Cohen’s kappa value of 0 or less indicates no agreement, 0.01–0.20 as none to slight, 0.21–0.40 as fair, 0.41–0.60 as moderate, 0.61–0.80 as substantial, and 0.81–1.00 as almost perfect agreement [75]. A Cohen’s kappa indicating almost perfect agreement between coders is the aim.

5.5 Results

secStudyResults A total of 50 participants were included in the study after exclusions due to technical issues (2 participants), or participants having had prior experience with Esperanto (2 participants). Participants were recruited from the university. Participants were assigned to conditions to balance for male/female gender. There was 1 participant who identified as non-binary. There were 24 participants in the tutor robot condition and 26 in the peer-learner robot condition.

Statistical results are reported to the nearest tenth as it is recommended to report one decimal more than the measured data [53]. Statistics reported are either in the form of mean and standard deviation (sd) in round brackets, or median and inter-quartile range (IQR) in round brackets.

5.5.1 Quantitative Results

Participant Demographics

Participant background demographics (gender, age, occupation, home faculty (of study or work), and language familiarity (with either the source languages of Esperanto or other languages) are shown in Table 5.5, Table 5.6, and Figures 5.9, and 5.10. For age, the mean and standard deviation is reported. For gender, counts for each gender category are reported. For occupation, counts for undergraduate, graduate, staff and faculty were reported. For department, counts for all six departments in the university were reported, with a seventh category labeled 'N/A' for participants who did not belong to any faculty. The six faculties are Applied Health Sciences (AHS), Arts (ARTS), Engineering (ENG), Environment (ENV), Mathematics (MATH), and Science (SCI).

Language familiarity results were separated into language familiarity with source languages (SL) of Esperanto (i.e., Dutch, French, German, and Italian) and familiarity with other languages (OL) (separated by language group). Figure 5.9 shows the distribution of participants' familiarity based on either OL only, both SL and OL, or SL only. Figure 5.10 shows the distribution of participants' familiarity with each source language. Participants can have familiarity with multiple source languages so the sum of participant counts for each condition in Figure 5.10 can exceed those in Figure 5.9.

The mean age in years for condition T was 23.3 (sd= 5.0) and for condition P was 22.6 (sd=2.8). The occupation and faculty distributions show predominately students (46 total - 24 Undergraduate students, 22 Graduate students), and mostly being from the faculties of Arts (11) or Engineering (20). The gender distribution shows more female participants than other genders (32 Female, 17 Male, 1 Other).

Participants in the tutor condition had greater familiarity with the source languages of Esperanto (i.e., Dutch, French, German, and Italian) in comparison to participants in the peer condition. For the peer condition, participants' familiarity with other languages was higher in comparison to the tutor condition.

Participant Personality Traits

Figure 5.11 shows the self-reported Big Five personality traits of participants in each condition. The ratings for each trait is as follows: Agreeableness (Tutor: Median=4.5 (IQR=1.3), Peer: Median=5.0 (IQR=2.0)), Conscientiousness (Tutor: Median=5.0 (IQR=2.3), Peer: Median=4.5 (IQR=1.5)), Emotional Stability (Tutor: Median=5.0 (IQR=1.5), Peer:

Demographic Characteristic	Tutor Robot Condition						
Age	mean			sd			
	23.3			5			
Gender	Male		Female		Non-binary		
	7		17		0		
Occupation	Undergraduate		Graduate		Staff		Faculty
	7		16		1		0
Department	AHS	ARTS	ENG	ENV	MATH	SCI	N/A
	0	9	11	0	1	1	2

Table 5.5: Tutor Robot Condition Demographic Characteristics of Participants

Demographic Characteristic	Peer-learner Robot Condition						
Age (years)	mean			sd			
	22.6			2.8			
Gender	Male		Female		Non-binary		
	9		16		1		
Occupation	Undergraduate		Graduate		Staff		Faculty
	18		4		4		0
Department	AHS	ARTS	ENG	ENV	MATH	SCI	N/A
	0	4	11	1	2	4	4

Table 5.6: Peer-Learner Robot Condition Demographic Characteristics of Participants

Median=4.5 (IQR=1.5)), Extraversion (Tutor: Median=4.3 (IQR=1.8), Peer: Median=4.0 (IQR=1.0)) Openness to New Experiences (Tutor: Median=5.0 (IQR=1.5), Peer: Median=4.5 (IQR=1.5)). A statistically significant difference was found in the Conscientiousness trait of participants between the two conditions ($p=0.0384$) where participants in the peer condition rated themselves lower than participants in the tutor condition. No significant difference between the two conditions was found for all other personality traits.

RoSAS Social Attribute Scales

Figure 5.12 compares the participant’s perception towards the Maple robot between the two conditions in terms of the three social robot attributes Competence, Discomfort, and Warmth. The ratings are as follows, Competence (Tutor: Median=5.6 (IQR=0.7), Peer:

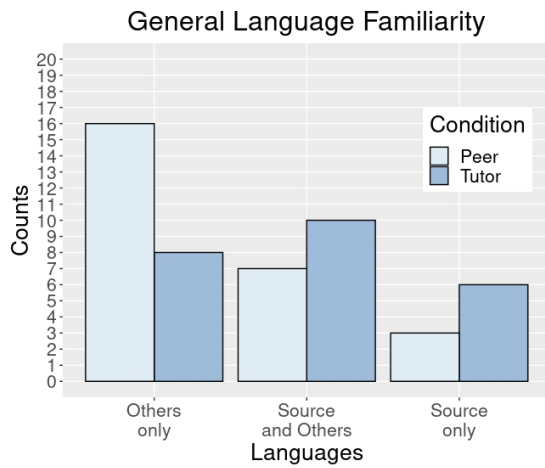


Figure 5.9:
General Language Familiarity of
Participants (Counts = Number of People)

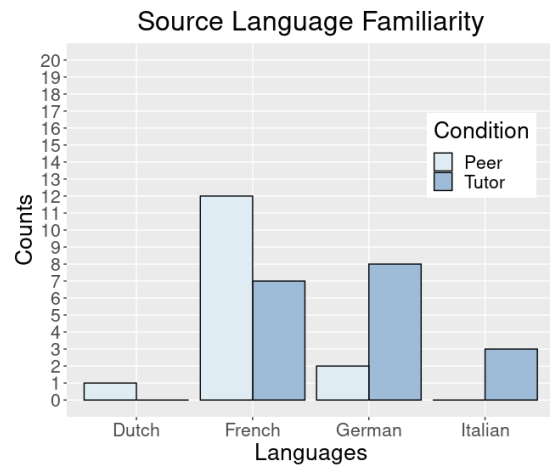


Figure 5.10:
Source Language Familiarity of
Participants (Counts = Number of People)

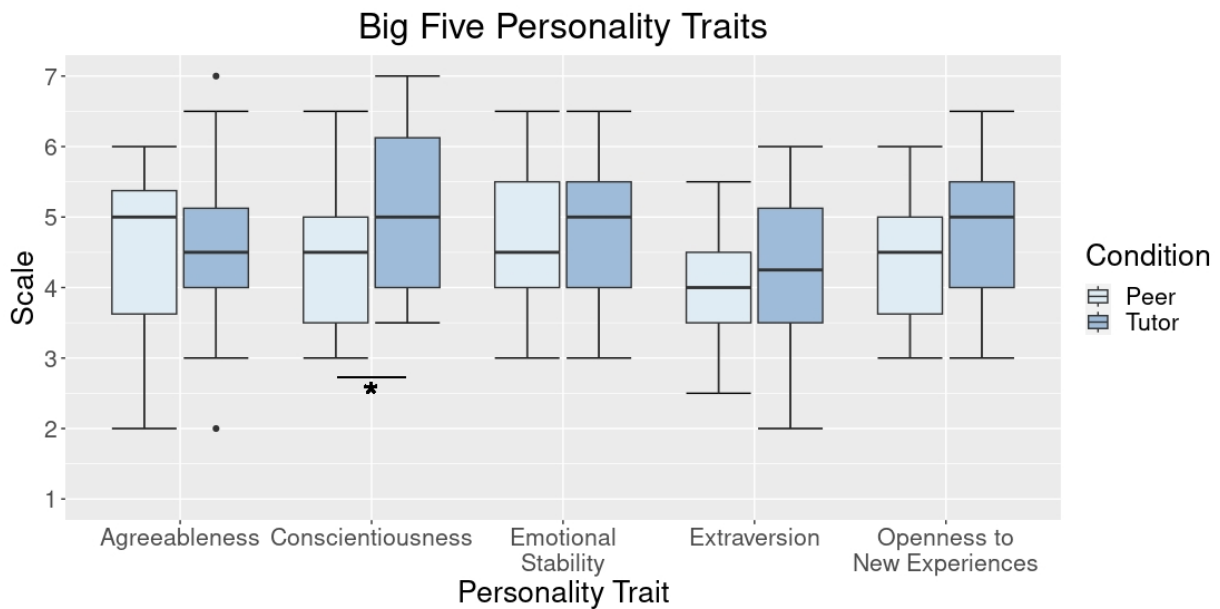


Figure 5.11: Big Five Personality Traits of Participants Comparing Experimental Conditions, with * representing $p \leq 0.05$. Dots represent outliers.

Median=4.5 (IQR=2.5)), Discomfort (Tutor: Median=1.8 (IQR=0.7), Peer: Median=2.0 (IQR=1.1)), Warmth (Tutor: Median=5.1 (IQR=1.8), Peer: Median=4.7 (IQR=1.5)).

The Competence attribute of the tutor robot was rated significantly higher ($p=0.00086$) than the peer-learner robot, where the tutor robot received a high score and the peer-learner robot received a moderate score. No significant difference was found in the Discomfort attribute between the tutor and peer-learner robot, where both conditions received low scores. There was also no significant difference found in the Warmth attribute between the tutor and peer-learner robot, where both robot roles received moderately high scores.

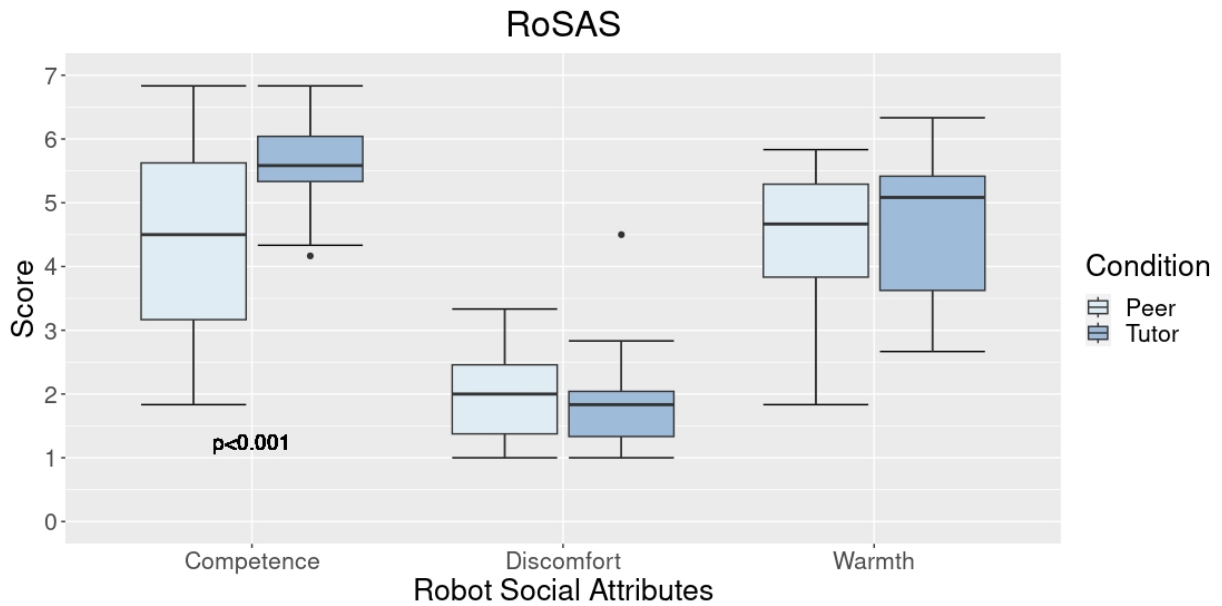


Figure 5.12: Social attributes of the Maple robot for each condition. *** indicates $p \leq 0.001$

UEQ Scales

Figure 5.13 compares the participant's user experience of the overall system (the robot and the learning interface) for both conditions in terms of six scales attractiveness, dependability, efficiency, novelty, perspicuity, and stimulation. The ratings for each scale are summarized in Table 5.7. No significant differences in UEQ scales between the T and P condition were found, however the p-value for the efficiency scale was 0.0551, very close to the threshold value of $p \leq 0.05$.

