Enhancing Social Learning in Humanoid Robots Taught by Non-Expert Human Teachers

by

Pourya Aliasghari

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2025

© Pourya Aliasghari 2025

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Yiannis Demiris Professor, Dept. of Electrical and Electronic Engineering, Imperial College London
Supervisors:	Kerstin Dautenhahn Professor, Dept. of Electrical and Computer Engineering, University of Waterloo
	Chrystopher L. Nehaniv Professor, Dept. of Systems Design Engineering and Dept. of Electrical and Computer Engineering, University of Waterloo
	Moojan Ghafurian Assistant Professor, Dept. of Systems Design Engineering, University of Waterloo
Internal Members:	Stephen L. Smith Professor, Dept. of Electrical and Computer Engineering, University of Waterloo
	Yash Vardhan Pant Assistant Professor, Dept. of Electrical and Computer Engineering, University of Waterloo
Internal-External Member:	Edith Law Associate Professor, David R. Cheriton School of Computer Science, University of Waterloo

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the 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.

Statement of Contributions

This thesis includes peer-reviewed material from two published journal papers ([**J1**] and [**J2**]), three conference papers ([**C1**], [**C2**], and [**C3**]), and one journal manuscript in preparation ([**J3**]) for which Pourya Aliasghari was the main investigator and the first author.

Research presented in Chapter 3 (Experiment 1):

Pourya Aliasghari designed the study, implemented the apparatus, performed the experiment, conducted the data analysis, and drafted the manuscripts. Moojan Ghafurian, Chrystopher L. Nehaniv, and Kerstin Dautenhahn supervised the project, and reviewed and edited the manuscripts.

Citations:

[**J1**] **Pourya Aliasghari**, Moojan Ghafurian, Chrystopher L. Nehaniv, and Kerstin Dautenhahn. How Non-experts Kinesthetically Teach a Robot over Multiple Sessions: Diversity in Teaching Styles and Effects on Performance. *International Journal of Social Robotics*, 11-12(16):2079–2105, 2024.

[**C1**] **Pourya Aliasghari**, Moojan Ghafurian, Chrystopher L. Nehaniv, and Kerstin Dautenhahn. Kinesthetic Teaching of a Robot over Multiple Sessions: Impacts on Speed and Success. In *Proceedings of the 14th International Conference of Social Robotics (ICSR '22)*, pages 160–170. Springer, 2022.

Research presented in Chapter 4 (system development):

Pourya Aliasghari designed the system, programmed the algorithms, implemented the framework, performed the tests, conducted the data analysis, and drafted the manuscripts. Moojan Ghafurian, Chrystopher L. Nehaniv, and Kerstin Dautenhahn supervised the project, and reviewed and edited the manuscripts.

Citations:

[**J2**] **Pourya Aliasghari**, Moojan Ghafurian, Chrystopher L. Nehaniv, and Kerstin Dautenhahn. A Biologically Inspired Program-level Imitation Approach for Robots. *IEEE Access*, 13:89481-89501, 2025.

[**C2**] **Pourya Aliasghari**, Moojan Ghafurian, Chrystopher L. Nehaniv, and Kerstin Dautenhahn. A Biologically Inspired Program-level Imitation Approach for Robots: Proof-of-Concept. In *Proceedings of the 33rd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN '24)*, IEEE, 2024.

Research presented in Chapter 5 (Experiment 2):

Pourya Aliasghari designed the study, implemented the apparatus, performed the experiment, conducted the data analysis, and drafted the manuscripts. Moojan Ghafurian, Chrystopher L. Nehaniv, and Kerstin Dautenhahn supervised the project, and reviewed and edited the manuscripts.

Citations:

[J3] **Pourya Aliasghari**, Chrystopher L. Nehaniv, Moojan Ghafurian, and Kerstin Dautenhahn. Who to imitate? Exploring the role of human teachers' awareness of a learner robot's limitations and capabilities in improving learning from demonstration. *(Manuscript in preparation for journal submission)*

[C3] **Pourya Aliasghari**, Chrystopher L. Nehaniv, Moojan Ghafurian, and Kerstin Dautenhahn. Improving Robot Learning Outcomes in Human-Robot Teaching: The Role of Human Teachers' Awareness of a Robot's Vision Constraints. In *Proceedings of the 34th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN '25)*, IEEE, 2025.

Other contributions:

Pourya Aliasghari published a first-authored conference paper ([**C4**]) that was a follow-up of his M.A.Sc. thesis. This research focused on the explainability of learner robots and was, therefore, not within the scope of this thesis. The experiment reported in this paper was conducted virtually during the first term of Pourya Aliasghari's Ph.D. program, when the labs were not open for in-person experiments. Pourya Aliasghari designed the study, implemented the apparatus, performed the experiment, conducted the data analysis, and drafted the manuscript. Moojan Ghafurian, Chrystopher L. Nehaniv, and Kerstin Dautenhahn supervised the project, and reviewed and edited the manuscript.

Citation:

[C4] **Pourya Aliasghari**, Moojan Ghafurian, Chrystopher L. Nehaniv, and Kerstin Dautenhahn. How Do Different Modes of Verbal Expressiveness of a Student Robot Making Errors Impact Human Teachers' Intention to Use the Robot? In *Proceedings of the 9th International Conference on Human-Agent Interaction (HAI '21)*, ACM, 2021.

Collaborations:

Pourya Aliasghari collaborated with Sahand Shaghaghi on part of their Ph.D. research projects related to the gaze behaviour and personality of robots. Pourya Aliasghari contributed to the programming of the iCub robot and system implementation. One conference paper ([**C5**]), one workshop paper ([**W1**]), and two journal manuscripts in preparation ([**J4**] and [**J5**]) have been resulted from this collaboration. Pourya Aliasghari also provided mentorship for M.A.Sc. research studies of Priyank Avijeet and Andrea Chakma in one conference publication ([**C6**]) and one journal manuscript in preparation ([**J6**]).

Citations:

[**J4**] Sahand Shaghaghi, **Pourya Aliasghari**, Bryan Tripp, Kerstin Dautenhahn, and Chrystopher L. Nehaniv. The Impact of Robot Role and Personality on Participants' Perception of the Robot in a Human-Robot Teaching Task. *(Manuscript under third round of journal review)*

[**J5**] Sahand Shaghaghi, **Pourya Aliasghari**, Bryan Tripp, Britt Anderson, Kerstin Dautenhahn, and Chrystopher L. Nehaniv. Social Gaze Space Interaction Architecture Design and Evaluation. *(Manuscript in preparation for journal submission)*

[**J6**] Andrea Chakma, **Pourya Aliasghari**, Arnima Shrestha, and Moojan Ghafurian. Designing Social Robot Identities Using Affect Control Theory. *(Manuscript in preparation for journal submission)*

[**C5**] Sahand Shaghaghi, **Pourya Aliasghari**, Bryan Tripp, Britt Anderson, Kerstin Dautenhahn, and Chrystopher L. Nehaniv. Towards an Interaction Architecture for the iCub Robot: Social Gaze Space Model Adaptation for Social Interaction. In *Proceedings of the 23rd IEEE-RAS International Conference on Humanoid Robots (Humanoids '24)*, IEEE, 2024.

[**C6**] Priyank Avijeet, **Pourya Aliasghari**, and Kerstin Dautenhahn. A Greeting Preferences in a Hospitality Context: A Cross-Cultural Study with a Social Robot. In *Proceedings of the 33rd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN '24)*, IEEE, 2024.

[**W1**] Sahand Shaghaghi, **Pourya Aliasghari**, Bryan Tripp, Kerstin Dautenhahn, and Chrystopher L. Nehaniv. Utilization of Non-verbal Behaviour and Social Gaze in Classroom Human-Robot Interaction Communications. In *Proceedings of WTF Workshop held in conjunction with the ACM conference on Conversational User Interfaces (CUI)*, 2023.

Abstract

Tools that assist with daily tasks are valuable. For example, with the aging population in Canada and worldwide, there is a growing demand for ways to help older adults perform daily activities independently. Socially intelligent robots can promote independence by assisting with routine tasks. While advanced robots may be capable of performing various specialized operations, it is not feasible for their designers to program them in advance to effectively carry out multi-step, complex tasks requiring high-level planning and coordination, 'out of the box' in new environments and for users with diverse preferences. To successfully integrate into domestic environments, robots must learn new task knowledge from human users. Many of our own skills as human beings have been acquired through social learning, i.e., learning via observation of or interaction with others, throughout our lifetime. Social learning for robots enables the transfer of skills without the need for explicit programming, allowing users to teach robots via natural, intuitive, and interactive methods.