UEQ Scale	Tutor		Peer	
	Median	IQR	Median	IQR
Attractiveness	2.1	1.2	1.9	1.0
Dependability	1.9	1.0	1.8	0.9
Efficiency*	1.3	1.0	0.5	2.1
Novelty	1.4	1.0	0.8	1.7
Perspicuity	2.3	1.0	2.3	0.9
Stimulation	1.5	0.6	1.1	1.8

*Efficiency $p=0.0551$ almost ≤ 0.05

Table 5.7: Median and IQR (Inter-Quartile Range) Results of UEQ Scales for System

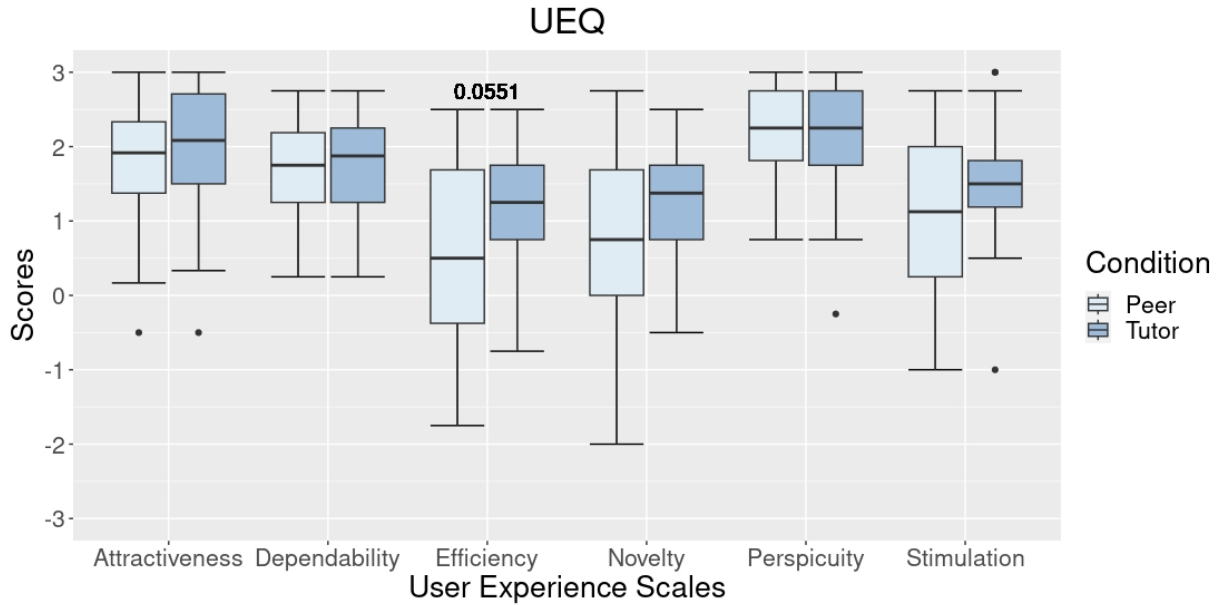


Figure 5.13: Boxplots of UEQ Scales for System

SUS Usability Score

Figure 5.14 shows the usability scores given by participants to the whole learning system (robot and tablet). No significant difference was found in the scores between the two conditions. For the tutor robot condition, the average score is $M=83.5$ ($SD=13.7$) and Median=85 (IQR=19.3). For the peer-learner robot condition, the average score is $M=78.9$ ($SD=13.1$) and Median=82 (IQR=17.5). For both conditions, the average, median, and

25% quartile ranges for the SUS scores are above the average benchmark of 68. This shows that in general, participants found the system in both conditions to have above average usability.

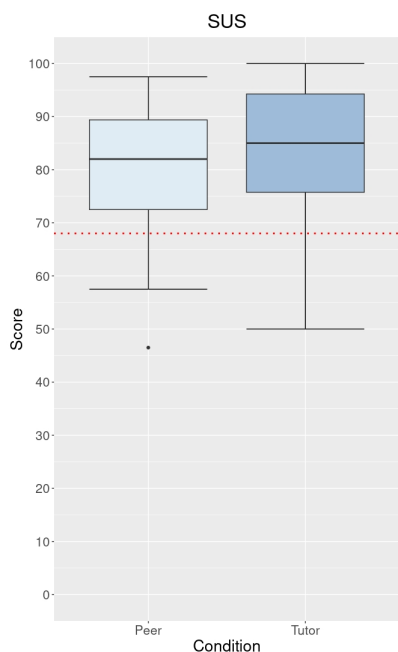
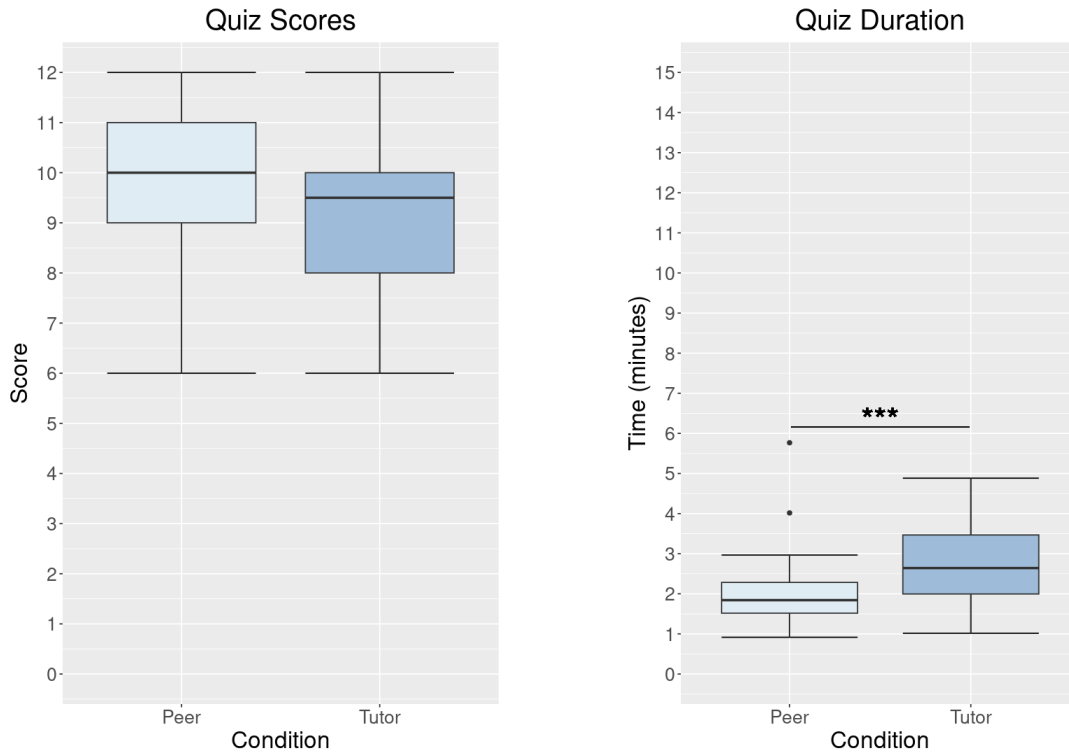


Figure 5.14: Usability Scores of Robot in Each Condition Rated by Participants

Quiz Results

Figure 5.15a and 5.15b show participants' total quiz scores and quiz duration (i.e., time taken to complete quiz) for each condition. While no statistically significant difference was found between scores across conditions, the performance of participants in the peer-learner robot condition was better than that of the tutor robot condition where the median quiz scores were 0.5 points higher, and the mean score was 0.3 points higher. A significant difference of $p \leq 0.001$ was found in the time participants took to complete the quiz, where participants in the peer condition took significantly less time where comparison of median test duration showed 30% faster completion times (24% faster times for the mean).



(a) Quiz scores.
Peer: Median=10 (IQR=2),
Tutor: Median=9.5 (IQR=2),
No significant difference found.

(b) Quiz duration (15 minute limit).
Peer: Median=1.84 (IQR=0.8),
Tutor: Median=2.64 (IQR=1.5),
*** indicates $p \leq 0.001$.

Figure 5.15: (a) Quiz scores and (b) Quiz completion time for each condition

5.5.2 Qualitative Results

In total 50 sets of written feedback were acquired for each open-ended question, each response for a question was treated as one data unit and codes were developed and associated to each response when the content of the response was related to a code. Therefore, one data unit (response) can have multiple codes and themes associated to it.

ICR Agreement

The Intercode Reliability (ICR) [88] was evaluated using the Cohen's kappa (as stated in Section 5.4.2). The observed linear weighted Kappa value is 0.87 (0.8655), indicating there

is almost perfect agreement between coders as it is greater than 0.8. This means that any future interpretation in the thematic analysis done by the first coder is objective and would match with the interpretation of the other independent coder (2nd coder).

Thematic Analysis Results

Four themes were identified from participants responses to the two open-ended questions. Table 5.8 summarizes the themes identified in each condition of the experiment. The following sections detail the findings on each theme.

Themes	Subthemes	Tutor Robot (T)		Peer-learner Robot (P)		All conditions	
1. Face, Movement and Voice of Robot	1.1 Facial Expressions	5	21%	3	12%	8	16%
	1.2 Movement and Gestures	8	33%	5	19%	13	26%
	1.3 Voice and Speech	5	21%	9	35%	14	28%
	Total number/percentage of participants whose response relates to face, movements and voice of robot	18	75%	17	65%	35	70%
2. Interaction with Robot	2.1 Perception of Robot Behavior	8	33%	12	46%	20	40%
	2.2 Robot Awareness, Capability, and Effectiveness	3	13%	5	19%	8	16%
	2.3 Desire for More Interaction Opportunities	8	33%	12	46%	20	40%
	2.4 Engagement with Robot	6	25%	11	42%	17	34%
	Total number/percentage of participants whose response relates to interaction with robot	25	104%	40	154%	65	130%
3. Overall System/Technology	3.1 Technology Acceptance	10	42%	5	19%	15	30%
	3.2 Robot Role and Contribution	1	4%	6	23%	7	14%
	Total number/percentage of participants whose response relates to overall system/technology	11	46%	11	42%	22	44%
4. Esperanto Lesson	4.1 Pronunciation Discrepancy and Inaccuracy	6	25%	4	15%	10	20%
	4.2 Learning Experience	8	33%	2	8%	10	20%
	4.3 Add Customization and Expand Scope	5	21%	2	8%	7	14%
	Total number/percentage of participants whose response relates to Esperanto lesson	19	79%	8	31%	27	54%

Table 5.8: Summary of Themes Identified in Participant Response to Open-Ended Questions

Theme 1 Face, Movement and Voice of Robot

55 participants made comments regarding their impression of the facial expressions, movements and gestures, and speaking voice of the robot.

Subtheme 1.1 Face of Robot

12 participants commented on the robot's face. Positive comments expressed finding the facial expressions of the robot pleasant and welcoming. Negative comments expressed finding the robot's eye blinking being slow, and that facial expressions are sometimes repetitive which creates discomfort.

"My first thought was how cute and approachable to robot's 'face' looks once you start interacting with it. (P)", "The repeated face gestures of the robot was engaging but also a bit creepy. (T)".

Subtheme 1.2 Movements of Robot

26 participants commented on the robot movements. Positive comments included movements being human-like, expressive and friendly, with certain gestures of the robot making participants feel encouraged and motivated. Negative comments included some robot movements being rigid or slow, unexpected, excessive, and/or distracting. A few commented on too much noise from motors while robot was moving.

"The hand gesture and movement of head of robot is good, it makes this more interactive and motivating. (T)", "too much movement in short spans can be unnerving at times. (P)"

Subtheme 1.3 Voice of Robot

14 participants commented on the robot's voice and speaking. 3 participants indicated they liked the robot voice, and 7 participants found the voice too robotic and would prefer a more human-like voice, and 5 participants' comments suggest increasing talking pace.

"voice of the robot was really soft and friendly which made me trust him/her more (T)", "I think the voice of the robot was very robotic, if it can be replaced with a more natural voice then it might become more interactive. (T)".

Theme 2 Interaction with Robot

71 participant responses had comments related to interaction with the robot. This theme describes perceptions of the interaction participants had with the robot. 5 subthemes were identified.

Subtheme 2.1 Perception of Robot Behavior

18 participants commented on their interaction with the robot leaving an impression or emotional impact, where 10 and 12 out of 18 participant responses held positive and negative connotations, respectively. Positive comments included . Negatives included feeling

interaction was inorganic, and decreasing willingness to pay attention to robot.

“I noticed that the robot gave a lot of “human” sort of reactions, such as it responded to me “okay”, it “talked” in a pretty natural way as well which helped me feel more comfortable. (P)”, “I thought the robot’s speech was a little slow ... but I thought it was a little slow as it did it (facial expressions) ... I noticed I was starting to get impatient waiting for the robot to complete the gesture (T)”.

Subtheme 2.2 Robot Awareness, Capability, and Effectiveness

8 participants made comments related to participants’ impression of robot’s self-awareness, capabilities and effectiveness relating to the Esperanto lesson.

“The most prominent feature about the robot was its awareness of when the presentation was finished speaking and when to speak. It also spoke quite clearly ... Although the robot was engaging when it spoke, I felt that it often did not provide much support when answering questions ... It seemed to struggle to accurately recall words and I learned to mostly ignore its suggestions when answering questions. (P)”, “The positive things I noticed about the study was the interactive nature of learning a new language with the robot. The quiz paired with a somewhat human-like learning experience with the robot made my learning more engaging. (T)”.

Subtheme 2.3 Desire for More Interaction Opportunities

20 participants expressed wanting more interaction in more situations and/or suggesting changes to increase the quality of interaction with the robot.

“That would be nice if I could shake the robot’s hand, I wanted to do this every time it was moving. (T)”, “(add) more sociable components, Maple as a companion encouraged learning, but only seems as valuable when the user would prefer not to learn alone (P)”

Subtheme 2.4 Engagement with Robot

17 participants made comments related to how much they engaged with the robot. Participants were not given the research area’s definitions for terminology such as engagement and interactive. These words, when used in participant responses, were interpreted by authors as referring to the same thing (engagement), and thus the authors assumed the use of words such as engagement and interactive to mean the same. Out of the 17 participants, 11 participants expressed interaction with the robot was engaging, enjoyable, and human-like, while responses from 6 participants indicated there was less attention paid to the robot or lack of engagement between the robot and participant.

”The positive things I noticed about the study was the interactive nature of learning a new language with the robot. The quiz paired with a somewhat human-like learning experience with the robot made my learning more engaging. (T)”, ”The facial expressions and movements were very engaging as it made it feel like speaking to an emotional being

rather than a computer. I felt that the robot was more than just code and motors - it certainly felt closer to speaking to a human than I expected. (P)”

Theme 3 Overall System/Technology

Subtheme 3.1 Technology Acceptance

15 participants expressed a positive experience and attitudes of satisfaction and acceptance towards the technology (robot & tablet learning system) being used for learning.

“I have a positive experience. The robot is a great invention for studying a new language. It looks creative (T)”, “It was fun to have someone learning the language alongside me. (P)”.

Subtheme 3.2 Robot Role and Contribution

7 participants’ responses included comments that expressed less attention paid to the robot, feeling the robot does not support learning, and that modifying the robot’s behavior to provide more guidance would have made it more useful learning-wise.

“if we can somehow integrate the screen with robot so that there is only one way to interact with robot, that would be better. As the focus for most of the time remained on the screen only, rather than the robot. (T)”, “In my opinion, the robot in its current state does not add much to the learning process and just serves more as a companion than a helper. (P)”.

Theme 4 Esperanto Lesson

Subtheme 4.1 Pronunciation Discrepancy and Inaccuracy

19 participants commented on the pronunciation discrepancy between the tutor robot teaching and when the pronunciation audio played, and the inaccuracy of the peer-learner robot’s pronunciation to the pronunciation audio played. Responses suggest providing pronunciation of vocabulary in the robot’s voice (from T) and making the robot pronounce the words better (from P) would improve the learning experience.

“The negative things I noticed was the inconsistency in the robots voice when saying the words in another language (T)”, “I took notice of the fact that the robot speaks the Esperanto words with an English accent. Since my native language is a romance language, I think my pronunciation in Esperanto was better than his in general. This made it a bit boring for me, but maybe for a native English speaker this would be relatable instead.

(P)”.

Subtheme 4.2 Learning Experience

11 participants commented on the pace, difficulty, interface, and learning components of the Esperanto lesson. Responses indicate most feel the lesson pace is too slow, there is split opinion on difficulty level between too easy and appropriate, and some find instructions confusing when doing practice problems.

“I found the questions in the lesson to be somewhat too easy and it made the lesson feel slow/boring as it went on. (P)”, “It went at a pace that was slow and easy to follow, though it could be because the words were easier and close to languages I already know. (P)”.

Subtheme 4.3 Add Customization and Expand Scope

8 participants suggested adding customization options such as difficulty level, captions, ability to repeat things, and adding more learning components such as spelling, writing, listening practice and feedback for user performance.

“implementing different methods of learning a word, spelling, listening & selecting (P)”, “It would be interesting to get feedback from a robot about my pronunciation.(T)”.

5.6 Discussion

5.6.1 Revisiting Hypotheses

H1 Participants will have higher quiz scores when robot assumes peer learner role

The analysis of the data shows no significant difference in the quiz scores between conditions which means H1 is not supported.

A few reasons could explain why no statistical significance difference was found. Firstly, the scope of the Esperanto lesson was limited to a set of beginner level vocabulary words, and the difficulty level of the Esperanto lesson was either at a comfortable level or too easy according to thematic analysis of participant responses (see Section 5.5.2). This is also reflected in the high quiz scores across both conditions (see Figure 5.15a). Thematic analysis also identified that some participants felt that in the lesson the tablet was sufficient and the robot was not needed for knowledge acquisition but rather served better in terms of companionship and encouragement (see Section 5.5.2). Thus, it can be argued that due to these reasons, learning results would not be significantly different depending

on the role the robot assumed. Secondly, the practice problems were of the same type as the majority of the quiz questions (multiple choice). Thus, participants had more practice in recognizing the correct word during the lesson and were prepared for that portion of the quiz, resulting in high scores. Thirdly, in the quiz, 8 out of 12 questions were multiple choice, which tested recognition, and 4 of 12 questions were fill-in-the-blank questions, which tested recall. Recognition type questions are easier than recall type questions since recognition memory requires less brain activity (i.e., less cognitive load) than recall [36]. Since the majority of quiz questions tested recognition, the cognitive load was relatively light.

H2 Participants will take less time to complete the quiz when the robot assumes peer learner role

A significant difference was found in the amount of time taken by participants to complete the quiz, whereby participants in the peer-learner robot condition completed the quiz 30 percent faster than those in the tutor robot condition 5.15b. Thus, H2 is accepted. This might be explained as follows.

Since quiz scores were very similar in both conditions, faster quiz completion times in condition P could indicate more meaningful learning occurred in participants who learned with a peer-learner robot. Meaningful learning is defined as "the acquisition of new meanings. New meanings, conversely, are the end-products of meaningful learning. That is, the emergence of new meanings in the learner reflects the prior operation and completion of a meaningful learning process" [20], which occurs when the learner is "actively and willfully trying to achieve a cognitive goal, they think and learn more because they are fulfilling an intention." [126]. The cognitive goal may not be initially expressed by the learner, but it must be accepted and adopted by the learner in order for learning to be meaningful, this is because when learners evaluate their learning in terms of their intentions, they understand more and are better able to use the knowledge that they have constructed in new situations" [20]. From this definition, the intention and cognitive goal in the participant potentially was cultivated by the peer-learner robot's remarks during the practice problem part of the lesson (either suggesting the correct answer, saying it does not know, or suggesting the wrong answer), and may have had the effect of prompting the participant to further critically reflect on the question and create a cognitive goal to actively retrieve learned knowledge. This can be seen in the following quotes from written responses of participants in the peer-learner condition, "I liked that the robot tried to lead you to the wrong answer on occasion. Keeps you engaged and critically thinking. (P)", and "The robot tries to fool you sometimes (P)". Active retrieval based learning was argued to enhance meaningful learning [59]. This may have resulted in a deeper consolidation of

learned knowledge, making them more assured of their answer. This might explain why it took participants in the peer-learner condition significantly less time to complete the quiz with the same scores while obtaining similar scores to participants in the tutor condition.

H3 Participants will perceive the robot in a peer learner role more positively (e.g., higher warmth, lower discomfort, higher competence)

From the calculated RoSAS scales, no significant difference was found in Warmth and Discomfort towards the robot between conditions, however, the Competence of the robot was rated significantly lower in the peer condition than in the tutor condition. This means that based on our data H3 is not supported.

H4 Robot assuming peer learner role will yield higher ratings in usability and user experience from participants

No significant differences were found in the SUS scores and UEQ scales between the two conditions. This means that in this study H4 cannot be supported.

5.6.2 Implications of Robot’s Role

The statistical results provided a solid baseline understanding of the system’s usability, the robots’ perceived social attributes, and the overall learning effectiveness. However, the thematic analysis added depth, revealing specific areas for improvement of the system and nuanced perceptions that vary among participants. This variance could have been influenced by a combination of demographic characteristics (such as age, occupation, and faculty) and personality traits, which shape individuals’ preferences, expectations, and perceptions. For example, participants’ age, field of study or work, and personalities could influence how they perceive and prefer different robot roles, forms of interactions, and learning structures in a language learning setting.

Regarding age and occupation, nearly all participants were students either in undergraduate or graduate studies. Being a student might imply that participants are more accustomed to a certain format of learning.

In terms of the distribution of participants across faculties, with Engineering and Arts being represented more than other faculties in the tutor condition, and a more diverse faculty representation in the peer-learner condition, could mean that participants’ technological proficiency, exposure, and learning preferences could have influenced their perception of the robot’s competence and the system’s usability.

The above-average SUS scores in both conditions suggest good usability, but the the-

matic analysis provided insight that participants desired more natural interactions and expressed specific preferences for robot movements and voice. This suggests a gap between overall usability satisfaction and specific interaction preferences, possibly influenced by individual personality traits or demographic factors. For example, participants from technology-oriented faculties might have higher expectations for robot movements and voice, affecting their perception of the system’s efficiency and novelty.

The tutor robot’s higher competence social attribute ratings from the RoSAS align with comments on its authoritative and knowledgeable presence from thematic analysis. In contrast, the peer-learner robot’s moderate competence score might reflect its role as a companion rather than an authority, as indicated by qualitative feedback.

5.7 Conclusion

In conclusion, this study found that while there was little to no difference in standardized metrics for participants’ self-reported perception towards the robot, participants who learned with a peer-learner robot were able to complete the knowledge assessment (Esperanto quiz) 30% faster than participants who learned with a tutor robot. This result might indicate that the robot assuming a peer-learner role promoted more meaningful learning in participants. This finding is important because it informs the direction of future development for the Maple robot and for RALL with adult learners. This study also found from participants’ written feedback that understanding the complex interplay between demographic factors, personality traits, and participant feedback is vital for refining human-robot interaction in educational settings. It is crucial not only to assess whether the robots and systems are generally usable or effective, but how different types of users perceive and interact with these technologies based on their unique characteristics and individual preferences. This understanding can inform more personalized and adaptive educational robots, leading to more effective and satisfying learning experiences.

Chapter 6

Thesis Conclusion

6.1 Summary of Contributions

The research makes a significant contribution to the field of HRI by developing software for robot hardware, which implements new functionality into an existing robot and thus creating a new social robot. This robot software is designed to be versatile and adaptable, making it useful for a wide range of projects in HRI. By developing this new robot platform, the research opens up possibilities for addressing important social needs like providing companionship, supporting education, and assisting in various interactive settings. Additionally, the platform serves as a valuable tool for researchers, giving them a standardized framework to conduct experiments and studies in HRI.

The research culminated in a paper submitted to the International Journal of Social Robotics (SORO). Furthermore, positive response to the developed robot platform was evident from informal showcases to elementary school students as well as a user study that involved university students and staff (see Section 5). Feedback from these interactions showed that people are interested in the Maple robot and see potential in this new robot platform for learning applications. This feedback confirms Maple robot’s usefulness for future HRI projects and applications. Overall, the research not only advanced knowledge in HRI but also laid groundwork for practical applications that can benefit society, which may include applications in language learning and more learning contexts.

6.2 Follow-up with Local Immigration Services

A meeting with the team at Reception House of Waterloo region [15] was scheduled (later in Winter 2024 term) to present the Maple robot, its features, and results of the "Language Crash Course with a Robot" user study see 5 for the purpose of continuing discussions on the interaction design with immigrant children (the original target population of this research, see Section 1). Part of the plan was to open a discussion to include the professional at outreach and immigrant families in user-centered design with expert advice such that there can be continued development of the Maple robot for HRI targeted for the original research goal (see Section 1). Initial communications with the Waterloo Region Immigration Services and Reception Housing before the meeting were positive and they showed significant interest and enthusiasm towards this research project.

On March 6th 2024, we successfully met with the team at Reception House and discussed with professionals / experts who work with immigrant families at the Waterloo Region Reception House about the biggest challenges the immigrants they work with encounter and to brainstorm how we can collaborate to apply our research project and Maple robot to try to help immigrant families with their challenges. The outreach team was shown demonstration videos shown to them of the Maple robot in each condition (tutor robot condition, peer-learner robot condition) of our "Language Crash Course with a Robot" user study (see Section 5) and was asked to answer a set of questions (see Section 6.2) to understand whether the Maple robot that was developed and the learning system is seen as a potentially helpful tool for immigrant families that they work with. The set of question are as follows:

1. What are things you liked or disliked (or have comments) about each condition in terms of teaching a language?
2. Do you think such a system, if adapted to teaching English, could be useful for immigrant children?
3. What ages of immigrant children do you think this robot and learning system would be suitable for? Can speak on varying levels of suitability (which ages most suitable, which ones maybe not, and reasoning behind that).
4. Which condition (i.e., robot role) would you think is better for use with immigrant children?