This thesis targets three key challenges in the social learning of robots: enabling non-expert humans to teach robots without external help, enabling robots to learn and perform multi-step tasks, and enabling robots to identify the most suitable teachers in their social learning.

The first phase of my research examines whether or not participants with no prior experience teaching a robot could become more proficient robot teachers through repeated human-robot teaching interactions. An experiment was conducted with twenty-eight participants who were asked to kinesthetically teach a Pepper robot various cleaning tasks across five repeated sessions. Analysis of the data revealed a diversity in non-experts' human-robot teaching styles in repeated interaction. Most participants significantly improved both the success rate and speed of their kinesthetic demonstrations after multiple rounds of teaching the robot.

The second phase introduces a novel, biologically inspired imitation approach enabling robots to understand and perform complex tasks using high-level programs that incorporate sequential regularities between sub-goals that a robot can recognize and achieve. To learn a new task, the system processes demonstrations to identify multiple possible arrangements of sub-goals that achieve the overall task goal. For task execution, the robot determines the optimal sequence of actions by evaluating the available sequences based on user-defined criteria, through mental simulation of the real task. This learning architecture was implemented on an iCub humanoid robot, and its effectiveness was evaluated across multiple scenarios.

In the third phase, I propose an attribute for identifying the most suitable teachers for a robot: human teachers' awareness of and attention to the robot's limitations and capabilities. I investigate the impact of this attribute on robot learning outcomes in an experiment with seventy-two participants who taught three physical tasks to an iCub humanoid robot. Teachers' awareness of the robot's visual limitations and learning capabilities was manipulated by offering the robot's visual perspective and by placing participants in the robot's position when labelling actions in demonstrations. Participants who could see the robot's vision output paid increased attention to ensuring that task objects in their demonstrations were visible to the robot. This emphasis on attention resulted in improved learning outcomes for the robot, as indicated by lower perception error rates and higher learning scores. I also propose a metric for robots to estimate the potential for receiving high-quality demonstrations from particular human teachers.

These findings demonstrate the feasibility of non-experts adapting to robot teaching through repeated exposure to human-robot teaching tasks, without formal training or external intervention, and also contribute to understanding factors in human teachers that lead to better learning outcomes for robots. Furthermore, I propose a robot learning approach that accommodates variations in human teaching styles, enabling robots to perform tasks with greater flexibility and efficiency. Together, these contributions advance the development of multifunctional and adaptable robots capable of operating autonomously and safely in human environments to assist individuals in various daily activities.

Acknowledgements

In the first place, I would like to thank my supervisors, Prof. Kerstin Dautenhahn, Prof. Chrystopher L. Nehaniv, and Prof. Moojan Ghafurian for their guidance and support throughout my Ph.D. program. Prof. Dautenhahn's patience and kindness, along with her extensive expertise in research and supervision, have been invaluable. Prof. Nehaniv's broad expertise, insightful guidance, and attention to detail have been crucial in shaping my research. Prof. Ghafurian's effort and dedication, along with her extensive expertise and constructive feedback, have made her an outstanding mentor for my academic and professional growth. I have learned greatly from each of you, and my research would not have been possible without your support.

I would like to extend my thanks to my examining committee members, Prof. Yiannis Demiris, Prof. Stephen L. Smith, Prof. Yash Vardhan Pant, and Prof. Edith Law. I greatly appreciate your willingness to serve on my committee, the time and effort you dedicated to this process, and the insightful feedback you provided.

I am grateful to all my fellow members of the Social and Intelligent Robotics Research Lab (SIRRL) for their kindness, support, and valuable feedback. I am grateful for all the collaborative efforts and friendships, which have made my Ph.D. journey productive and memorable. In connection with my research, I would like to thank Delara Forghani and Ali Yamini for their assistance with data encoding; Alicia Pan and Neil Fernandes for their help in testing my experimental setup; and Ali Ayub, Samira Rasouli, Ali Noormohammadi Asl, and Elaheh Sanoubari for their insightful advice.