5. A considerable amount of university student participant feedback found the lesson too easy. Given that the participants were all at least university students or older/more experienced, do you think this difficulty level is good for children? Or is it also too easy and it should start with higher difficulty than this?
6. From participant feedback, there were comments on wanting the robot to 1) interact even more with the participant and react more to the participant's behavior, and 2) integrate more learning components such as practice spelling, feedback on pronunciation, and more grammar. Do you think this would be something good and which do you think is more important?
7. There was also participant feedback that the robot's behavior and the lesson pace are too slow, or some movements are repetitive and can be boring (sometimes off putting) do you think immigrant children would also feel that way? Why or why not?
8. What features do you think should be added to the current system that would really be useful for children (that was not observed or suggested by participant feedback mentioned earlier)?

The professionals'/experts' at Reception House responses to our questions and feedback for the Maple robot platform was as follows:

In response to the questions about the language teaching system and its adaptation for immigrant children (Questions 1 and 2), Reception House expresses positive views about the visual engagement provided by the robot. They believe that the robot's engaging and friendly appearance could effectively hold the attention of children, particularly those who may have faced challenges in traditional learning structures due to their background as refugees. They also see potential benefits in using the robot to teach not only language but also behavioral skills, particularly for children who may speak different native languages. They suggest that incorporating such content could help support the children's learning of English while also focusing on important social and behavioral aspects.

In response to the question about the suitable ages for the robot and learning system (Question 3), Reception House suggests that it could appeal to a broad range of ages, possibly from early childhood to early teens. They mention that younger children, particularly those aged two to six, would likely find the system very attractive. However, they also believe that older children, possibly up to early teens, could still find it engaging and cool to interact with. Reception House emphasizes the importance of developing educational

content tailored to the needs of these children, particularly focusing on behavioral aspects such as sharing and respect, which they see as crucial for their educational development.

In response to the question about which robot role would be better for use with immigrant children (Question 4), Reception House expresses a preference for the peer-learner role over the tutor role for the robot. They believe that the peer-learner role, which Reception House perceived as being characterized by soft tones and relatable language, is more likely to be mimicked by children, as the children can see it as relatable to themselves. Reception House points out that small inherent differences between the peer-learner and tutor role behavior (e.g. the use of more polite language in the peer-learner role), could impact children's behavior while interacting with the robot. They also highlight the significant challenges that many immigrant children, especially refugees, face due to the traumas they have experienced. These challenges can manifest in difficulties with emotional regulation and adapting to formal social structures, potentially affecting their interactions with the robot.

In response to the question regarding the difficulty level of the lessons for immigrant children compared to the feedback from university student participants (Question 5), Reception House offers insights based on their observations. They suggest that the difficulty level of the lessons may vary depending on the age and background of the children. Younger children, particularly those in temporary accommodation, may pick up basic vocabulary very quickly, especially if they are frequently exposed to common words like "hello" and "goodbye." However, as children grow older, they may find more common phrases challenging, although still manageable. Regarding the practice problems in the Esperanto lesson, the respondent highlights potential challenges for children with limited schooling, particularly in reading comprehension. They emphasize the importance of adapting the content and user interface to accommodate varying literacy levels among immigrant children, suggesting that while they may be proficient in spoken language, their written language skills could be significantly different. This also circles back to the idea of the robot speaking the immigrant's native language to facilitate understanding of the interface, indicating a potential solution to address language barriers and enhance accessibility. In addition to discussing the difficulty level of the lessons, the respondent brings up the social integration challenges faced by refugee immigrant children, emphasizing that they often struggle more with social behaviors than academic aspects. They point out specific social behaviors that refugee children may find difficult to adapt to or learn, such as emotional regulation and identifying their feelings when experiencing strong emotions. This includes learning how to differentiate between different emotions and how to appropriately express and manage them. Another challenge mentioned is adapting to social norms and expectations in Canadian schools, particularly in terms of rough play versus zero-tolerance policies on fighting

and bullying. Refugee children may need to relearn social boundaries and norms, which can be a significant challenge in their adjustment process.

In response to the question about the interaction level with children using the language teaching system (Questions 6 and 7), Reception draws from their experience in running children’s programming in temporary accommodation. While not an expert on the topic, they suggest that a peer model with more movement and behavior mimicry would likely be more beneficial and engaging for children. They emphasize the importance of focusing on behavior rather than just language learning, suggesting that children may be more likely to mimic the robot’s behavior if it reflects excitement or engagement. Additionally, the respondent highlights the challenges faced by parents in adjusting their parenting styles to Canadian norms, suggesting that the robot’s behavior could serve as a positive model for both children and parents.

In response to the question about features to be added to the current system for children (Question 8), Reception House highlights the struggles faced by parents in adjusting their parenting styles to Canadian norms, particularly due to cultural differences. They note that children often perceive the world differently from their parents, leading to conflicts in behavior expectations. Reception House suggests that robots could be agents of change, could play a role in facilitating communication between children and parents. They emphasize the importance of addressing these cultural and generational differences to support children’s mental well-being. Additionally, the respondent emphasizes the need for the system to foster a balance between traditional values and the expectations of Canadian society, advocating for a more holistic approach to addressing the challenges faced by immigrant families.

Overall, it was found that the Reception House works with refugee immigrant families and are often traumatized and distrustful of government systems due to their past experiences in their native country. The main challenge that refugee immigrant families face is with social integration, such as problems when adapting to the new environment’s social norms, moral values, and finding a compromise between their mindset from their native culture and the mindset of the culture they immigrated to. This finding is in contrary to the main challenges and concerns that the general immigrant population face which are receiving quality education and financial stability, these concerns are actually not commonly mentioned by the refugee immigrant families that Reception House works with. This is an interesting finding as existing literature on RAL for teaching social skills focuses on populations of children with special needs such as Autism, however from the discussion with Reception House the need to learn social skills proves to also be needed in other populations. This discovery sparks future research possibilities, exploring how social robots can help humans learn about and understand diverse cultural social etiquette and values.

6.3 Limitations

One of the limitations of this work is the study only collected results based on short-term interactions which fails to capture the potential long-term effects on learning and the evolution of interaction and relationship between the human and the robot/system over time, which are crucial for understanding sustained engagement and educational impact. Our results may also be prone to novelty effect.

Another limitation is the reliance on self-reported measures for assessing personality traits and user experience which introduces the possibility of bias, e.g. participants might respond based on perceived social expectations rather than their actual experiences and thoughts. While the study provides a rich qualitative analysis, its quantitative data might not adequately reflect the complex ways individual differences influence learning with robots. These measurement tools and methods, therefore, may not provide a complete picture of the interaction dynamics between humans and robots.

In its current state of development of Maple the social robot, technological constraints significantly limit the study’s findings. The robot behavior was automated with a set of predetermined behaviors, so that the robot could operate autonomously alongside the tablet and in interaction with the user. These behaviors were either sequenced or triggered by basic conditions such as the participant’s performance on answers to practice problems. Lacking the ability to process speech or visual input reliably, the robot offered less spontaneous, natural, and personalized interactions. This restriction limited opportunities for participants to engage meaningfully with the robot and prevented them from initiating interactions, however, the results reveal features and improvements that should be prioritized in the future development of the robot.

The generalizability of the study’s findings is another limitation. Conducted in the context of learning Esperanto, a relatively uncommon language, and given the low difficulty of the lesson content, the results might not be applicable to more complex and challenging learning tasks. Moreover, the controlled and lab-based environment of the study may not reflect real-world educational settings where factors such as group dynamics and environmental distractions can significantly influence the learning process.

Lastly, the study does not sufficiently consider participants’ learning preferences such as learning environments they are most accustomed to and/or prefer, study habits, and modes of learning they have experienced. This information could reveal psychological and behavioral aspects that could influence the learning experience and expectations about the robot’s role, behavior, and learning structure, as well as how these might affect their interaction and learning outcomes. Due to only having short-term interactions as mentioned

earlier, the study also fails to observe how user perceptions and robot effectiveness might evolve as users become more accustomed to learning with robots over time, a factor that could significantly alter the long-term effectiveness of robotic tutors.

6.4 Literature on Using Robots or Digital Technologies to Assist Immigrants and Refugees

Table 6.1 and Table 6.2 show a summary of publications on using robots (digital technologies instead of robots in [45]) to assist immigrants and/or refugees in learning or transition. From the summary, it can be observed that the background, age range, and situation of the participants varied, where [42] focused on teaching unaccompanied immigrant children gestures, [18] focused on university students, [45] focused on older immigrants/refugees, and [64] focused on immigrants in Finland, [73] focused on male refugees from Syria, and [72] focused on Spanish speaking English language learners (ELLs). The work [16] was a workshop that focused on social robots for social and cultural inclusion, and [45] was a review that focused on digital technologies instead of social robots for use with immigrants.

The results of those works that conducted a user study all found positive results in the users using the robot, such as positive impact on learning outcomes, successful engagement of participant with robot, positive user experience and attitude towards the robot. The limitations found in all studies was the limited number of participants and lack of investigation into using their implementation on different populations (such as different age groups or background of immigrants), thus there is limited generalizability for their findings. Another common limitation found was that all studies were early stage research which meant the results do not represent a comprehensive investigation and thus findings may change once future work is conducted. A general limitation identified in this niche area of research is that there is limited existing work and the studies conducted all are very different, each study focused on a different subset of immigrants, different subject for immigrants to learn, and very different study designs and objectives, which makes it difficult to evaluate each work against each other, as well as evaluate the work in this thesis to them. However, the participatory design approach taken in [72] with the elementary school students and professionals provides guidance for our future work (described later in 6.5).

Despite the limitations, the findings indicate a promising application for social robots to help immigrants and refugees with language and social integration, and are encouraging for continuing the research direction described in this thesis.

Paper Title	Purpose	Target Population
Social Robots supporting the Inclusion of Unaccompanied Migrant Children: Teaching the Meaning of Culture-Related Gestures [42]	Explore how social robots can tutor second language and support cultural integration	Unaccompanied migrant children in a new country dealing with language and cultural barriers
Goes to the Heart: Speaking the User's Native Language [18]	Develop a social robot to support non-native speakers in navigating local services	International students, refugees, and asylum seekers facing language and bureaucratic challenges in a new country
SOcial and Cultural IntegrAtion with Personalized Interfaces (SOCIALIZE) 2022 [16]	Foster social and cultural inclusion using interactive techniques and social robots	Vulnerable and at-risk categories, including individuals from diverse nationalities and cultural backgrounds
Understanding the Impact of Digital Technology on the Well-being of Older Immigrants and Refugees [45]	Assess how digital technology affects the well-being of older immigrants and refugees	Older immigrants and refugees, facing challenges like language barriers, cultural adaptation, and isolation
Customized Robot-Assisted Language Learning (RALL) for Immigrants in the Workplace [64]	Support immigrants' work-related language learning using robots	Immigrants in Finland, specifically those needing to learn the local language for work and integration
Social Robots as a Means of Integration? An Explorative Acceptance Study considering Gender and Non-verbal Behaviour [73]	Investigate the acceptance of social robots teaching culture-specific behaviors to refugees	Male Syrian refugees, who have experienced forced displacement and are adapting to a new cultural and social environment
Designing for culturally responsive social robots: An application of a participatory framework [72]	Gather culturally responsive requirements for child-robot interactions. Engage ELLs in design decisions for social robots.	Spanish speaking ELLs elementary school children from South American countries

Table 6.1: Summary of Literature on Learning Support for Immigrants and Refugees - Purpose, Target Population

Paper Title	Design	Results	Limitations	Future Work
Social Robots supporting the Inclusion of Unaccompanied Migrant Children: Teaching the Meaning of Culture-Related Gestures [42]	Qualitative user study	Positive impact on learning gestures and engaging children	Limited number of participants and categories of subjects	Expand the experiment to more participants
Goes to the Heart: Speaking the User's Native Language [18]	Field user study	User study pending	Not specified	Extend to support refugees and asylum seekers
SOcial and Cultural IntegrAtion with Personalized Interfaces (SOCIALIZE) 2022 [16]	Workshop overview	Insights into interactive techniques and social robots for inclusion	Early-stage research	Develop and apply in inclusive settings
Understanding the Impact of Digital Technology on the Well-being of Older Immigrants and Refugees [45]	Systematic literature review	Digital technology critical for well-being but underused by target group	Limited to English-language studies	More studies on digital technology use among older refugees
Customized Robot-Assisted Language Learning (RALL) for Immigrants in the Workplace [64]	Qualitative user study	Positive user experience and learning outcome	Limited number of participants and language proficiency levels	Comprehensive studies with diverse participants
Social Robots as a Means of Integration? An Explorative Acceptance Study considering Gender and Non-verbal Behaviour [73]	Explorative user study	Generally positive attitude towards educational robots	Culture-specific manipulations, limited to male Syrian refugees	Systematic investigation of non-verbal behavior and other measures
Designing for culturally responsive social robots: An application of a participatory framework [72]	Explorative user study	Students found the social robot engaging but hesitant to converse. Teachers suggested peer-to-peer conversations and stimulating student responsiveness.	Generalization of cultural traits in participants	Enhance participatory design, in-depth cultural consideration, integration in real educational contexts

Table 6.2: Summary of Literature on Learning Support for Immigrants and Refugees - Design, Results, Limitations, Future Work

6.5 Future Work

Future work should address the limitations discussed and address the findings from meeting with local communities.

Future work to address the limitations discussed may be done by incorporating a broader range of interaction conditions, extending the duration of interactions, employing a mix of measurement tools, enhancing technological capabilities, and exploring more diverse educational contexts and languages. Improving the robot’s interactive capabilities and adding features to allow for more nuanced and participant-initiated interactions and developing the ability to perceive and respond to a wider range of user inputs and behaviors will provide deeper insights into the potential of robots as interactive learning tools and the dynamics of user-robot relationships. Additionally, a deeper examination of psychological and behavioral factors and their long-term evolution will provide more comprehensive insights into the effectiveness and potential of learning with and alongside social robots.

Future work to address findings from local communities is to continue discussion with local communities such as Reception House with the original target population and experts who can take part in co-design and user-centered design processes. Future work in collaboration with experts and immigrant families could focus on several areas based on the insights provided by Reception House. This could include:

- Developing and testing educational content that integrates behavioral skills alongside language learning to better support immigrant children’s holistic development.
- Exploring the feasibility and effectiveness of implementing a peer-learner role for robots in educational settings, with a focus on its impact on children’s behavior and in social situations.
- Designing user interfaces and content delivery methods that accommodate varying literacy levels and social integration challenges among immigrant children, ensuring accessibility and inclusivity.
- Investigating the potential for robots to facilitate communication between children and parents from different cultural backgrounds, and exploring strategies to promote understanding and collaboration within immigrant families.
- Collaborating with educators, psychologists, and cultural experts to develop culturally responsive interventions and support systems that address the unique needs of immigrant children and their families.

Overall, future work in these areas has the potential to enhance the effectiveness of social robot for use with immigrant children by working on using social robots to aid in social integration and language acquisition, and improve mutual understanding and communication between immigrant children and immigrant parent.

References

- [1] Aldebaran - nao robot. Accessed on 2023-12-20.
- [2] Cozmo robot - digital dream labs store. Accessed on 2023-12-20.
- [3] Emo - the next generation robot by living.ai. Accessed on 2023-12-20.
- [4] Luxai - qtrobot humanoid social robot for research and teaching. Accessed on 2023-12-20.
- [5] Miro-e website. Accessed on 2023-12-20.
- [6] Misty robotics - misty ii. Accessed on 2023-12-20.
- [7] React - a javascript library for building user interfaces. Accessed on 2023-12-20.
- [8] State of manufacturing: Vector 2.0. Accessed on 2023-12-20.
- [9] Vector robot - digital dream labs store. Accessed on 2023-12-20.
- [10] *Anderson–Darling Test*, pages 12–14. Springer New York, New York, NY, 2008.
- [11] Dynamixel sdk. https://emanual.robotis.com/docs/en/software/dynamixel/dynamixel_sdk/overview/, 2023. Accessed: 2023-11-22.
- [12] Engineer kit1. <https://emanual.robotis.com/docs/en/edu/engineer/kit1/>, 2023. Accessed: 2023-11-22.
- [13] U2d2. <https://emanual.robotis.com/docs/en/parts/interface/u2d2/>, 2023. Accessed: 2023-11-22.
- [14] U2d2 power hub. https://emanual.robotis.com/docs/en/parts/interface/u2d2_power_hub/, 2023. Accessed: 2023-11-22.

- [15] Reception house. Website, 2024. Retrieved from <https://receptionhouse.ca/>.
- [16] Francesco Agrusti, Fabio Gasparetti, Cristina Gena, Giuseppe Sansonetti, and Marko Tkalcic. Social and cultural integration with personalized interfaces (socialize). In *26th International Conference on Intelligent User Interfaces - Companion*, IUI '21 Companion, page 9–11, New York, NY, USA, 2021. Association for Computing Machinery.
- [17] Living AI. Emo - desktop pet robot, 2023. Accessed: December 10, 2023.
- [18] Shaul Ashkenazi, Gabriel Skantze, Jane Stuart-Smith, and Mary Ellen Foster. Goes to the heart: Speaking the user’s native language. In *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, HRI '24, page 214–218, New York, NY, USA, 2024. Association for Computing Machinery.
- [19] Suhaib Aslam, Edwin Christian Dertien, and Jelle van Dijk. Designerly scaffoldings for co-designing social robots and human-robot interaction. 8 2020. 1st International Workshop on Designerly HRI Knowledge 2020 ; Conference date: 03-09-2020 Through 03-09-2020.
- [20] David P. Ausubel. *The Nature of Meaning and Meaningful Learning*, pages 67–100. Springer Netherlands, Dordrecht, 2000.
- [21] Bruce D. Baker. *How Money Matters for Schools*. Learning Policy Institute, Palo Alto, CA, 2017. This work is licensed under the Creative Commons Attribution—Noncommercial 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc/4.0/>.
- [22] Christoph Bartneck, Tony Belpaeme, Friederike Eyssel, Takayuki Kanda, Merel Keijsers, and Selma Šabanović. *Human-Robot Interaction: An Introduction*. Cambridge University Press, Cambridge, May 2020. <https://doi.org/10.1017/9781108676649>.
- [23] Paul Baxter, Emily Ashurst, Robin Read, James Kennedy, and Tony Belpaeme. Robot education peers in a situated primary school study: Personalisation promotes child learning. *PLOS ONE*, 12:1–23, 05 2017.
- [24] Tony Belpaeme, Paul Baxter, Robin Read, Rachel Wood, Heriberto Cuayahuitl, Bernd Kiefer, Stefania Racioppa, Ivana Kruijff-Korabayova, Georgios Athanasopoulos, V. Enescu, Rosemarijn Looije, Mark Neerincx, Yiannis Demiris, Raquel Ros, Aryel Beck, Lola Cañamero, Antoine Hiolle, Matthew Lewis, Ilaria Baroni, and Remi

- Humbert. Multimodal child-robot interaction: Building social bonds. *Journal of Human-Robot Interaction*, 1:33–53, 12 2012.
- [25] Tony Belpaeme, James Kennedy, Aditi Ramachandran, Brian Scassellati, and Fumihide Tanaka. Social robots for education: A review. *Science Robotics*, 3(21):eaat5954, 2018.
 - [26] Tony Belpaeme and Fumihide Tanaka. *Social Robots as educators*, chapter 7. OECD Publishing, 2021.
 - [27] Aude Billard and Daniel Grollman. *Human-Robot Interaction*, pages 1474–1476. Springer US, Boston, MA, 2012.
 - [28] David Boud, Ruth Cohen, and Jane Sampson. Peer learning and assessment. *Assessment Evaluation in Higher Education*, 24:413–426, 12 1999.
 - [29] Virginia Braun and Victoria Clarke. Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2):77–101, 2006.
 - [30] Cynthia Breazeal. Kismet overview. <https://web.media.mit.edu/~cynthiab/research/robots/kismet/overview/overview.html>. Accessed on: 2023-12-20.
 - [31] Cynthia Breazeal. Toward sociable robots. *Robotics and Autonomous Systems*, 42(3):167–175, 2003. Socially Interactive Robots.
 - [32] Cynthia Breazeal, Atsuo Takanishi, and Tetsunori Kobayashi. Social robots that interact with people. *Springer handbook of robotics*, pages 1349–1369, 2008.
 - [33] Encyclopædia Britannica. Esperanto, n.d. Accessed 2023-11-22.
 - [34] Elizabeth Broadbent. Interactions with robots: The truths we reveal about ourselves. *Annual Review of Psychology*, 68(1):627–652, 2017. PMID: 27648986.
 - [35] John Brooke. Sus: A quick and dirty usability scale. *Usability Eval. Ind.*, 189, 11 1995.
 - [36] Roberto Cabeza, Shitij Kapur, Fergus Craik, Anthony McIntosh, Sylvain Houle, and Endel Tulving. Functional neuroanatomy of recall and recognition: A pet study of episodic memory. *Journal of Cognitive Neuroscience*, 9:254–265, 03 1997.

- [37] Colleen M. Carpinella, Alisa B. Wyman, Michael A. Perez, and Steven J. Stroessner. The robotic social attributes scale (rosas): Development and validation. In *2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 254–262, 2017.
- [38] Xavier Caré. Innorobo 2015 - nao (cropped), 2015. Accessed: December 10, 2023.
- [39] Chatchai Chirapornchai, Paul Bremner, and Joseph E. Daly. Helper’s high with a robot pet. In *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction, HRI ’21 Companion*, page 229–233, New York, NY, USA, 2021. Association for Computing Machinery.
- [40] K. Dautenhahn. Human-robot interaction, January 1 2014.
- [41] Kerstin Dautenhahn. Methodology & themes of human-robot interaction: A growing research field. *International Journal of Advanced Robotic Systems*, 4(1):15, 2007.
- [42] Berardina De Carolis, Giuseppe Palestra, Cosimo Della Penna, Marina Cianciotta, and Anna Cervellone. Social robots supporting the inclusion of unaccompanied migrant children: Teaching the meaning of culture-related gestures. *Journal of e-Learning and Knowledge Society*, 15(2), June 2019.
- [43] DevX. C plus plus programming language. <https://www.devx.com/terms/c-plus-plus-programming-language/>, 2023. Accessed: 2023-11-22.
- [44] Yerassyl Diyas, Dmitriy Brakk, Yernar Aimambetov, and Anara Sandygulova. Evaluating peer versus teacher robot within educational scenario of programming learning. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 425–426, 2016.
- [45] Prince Chiagozie Ekoh, Tochukwu Jonathan Okolie, Fidel Bethel Nnadi, Oluwagbemiga Oyinlola, and Christine A Walsh. Understanding the impact of digital technology on the well-being of older immigrants and refugees: A scoping review. *DIGITAL HEALTH*, 9:20552076231194947, 2023.
- [46] Olov Engwall and José Lopes. Interaction and collaboration in robot-assisted language learning for adults. *Computer Assisted Language Learning*, 35(5-6):1273–1309, 2022.
- [47] David Feil-Seifer and Maja Mataric. *Human-Robot Interaction*, pages 4643–4659. 04 2009.

- [48] Joseph L. Fleiss. Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76(5):378–382, 1971.
- [49] N. M. Fraser and G. N. Gilbert. Simulating speech systems. *Computer Speech & Language*, 5(1):81–99, 1991.
- [50] Nithya J Gogtay and Urmila M Thatte. Principles of correlation analysis. *Journal of the Association of Physicians of India*, 65(3):78–81, 2017.
- [51] Michel Goossens, Frank Mittelbach, and Alexander Samarin. *The L^AT_EX Companion*. Addison-Wesley, Reading, Massachusetts, 1994.
- [52] Samuel D Gosling, Peter J Rentfrow, and William B Swann. A very brief measure of the big-five personality domains. *Journal of Research in Personality*, 37(6):504–528, 2003.
- [53] Farrokh Habibzadeh. Statistical data editing in scientific articles. *J Korean Med Sci*, 32(7):1072–1076, 2017.
- [54] Abid Haleem, Mohd Javaid, Mohd Asim Qadri, and Rajiv Suman. Understanding the role of digital technologies in education: A review. *Sustainable Operations and Computers*, 3:275–285, 2022.
- [55] Anna Henschel, Guy Laban, and Emily S. Cross. What makes a robot social? a review of social robots from science fiction to a home or hospital near you. *Current Robotics Reports*, 2(1):9–19, 2021.
- [56] Ruud Hortensius, Bishakha Chaudhury, Martin Hoffmann, and Emily Cross. Tracking human interactions with a commercially-available robot over multiple days. *Open research Europe*, 2:97, 2022. Epub ahead of print.
- [57] This is Engineering. Rabbit robot, 2019. Accessed: December 10, 2023.
- [58] Takayuki Kanda, Takayuki Hirano, Daniel Eaton, and Hiroshi Ishiguro. Interactive robots as social partners and peer tutors for children: A field trial. *Human Computer Interaction (Special issues on human-robot interaction)*, 19:61–84, 06 2004.
- [59] Jeffrey D. Karpicke and Phillip J. Grimaldi. Retrieval-based learning: A perspective for enhancing meaningful learning. *Educational Psychology Review*, 24(3):401–418, 2012.