I would also like to thank the iCub facility at the Istituto Italiano di Tecnologia (IIT) in Italy for their technical support during my research with the iCub robot. I am especially grateful to Antonio Consilvio and Antonio Azocar for travelling to Canada to provide on-site assistance.

I am also thankful to my wonderful parents for believing in me, and to my dear friends for their continued support and encouragement, and for the joyful moments we have shared.

Lastly, my deepest gratitude goes to my dearest wife, Fatemeh, for her endless love, compassion, care, and support. This thesis would not have been possible without your constant encouragement. You have been an invaluable source of motivation and my emotional support during the difficult days. I am deeply grateful to have you by my side throughout this journey.

This research was undertaken, in part, thanks to funding from the Canada 150 Research Chairs Program. I would also like to thank the participants who took part in my studies.

Dedication

To my beloved wife and my best friend, Fatemeh, for whom I am incredibly grateful.

To my parents, Akhtar and Hamid, who always supported me through every challenge.

Table of Contents

Examining Committee Membership ii

Author's Declaration iii

Statement of Contributions iv

Abstract vii

Acknowledgements ix

Dedication x

List of Figures xvii

List of Tables xx

List of Abbreviations xxii

1 Introduction 1

1.1 Motivation 1

1.2 Challenges in Social Learning of Robots 2

1.2.1 Non-expert Humans Teaching Robots 2

1.2.2 Robots Learning Multi-step Tasks 3

1.2.3 Robots Identifying Most Suitable Teachers . . . 4
1.3 Research Objectives . . . 5
1.4 Contributions . . . 6
1.5 Thesis Outline . . . 9

2 Background Literature **11**
2.1 Social Learning in Animals and Humans . . . 12
2.1.1 Social Learning Mechanisms . . . 13
2.1.2 Social Learning Strategies . . . 18
2.1.3 Program-Level Imitation in Primates . . . 21
2.1.4 Teacher's Awareness of the Capabilities of Students . . . 24
2.2 Social Learning and Robots . . . 25
2.2.1 Non-expert Humans Teaching Robots . . . 26
2.2.2 Human Teachers' Behaviour When Teaching Robots . . . 28
2.2.3 Robots Learning Task Structures . . . 29
2.2.4 An Example Biologically-Inspired Robot Learning System . . . 31
2.2.5 Robots Identifying Suitable Human Teachers . . . 32
2.3 Visual Perspective-Taking in Humans and Robots . . . 33
2.4 Positioning My Work Within the Literature . . . 34

3 Experiment 1: How Non-experts Kinesthetically Teach a Robot over Multiple Sessions **37**
3.1 Research Questions . . . 38
3.2 Methodology . . . 39
3.2.1 Robotic Implementation . . . 40
3.2.2 Experimental Procedure . . . 44
3.2.3 Data Collection and Analysis . . . 46
3.2.4 Participants . . . 52

3.2.5 Data Reliability . . . 52
3.2.6 Data Analysis and Statistical Modelling . . . 54
3.3 Results . . . 56
3.3.1 Success of Demonstrations . . . 56
3.3.2 Time Spent on Providing the Demonstrations . . . 59
3.3.3 Amount of Gaze Fixations on Robot's Hand . . . 60
3.3.4 Behavioural Data on Manipulation and Teaching Strategy . . . 61
3.3.5 NASA-TLX Workload Items . . . 63
3.3.6 Other Observations . . . 64
3.4 Discussion . . . 65
3.4.1 Teaching Success and Duration . . . 65
3.4.2 Human Gaze Patterns During Teaching . . . 69
3.4.3 Behavioural Data on Manipulation and Teaching Strategy . . . 70
3.4.4 Human Teacher's Workload . . . 71
3.4.5 Other Reflections . . . 72
3.5 Limitations . . . 75
3.6 Future Directions . . . 76
3.7 Summary and Conclusion . . . 77

4 System Development: A Biologically Inspired Program-Level Imitation Approach for Robots 79

4.1 Research Questions . . . 80
4.2 Program-level Imitation Architecture . . . 80
4.2.1 Demonstration Observer . . . 81
4.2.2 Program Generator . . . 84
4.2.3 Task Progress Observer . . . 88
4.2.4 Sequence Selector and Action Planner . . . 91
4.3 Robotic Implementation . . . 93