- [60] James Kennedy, Paul Baxter, Emmanuel Senft, and Tony Belpaeme. Social robot tutoring for child second language learning. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 231–238, 2016.
- [61] Donald Knuth. *The T_EXbook*. Addison-Wesley, Reading, Massachusetts, 1986.
- [62] Elly A. Konijn and Johan F. Hoorn. Robot tutor and pupils’ educational ability: Teaching the times tables. *Computers & Education*, 157:103970, 2020.
- [63] Elly A. Konijn, Brechtje Jansen, Victoria Mondaca Bustos, Veerle L. N. F. Hobbelen, and Daniel Preciado Vanegas. Social robots for (second) language learning in (migrant) primary school children. *International Journal of Social Robotics*, 14(3):827–843, Apr 2022.
- [64] Susanna Kouri, Ellinoora Köpman, Aino Ahtinen, and Valentina Ramirez Millan. Customized robot-assisted language learning to support immigrants at work: Findings and insights from a qualitative user experience study. In *Proceedings of the 8th International Conference on Human-Agent Interaction, HAI ’20*, page 212–220, New York, NY, USA, 2020. Association for Computing Machinery.
- [65] Klaus Krippendorff. Computing krippendorff’s alpha-reliability. 2011.
- [66] Leslie Lamport. *L^AT_EX — A Document Preparation System*. Addison-Wesley, Reading, Massachusetts, second edition, 1994.
- [67] Bettina Laugwitz, Theo Held, and Martin Schrepp. Construction and evaluation of a user experience questionnaire. volume 5298, pages 63–76, 11 2008.
- [68] Meira Levinson, Tatiana Geron, and Harry Brighthouse. Conceptions of educational equity. *AERA Open*, 8:23328584221121344, 2022.
- [69] Chia-Yu Liu and Hung-Ling Chen. Effects of peer learning on learning performance, motivation, and attitude. *International Journal of Education Economics and Development*, 11, 02 2020.
- [70] A. Lockerd and Cynthia Breazeal. Tutelage and socially guided robot learning. volume 4, pages 3475 – 3480 vol.4, 01 2004.
- [71] John R. Logan and Julia Burdick-Will. School segregation and disparities in urban, suburban, and rural areas. *The Annals of the American Academy of Political and Social Science*, 674(1):199–216, November 2017. GR - P2C HD041020/HD/NICHD NIH HHS/United States, GR - R24 HD041020/HD/NICHD NIH HHS/United States.

- [72] Belinda Louie, Elin Björling, Annie Kuo, and Patrícia Alves-Oliveira. Designing for culturally responsive social robots: An application of a participatory framework. *Frontiers in Robotics and AI*, 9:1–12, 10 2022.
- [73] Birgit Lugrin, Jessica Dippold, and Kirsten Bergmann. Social robots as a means of integration? an explorative acceptance study considering gender and non-verbal behaviour. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2026–2032, 2018.
- [74] Louis Major, Gill A. Francis, and Maria Tsapali. The effectiveness of technology-supported personalised learning in low- and middle-income countries: A meta-analysis. *British Journal of Educational Technology*, 52(5):1935–1964, 2021.
- [75] Mary L. McHugh. Interrater reliability: The kappa statistic. *Biochem Med (Zagreb)*, 22(3):276–282, 2012.
- [76] Cristian Mejia and Yuya Kajikawa. Bibliometric analysis of social robotics research: Identifying research trends and knowledgebase. *Applied Sciences*, 7(12), 2017.
- [77] Kenneth W. Miller. It’s not nice to fool humans. *IT Professional*, 12(1):51–52, 2010.
- [78] MobileSyrup. Cozmo robot, 2017. Accessed: December 10, 2023.
- [79] P. G. Moore. The two-sample t-test based on range. *Biometrika*, 44(3/4):482–489, 1957.
- [80] Omar Mubin, Catherine Stevens, Suleman Shahid, Abdullah Mahmud, and Jian-Jie Dong. A review of the applicability of robots in education. *Technology for Education and Learning*, 1, 06 2013.
- [81] Ágoston Nádasdy. Language families. In E. Pásztor, J. Vajda, and F. Loew, editors, *Language and Speech*, volume 56 of *Acta Neurochirurgica Supplementum*. Springer, Vienna, 1993.
- [82] United Nations. Goal 4: Quality education.
- [83] Jeannie Oakes, Peter Cookson, Janel George, Stephanie Levin, Desiree Carver-Thomas, Fred Frelow, and Barnett Berry. *Adequate and Equitable Education in High-Poverty Schools: Barriers and Opportunities in North Carolina*. 2019.
- [84] OECD. *OECD Digital Education Outlook 2021*. 2021.

- [85] Open Source Robotics Foundation. Documentation - ros wiki. <https://wiki.ros.org/>, 2023. Accessed: 2023-11-22.
- [86] Open Source Robotics Foundation. Ros - robot operating system. <https://www.ros.org/>, 2023. Accessed: 2023-11-22.
- [87] Open Source Robotics Foundation. Ros noetic ninjemys. <https://wiki.ros.org/noetic>, 2023. Accessed: 2023-11-22.
- [88] Cliodhna O’Connor and Helene Joffe. Intercoder reliability in qualitative research: Debates and practical guidelines. *International Journal of Qualitative Methods*, 19:1609406919899220, 2020.
- [89] Matthew K. X. J. Pan, Elizabeth A. Croft, and Günter Niemeyer. Validation of the robot social attributes scale (rosas) for human-robot interaction through a human-to-robot handover use case. 2017.
- [90] Hae Won Park, Rinat Rosenberg-Kima, Maor Rosenberg, Goren Gordon, and Cynthia Breazeal. Growing growth mindset with a social robot peer. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, HRI ’17, page 137–145, New York, NY, USA, 2017. Association for Computing Machinery.
- [91] Hannah R. M. Pelikan, Mathias Broth, and Leelo Keevalik. ”are you sad, cozmo?”: How humans make sense of a home robot’s emotion displays. In *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, HRI ’20, page 461–470, New York, NY, USA, 2020. Association for Computing Machinery.
- [92] Jean Piaget. *The Equilibration of Cognitive Structures: The Central Problem of Intellectual Development*. University of Chicago Press, Chicago, 1985. Original work published 1975.
- [93] Harry O. Posten. The robustness of the two—sample t—test over the pearson system. *Journal of Statistical Computation and Simulation*, 6(3-4):295–311, 1978.
- [94] Qualtrics. Qualtrics. Available online, 2005–2024. Development Company: Qualtrics, Location: Provo, Utah, USA.
- [95] Natasha Randall. A survey of robot-assisted language learning (rall). *ACM Transactions on Human-Robot Interaction*, 9(1):Article 7, December 2019.
- [96] Martin S. Ridout, Clarice G. B. Demétrio, and David Firth. Estimating intraclass correlation for binary data. *Biometrics*, 55(1):137–148, 1999.

- [97] Laurel D. Riek. Wizard of oz studies in hri: A systematic review and new reporting guidelines. *Journal of Human-Robot Interaction*, 1(1):119–136, 2012.
- [98] Laurel D. Riek and Rachel Watson. The age of avatar realism. *IEEE Robotics & Automation Magazine*, 17(4):37–42, 2010.
- [99] ROBOTIS. Dynamixelsdk. <https://github.com/ROBOTIS-GIT/DynamixelSDK>, 2023. Accessed: 2023-11-22.
- [100] INC. ROBOTIS. Robotis engineer kit 1 [us]. <https://www.robotis.us/robotis-engineer-kit-1-us/>, 2023. Accessed: 2023-11-22.
- [101] Nicole Salomons, Kaitlynn Taylor Pineda, Adérónké Adéjàre, and Brian Scassellati. “we make a great team!”: Adults with low prior domain knowledge learn more from a peer robot than a tutor robot. In *2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 176–184, 2022.
- [102] Mauro Sarrica, Sonia Brondi, and Leopoldina Fortunati. How many facets does a “social robot” have? a review of scientific and popular definitions online. *Information Technology & People*, 33(1):1–21, 2019.
- [103] Joe L. Saupe. Some useful estimates of the kuder-richardson formula number 20 reliability coefficient. *Educational and Psychological Measurement*, 21(1):63–71, 1961.
- [104] Jeremy Sawyer, Rita Obeid, Dennis Bubnitz, Anna Schwartz, Patricia Brooks, and Aaron Richmond. Which forms of active learning are most effective: Cooperative learning, writing-to-learn, multimedia instruction, or some combination? *Scholarship of Teaching and Learning in Psychology*, 3:257–271, 12 2017.
- [105] Brian Scassellati, Jake Brawer, Katherine Tsui, Setareh Nasihati Gilani, Melissa Malzkuhn, Barbara Manini, Adam Stone, Geo Kartheiser, Arcangelo Merla, Ari Shapiro, David Traum, and Laura-Ann Petitto. Teaching language to deaf infants with a robot and a virtual human. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI ’18, page 1–13, New York, NY, USA, 2018. Association for Computing Machinery.
- [106] Martin Schrepp. User experience questionnaire handbook, 09 2015.
- [107] Martin Schrepp, Andreas Hinderks, and Jörg Thomaschewski. Applying the user experience questionnaire (ueq) in different evaluation scenarios. pages 383–392, 06 2014.

- [108] Trenton Schulz, Jim Torresen, and Jo Herstad. Animation techniques in human-robot interaction user studies: A systematic literature review. *J. Hum.-Robot Interact.*, 8(2), jun 2019.
- [109] William A Scott. Reliability of content analysis: The case of nominal scale coding. *Public opinion quarterly*, pages 321–325, 1955.
- [110] Amazon Web Services. List of supported languages and voices, n.d. Accessed 2023-11-22.
- [111] Renato Silva, Claudio Policastro, Giovana Zuliani, Ednaldo Pizzolato, Roseli Aparecida, and Francelin Romero. Concept learning by human tutelage for social robots. *Learning and Nonlinear Models*, 6, 01 2008.
- [112] Statistics Canada. Fewer babies born as canada’s fertility rate hits a record low in 2020. [Online]. Available: <https://www.statcan.gc.ca/o1/en/plus/960-fewer-babies-born-canadas-fertility-rate-hits-record-low-2020>, May 2022. Accessed 27 November 2022.
- [113] Statistics Canada. Focus on geography series, 2021 census of population. [Online]. Available: <https://www12.statcan.gc.ca/census-recensement/2021/as-sa/fogs-spg/page.cfm?topic=9&lang=E&dguid=2021A000011124>, October 2022. Accessed 27 November 2022.
- [114] Statistics Canada. Immigrants make up the largest share of the population in over 150 years and continue to shape who we are as canadians. [Online]. Available: <https://www150.statcan.gc.ca/n1/daily-quotidien/221026/dq221026a-eng.htm>, October 2022. Accessed 27 November 2022.
- [115] Statistics Canada. Immigration as a source of labour supply. [Online]. Available: <https://www150.statcan.gc.ca/n1/daily-quotidien/220622/dq220622c-eng.htm>, June 2022. Accessed 27 November 2022.
- [116] Statistics Canada. In the midst of high job vacancies and historically low unemployment, canada faces record retirements from an aging labour force: number of seniors aged 65 and older grows six times faster than children 0-14. [Online]. Available: <https://www150.statcan.gc.ca/n1/daily-quotidien/220427/dq220427a-eng.htm>, April 2022. Accessed 27 November 2022.
- [117] Statistics Canada. A portrait of citizenship in canada from the 2021 census. [Online]. Available: <https://www12.statcan.gc.ca/census-recensement/2021/>

- [as-sa/98-200-X/2021008/98-200-X2021008-eng.cfm](#), November 2022. Accessed 27 November 2022.
- [118] Fumihide Tanaka and Shizuko Matsuzoe. Children teach a care-receiving robot to promote their learning: Field experiments in a classroom for vocabulary learning. *Journal of Human-Robot Interaction*, 1, 08 2012.
 - [119] Ubuntu. Ubuntu Server 20.04. <https://ubuntu.com/blog/ubuntu-server-20-04>, Accessed 2023-12-19.
 - [120] UNESCO. The global learning crisis: Why every child deserves a quality education. Technical report, 7, place de Fontenoy, 75352 Paris 07 SP, France, 2013.
 - [121] UNICEF. Early childhood development overview. [Online]. Available: <https://data.unicef.org/topic/early-childhood-development/overview/#:~:text=Early%20childhood%2C%20which%20spans%20the,social%2C%20emotional%20and%20physical%20development>, May 2022. Accessed 27 November 2022.
 - [122] United Nations Statistics Division. The sustainable development goals report 2022. Technical report, 2022. Accessed on: April 25, 2024.
 - [123] Rianne van den Berghe, Josje Verhagen, Ora Oudgenoeg-Paz, Sanne van der Ven, and Paul Leseman. Social robots for language learning: A review. *Review of Educational Research*, 89(2):259–295, April 2019. Article reuse guidelines: sagepub.com/journals-permissions.
 - [124] Wikipedia. Microsoft text-to-speech voices, n.d. Accessed 2023-11-22.
 - [125] Jefferson William. Vector robot by anki, 2019. Accessed: December 10, 2023.
 - [126] Kam Po Wong. Facilitating a meaningful learning experience for students by multi-media teaching approach. 2015.
 - [127] Richard Mark Wood. A review on education differences in urban and rural areas. *Telfer School of Management; the University of Ottawa*, March 2023. Corresponding Author’s E-mail: WoodRM@edu.in.
 - [128] Manfei Xu, Drew Fralick, Julia Z Zheng, Bokai Wang, Xin M Tu, and Changyong Feng. The differences and similarities between two-sample t-test and paired t-test. *Shanghai Arch Psychiatry*, 29(3):184–188, 2017.

- [129] Marie Yamamoto, Yue Hu, Enrique Coronado, and Gentiane Venture. Impression evaluation of robot’s behavior when assisting human in a cooking task. In *2021 30th IEEE International Conference on Robot Human Interactive Communication (RO-MAN)*, pages 743–748, 2021.
- [130] Cristina Zaga, Manja Lohse, Khiet P. Truong, and Vanessa Evers. The effect of a robot’s social character on children’s task engagement: Peer versus tutor. In Adriana Tapus, Elisabeth André, Jean-Claude Martin, François Ferland, and Mehdi Ammi, editors, *Social Robotics*, pages 704–713, Cham, 2015. Springer International Publishing.
- [131] Tom Ziemke, Nicklas Bergfeldt, Gunnar Búason, Tarja Susi, and Henrik Svensson. Evolving cognitive scaffolding and environment adaptation: A new research direction for evolutionary robotics. *Connect. Sci.*, 16:339–350, 12 2004.
- [132] Selma Šabanović. Robots in society, society in robots. *International Journal of Social Robotics*, 2(4):439–450, Dec 2010.

APPENDICES

Appendix A

Robot Development Supplementary Documentation

A.1 Motor Properties

Joint ID	Motor ID	Motor Location (joint or movement)	Rotation Plane	Min. Limit	Max. Limit	Zero Position	Direction of Positive Orientation
Joint3	1	Right shoulder (shoulder-elbow)	x-z	1050	3100	2050	CCW, forwards/up
Joint4	2	Right shoulder (shoulder-elbow)	y-z	1600	3100	2050	CW, outwards/up
Joint7	3	Left shoulder (shoulder)	x-z	1050	3100	2050	CCW, backwards
Joint8	4	Left shoulder (shoulder-elbow)	y-z	1600	3100	2050	CW, Inwards,
Joint5	5	Right forearm, (elbow-hand)	x-y	0	4095	2050	CW, outwards
Joint6	6	Right forearm (hand)	x-z	1500	3050	2050	CW, outwards
Joint9	7	Left forearm (elbow-hand)	x-y	0	4095	2050	CW, inwards
Joint10	8	Left forearm (hand)	x-z	1500	3050	2050	CCW, outwards,
Joint12	9	Head (forward-back)	x-z	900	2050	2050	CW, backwards
Joint11	10	Head (left-right)	y-z	1550	2550	2050	CCW, to its left
Joint2	11	Torso (left-right)	x-y	0	3800*	2050	CCW, to its left
Joint1	12	Torso (forward-back)	x-z	1025	2050	2050	CW, backwards

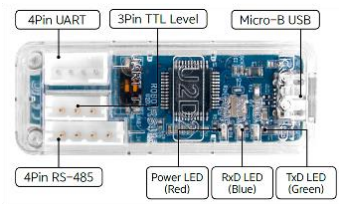
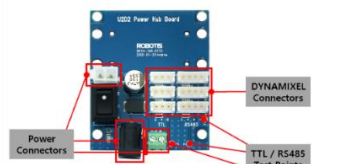
*(due to wire length, else max. is 4095)

CW = clockwise, CCW = counterclockwise

Table A.1: Motor Position Properties and Joint Mapping Summary for Maple Robot

A.2 Maple Robot SOP

MME Standard Operating Procedure (SOP)

Name	Maple	Maple robot:  Raspberry Pi:  U2D2 USB communication converter:  U2D2 power hub: 
Description	Humanoid robot with static base, 6 Dynamixel 2XL430-W250-T motors, Raspberry Pi 4B and U2D2 USB communication converter & power hub	
Manufacturer and Model	- Robotis Engineer Kit 1 – Dr.R configuration - Robotis Dynamixel Servos 2XL-W230 - Raspberry Pi 4B (aarch64) - U2D2 USB communication converter - U2D2 power hub - Robotis CM-550 controller (not in use)	
Location	E7 3431/3433	
SOP Creation Date	2023-01-24	
SOP Created By	- Victoria (Ya-Ting) Yang - y546yang@uwaterloo.ca	
SOP Revision Date		
SOP Revised By		
SOP Location	inv.mme.uwaterloo.ca - E7 3431 - A.I.R Lab Teams Equipment Channel	
Manual Location	- usually inv.mme.uwaterloo.ca - Hard copy at E7 3431/3433 - Online: https://emanual.robotis.com/docs/en/edu/engineer/kit1/	
Equipment Owner	- A.I.R Lab - Yue Hu (Assistant Professor, MME)	
Support Technicians	Robotis: https://www.robotis.us/	

Significant Hazards	• https://emanual.robotis.com/docs/en/edu/engineer/kit1/ • Keep the robot away from the face and body when the robot is operating. • Be careful for getting fingers or part of the body stuck in the robot joints. • Do not operate or store the robot under the direct sunlight. • Do not operate or store the robot near water or heat source. • Do not tamper or disassemble components. • Keep the robot and parts away from infants or younger children. • Do not impact or poke the robot with sharp objects. • The battery must be disconnected from the robot when not used or charged with the designated charger. • Do not disassemble or impact the battery or charger. • Do not heat the battery and avoid contact with fire and liquids. • Do not place battery in the microwave, laundry machine, refrigerator, or dryer. • Do not use damaged batteries (deformed, swollen, external damages). • Do not short the battery. • Do not reverse the polarity of the battery when charging. • Do not charge the battery when it is hot. Let the battery cools down to the room temperature before charging. • Do not store the battery in hot or humid place. • Do not charge multiple batteries with the charger at the same time. • Do not connect the battery to the charger when the charger is not connected to the power source.
Administrative Controls	• all users must be authorized by the equipment owner • who and when the equipment can be used: ○ Students working on the child HRI project can use ○ Must contact Yue Hu and Victoria Yang to coordinate use
Engineering Controls	• Power OFF on back of robot • Battery disconnect on robot base • All arms must be secured to the body as per installation specifications
PPE Required	• none
Relevant Standards and Codes	• none
Relevant MSDS	• all MSDSs can be found at msds.mme.uwaterloo.ca • list all relevant MSDS by chemical name • if none... enter "none"
Accident Procedure	• inform equipment owner and support technicians • See safety posters in this lab (E7 3431)
Emergency Shutdown Procedure	• The battery must be disconnected from the robot when not used or charged with the designated charger. • • provide direction for what to do to leave the equipment in a safe state in case the equipment must be abandoned quickly • how to avoid the equipment developing into a hazard when unattended • examples of a causal situation are: ○ fire alarm ○ medical emergency ○ building evacuation

Phone holder

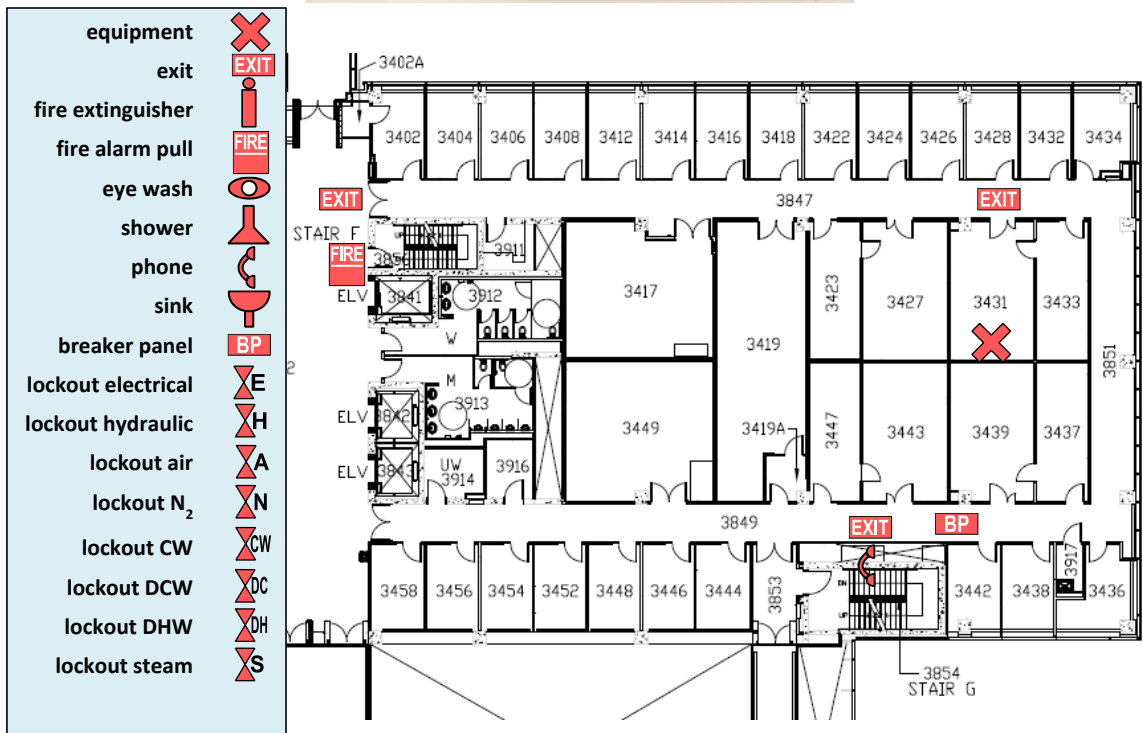
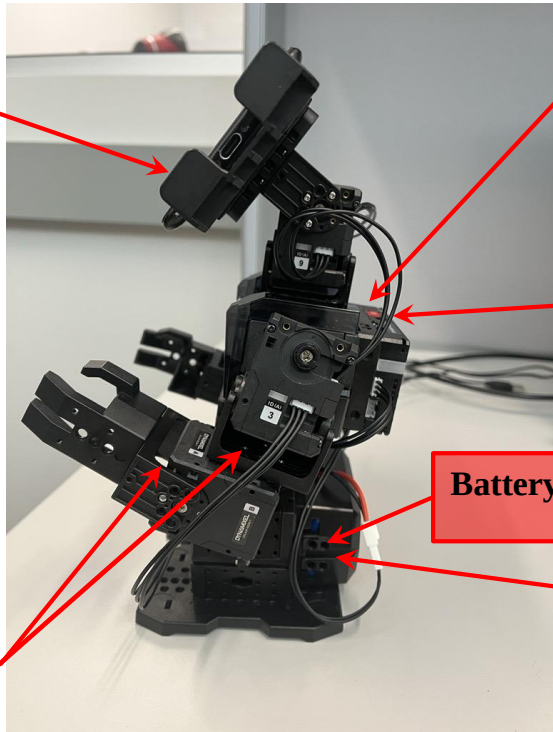
On/Off Switch

Power supply plug

Battery

Caution Hot

Pinch point



Pre-start Checklist

Document everything that needs to be done before starting the equipment or process. Include items such as:

- items to inspect
 - check if robot joints are not bent oddly
 - check that Raspberry Pi, U2D2 devices, and either battery or charger is plugged in
- items to have on hand
 - laptop with ubuntu, ROS and SDK installed
 - manual, documentation
- people to notify
 - Yue Hu
 - Victoria Yang
- Robot has enough empty space to allow free limb movement, and a large enough flat surface such that it wont fall over

Start-up Procedure

Document the start-up procedure. Include items such as:

- typical problems to watch for
 - motors not reading and writing (i.e., robot not moving even when motion scripts are executed)
- resolutions for typical problems
 - restart or check roscore is running on ros master (Raspberry Pi)
 - turn off power, unplug Raspberry Pi; then turn on power, plug in Raspberry Pi again
 - run launch file and scripts again
- what indicates all is going well
 - nodes and scripts are running and motors are moving according to scripts

Operating Procedure

- Physical setup:
 - All motors are linked together via 3-pin JST cables
 - Plug 3-pin JST cable (for chain of motors), and the U2D2 signal converter into U2D2 hub
 - Plug in power supply and turn hub power switch on
 - Connect signal converter (has micro-USB port) to computer or Raspberry Pi (has USB A port) using the USB A to micro-USB cable
- Software setup - Ubuntu:
 - Have ROS1 Noetic installed, a ROS workspace setup, and Dynamixel SDK installed:
 - ROS1 Noetic:
 - Regular install: <http://wiki.ros.org/noetic/Installation/Ubuntu>
 - Two line command install (If you have Ubuntu 20.04 LTS Focal Fossa):
<http://wiki.ros.org/ROS/Installation/TwoLineInstall/>
 - Setup a workspace:
<http://wiki.ros.org/ROS/Tutorials/InstallingandConfiguringROSEnvironment>
 - Dynamixel SDK:
https://emanual.robotis.com/docs/en/software/dynamixel/dynamixel_sdk/download/#repository
 - Checking connection port, in terminal:
 - \$ cd /dev
 - \$ ls -la ttyA*
 - Or do \$ ls -la ttyU*

- Or do \$ ls -la tty*
- identify which port is connected (on VMWare it was either ttyUSB0 or ttyUSB1, on dual boot/native ubuntu it's usually ttyACM0#, # being zero, 1 or other number. Can be different for others, in that case try tty*)
- \$ sudo chmod 777 <port>
- \$ ls -la /<port>
 - check if the permissions now show read, write, execute for all groups
- Navigate to Dynamixel SDK package and open read_write_node.cpp file
- In read_write_node.cpp file, make sure in "default setting" section, the default device name defined is the same as the port connected, if not change it to the correct port

Shutdown Procedure

1. stop running any relevant nodes/scripts and roscore on ROS
2. disconnect the micro-USB cable from the signal converter (i.e., disconnect motors from the set-up)
3. switch off power on the hub
4. disconnect power supply

Clean-up

Document the clean-up by creating a post in A.I.R Lab Equipment Teams channel that included information on usage according to specified format, and description of scope of the cleanup that was done.

Lockout

Make sure robot is turned off and disconnected from all other devices and power supply. Store in a place that is dry and away from other objects that can fall or push it. See [Health & Safety Ontario - Safeguarding/Lockout](#) for other general guidance.

Maintenance and Repair

- Charge battery to ensure robot is in good state for usage.
- Ensure servo motors are in zero position when not in use.
- Contact Robotis for other maintenance and repair

SOP Accepted by Equipment Owner

X

owner

digital signature (preferable) or pen on paper...

name	signature	date (yyyy-mm-dd)
------	-----------	-------------------

Authorized Trainers

X	X
trainer	owner

X	X
trainer	owner

X	X
trainer	owner

digital signature (preferable) or pen on paper...

trainer	authorized by	date
---------	---------------	------

name	name
------	------

signature	signature
-----------	-----------

name	name
------	------

signature	signature
-----------	-----------

name	name
------	------

signature	signature
-----------	-----------

Authorized Operators

X

operator

X

trainer

X

operator

X

trainer

X

operator

X

trainer

X

operator

X

trainer

digital signature (preferable) or pen on paper...

operator	trainer	date
----------	---------	------

name

name

signature

signature

name	name
signature	signature
name	name
signature	signature

Appendix B

User Study Supplementary Data

B.1 Participant Demographics

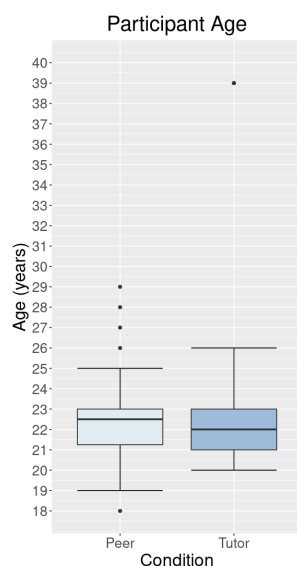


Figure B.1: Age Distribution

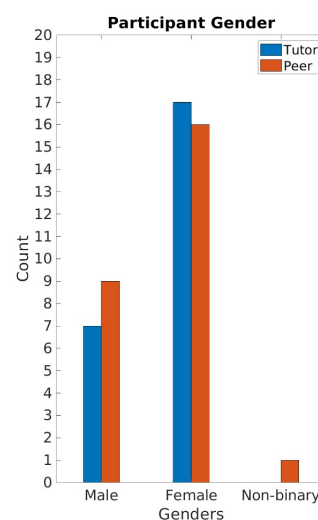


Figure B.2: Gender Demographics

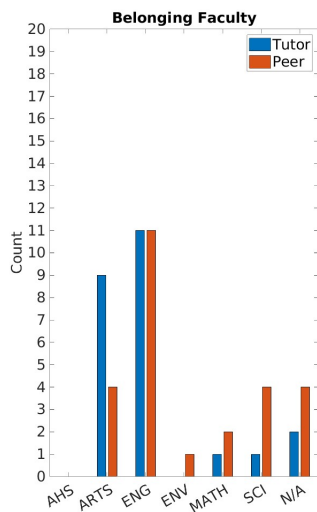


Figure B.3: Faculty Demographics

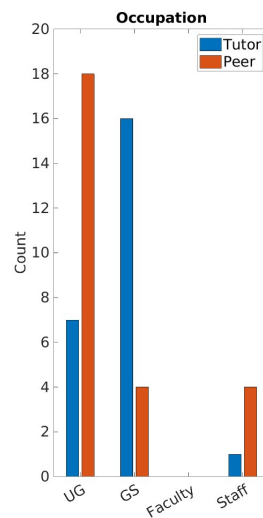


Figure B.4: Occupation Demographics

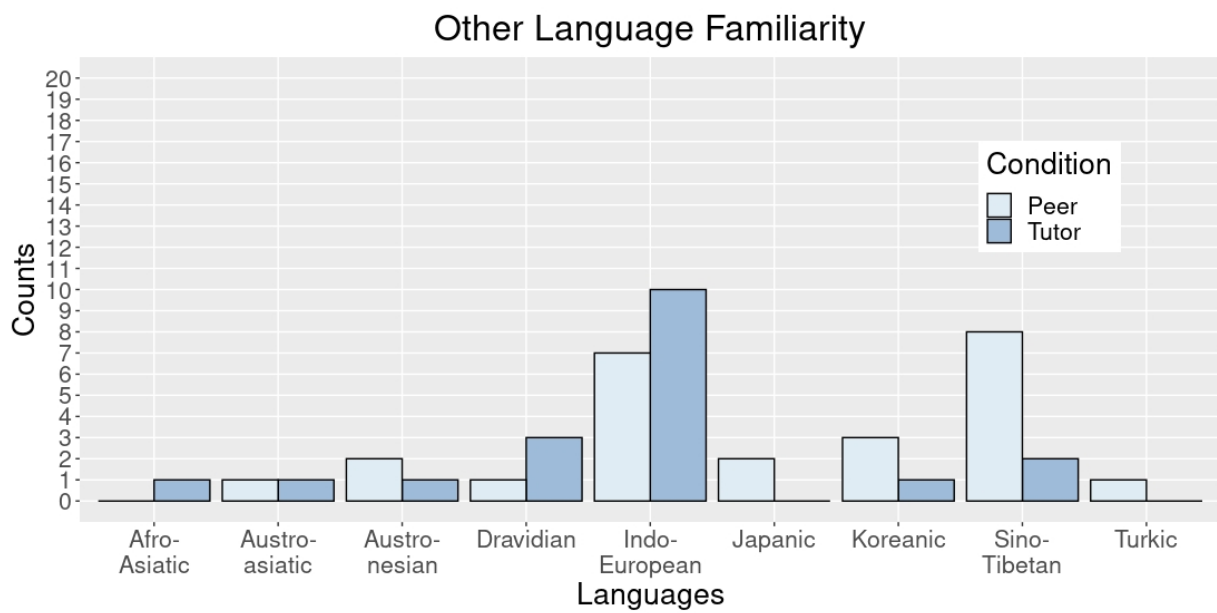


Figure B.5: Familiarity with Other Languages

B.2 Thematic Analysis - Additional Notable Participant Quotes

The following are additional quotes from participants' responses to open-ended questions which were notable and convey the themes and subthemes identified in the thematic analysis.

Theme 1 Face, Movement, and Voice of Robot

Subtheme 1.1 Facial Expressions

"I like that robot show some sleepy expression during explanation about one of word which is related to this emotion (T)", "The robot would have lots of different facial expressions that I found captured my attention more. (T)", "The robot's slow inhuman blinking was a little distracting (T)".

Subtheme 1.2 Movement and Gestures

"I noticed I was starting to get impatient waiting for the robot to complete the gesture (T)", "Some of movements was little unexpected and abrupt. It can scare (T)", "too much movement in short spans can be unnerving at times. (P)"

"I like the movement of the robots hands and arms since it made the robot seem friendlier. (P)", "sometime the motion of robot was unnecessary which actually caused some distraction. (T)"

Subtheme 1.3 Voice and Speech

"I did not think the voice of the robot matched the body (T)", "The text to speech voice of the robot/teacher was unnatural and becoming somewhat annoying. (P)"

"i like how the robot had a warm tone (T)"

Theme 2 Interaction with Robot

Subtheme 2.1 Perception of Robot Behavior

"... and it didn't complain if I chose a different answer. This made me feel a little more comfortable, because it felt like I wasn't being judged if I got the answer wrong. (P)", "it seemed truly friendly and encouraging. the interaction with the robot made me more happy (T)", "I noticed that the robot would nod at me when i answered the question correctly, which I found is somewhat encouraging. (T)"

Subtheme 2.2 Robot Awareness, Capability, and Effectiveness

“In my opinion, the robot in its current state does not add much to the learning process and just serves more as a companion than a helper. (P)”, “The experience with the robot could be improved if the robot was better at learning the language (P)”, “Maple sometimes felt like they were sabotaging the user’s answers vs. trying to help deduce the possible options (P)”.

Subtheme 2.3 Desire for More Interaction Opportunities

“If the robot is able to provide a bit more expression, this would also make it seem more natural (ex., thanking the participant/student and expressing that they hoped they had fun rather than going to sleep right away) (T)”, “Maybe if the robot be more creative in its response ... I don’t know if I say something that is not expected, how would the robot react? (P)”, “I didn’t make any mistakes but I wondered how it would respond if I had made a mistake, would it scold me or be very nice. Would it encourage to try again and ask for clarification. (T)”

Subtheme 2.4 Engagement with Robot

“Robot’s facial expressions and overall body language were very human-like and interactive (P)”, “The movement of the robot made it more engaging (T)”,

“I liked that the robot took a turn trying to say the word as well. I thought it was engaging. I liked that the robot tried to lead you to the wrong answer on occasion. Keeps you engaged and critically thinking. (P)”

“I found my eyes focused more on the computer screen than the robot’s anthropomorphic features. I felt kind of bad about this because it felt like I wasn’t paying enough attention to the robot’s movements and expressions. (T)” **Theme 3 Overall System/Technology**

Subtheme 3.1 Technology Acceptance

“It’s interesting, I love the robot design and interface. (P)”. “The robot looks really nice and cute, I was smiling all the time (T)”

Subtheme 3.2 Robot Role and Contribution

“the two system (ppt & robot) is connect together, but they are separated in terms of support my study, I feel like it is possible to learn the language just using one of the system (the ppt would be sufficient) (P)”, “More interactivity with the robot itself. Maybe the robot teaches me the language and rewards me instead of being my colearner? (P)”.

Theme 4 Esperanto Lesson

Subtheme 4.1 Pronunciation Discrepancy and Inaccuracy

“I noted that when pronouncing words, it was not able to speak with an accent similar to the voice played in the presentation. This could be seen as a criticism, but it would be unfair to say that humans are not guilty of this as well. (P)”.

Subtheme 4.2 Learning Experience

“the learning was at a good pace for me. (T)”, “at first it was confusing when the button to hear the answer and the button to choose the answer were attached. (T)”

Subtheme 4.3 Add Customization and Expand Scope

“Introduce difficulty level. (T)”, “It can be useful to see the transcription of the words. (T)”.

Appendix C

Ethics Clearance Certificate

UNIVERSITY OF WATERLOO

Notification of Ethics Clearance to Conduct Research with Human Participants

Principal Investigator: Yue Hu (Mechanical and Mechatronics Engineering)

Student investigator: Ya-Ting Yang

Co-Investigator: Kerstin Dautenhahn

File #: 44531

Title: Language Study Involving University Students/Staff/Faculty and Robot

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

Initial Approval Date: 04/20/23 (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: 04/21/24 (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



Heather Dekker, Ethics Advisor, hdekker@uwaterloo.ca, 519-888-4567, ext. 41506

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

Documents reviewed and received ethics clearance for use in the study and/or received for information:

file: MME_SOP_RobotisEngineerKit1_v1_20230320.pdf

file: pre_session_questionnaire_v1_20230320.pdf

file: post_session_questionnaire_v1_20230320.pdf

file: quiz_bank_v2_20230417.pdf

file: experiment_setup_diagram_v1_20230320.pdf

file: debrief_and_appreciation_v2_20230417.pdf

file: study_information_v2_20230417.pdf

file: study_poster_v2_20230410.pdf

file: recruit_email_dissemination_v2_20230417.pdf

file: recruit_email_consent-schedule_v2_20230417.pdf

file: participant_facing_feedback_appreciation_v1_20230417.pdf

file: consent_form_v2_20230417.pdf

Approved Protocol Version 2 in Research Ethics System

This is an official document. Retain for your files.

You are responsible for obtaining any additional institutional approvals that might be required to complete this study.