4.3.1 Abstract Physical Tasks . . . 94
4.3.2 Demonstrations Observer for iCub . . . 94
4.3.3 Task Progress Observer for iCub . . . 96
4.3.4 Sequence Selector for iCub . . . 97
4.3.5 Action Planner for iCub . . . 98
4.4 System Validation . . . 99
4.4.1 Learning and Adaptation . . . 99
4.4.2 Error Handling . . . 105
4.4.3 Task Execution . . . 110
4.5 Discussion . . . 113
4.6 Limitations . . . 118
4.7 Future Directions . . . 119
4.8 Summary and Conclusion . . . 120

5 Experiment 2: Exploring Humans' Awareness of a Robot's Limitations and Capabilities as an Indicator of Suitable Teachers for the Robot 121
5.1 Research Questions . . . 122
5.2 Methodology . . . 124
5.2.1 Resources Provided to Participants . . . 124
5.2.2 Abstract Physical Tasks . . . 126
5.2.3 Robot Learning and Performance Mechanism . . . 131
5.2.4 Experimental Procedure . . . 133
5.2.5 Measures and Metrics . . . 136
5.2.6 Participants . . . 141
5.2.7 Statistical Analysis . . . 142
5.3 Results . . . 144
5.3.1 Attention to the Visibility of Demonstrations . . . 144
5.3.2 Attention to Demonstrating Primitive Actions . . . 145

5.3.3 Quality of Recorded Demonstrations . . . 146
5.3.4 Robot Learning Outcomes . . . 159
5.3.5 Perceptions of Robot's Social Attributes . . . 164
5.3.6 Perceptions of Teaching Workload . . . 166
5.3.7 Effectiveness of PLI when Non-experts Teach a Robot . . . 166
5.3.8 Characteristics of Demonstrations Over Time . . . 172
5.3.9 Patterns to Find Better Teachers for the Robot . . . 175
5.4 Discussion . . . 177
5.4.1 Attention to the Visibility of Demonstrations . . . 177
5.4.2 Attention to Demonstrating Primitive Actions . . . 180
5.4.3 Quality of Recorded Demonstrations . . . 180
5.4.4 Robot Learning Outcomes . . . 183
5.4.5 Perceptions of Robot's Social Attributes . . . 184
5.4.6 Perceptions of Teaching Workload . . . 184
5.4.7 Effectiveness of PLI when Non-experts Teach a Robot . . . 185
5.4.8 Characteristics of Demonstrations Over Time . . . 186
5.4.9 Patterns to Identify Better Teachers for the Robot . . . 187
5.5 Limitations . . . 188
5.6 Future Directions . . . 189
5.7 Summary and Conclusion . . . 190

6 Conclusion **192**
6.1 Summary of Findings . . . 192
6.2 Contributions to Original Knowledge . . . 198
6.3 Overall Limitations and Future Work . . . 199

References **202**

APPENDICES **225**

A Ethics Clearance Certificates **226**

B Experiment 1 Additional Information **231**

C Experiment 2 Additional Information **235**

C.1 The Goal-Highlighting Activities 236

C.2 Online Interface for Collecting Questionnaire Data 257

C.3 Other Experimental Material and the Setup 268

List of Figures

1.1 Three phases of research presented in this thesis 7

2.1 An example of social learning in animals 13
2.2 Classification tree for social learning mechanisms 16
2.3 Social learning mechanisms based on their required cognitive load 17
2.4 Social learning strategies categorization 20
2.5 Organization processing nettle leaves by mountain gorillas 23

3.1 Layout and sequence of the experimental setup 41
3.2 Pepper robot used in the experiment 43
3.3 Computational method to estimate success of demonstrations 48
3.4 Processed scene camera video from the gaze recording device 50
3.5 Average subjective and computational success of demonstrations 57
3.6 Average success of demonstrations for each cluster of participants 58
3.7 Average time spent on teaching each task 60
3.8 Average percentage of participants' gaze fixations on the robot hand 62
3.9 Average percentage of the time teaching with a particular strategy 63
3.10 Average ratings of NASA-TLX questionnaire items 66

4.1 Overall architecture of our proposed PLI architecture 82
4.2 Layout of the abstract tasks used for system evaluation 95
4.3 Process of robot adaptation to the abstract potato boiling task 101

4.4 Process of robot adaptation to the abstract organization task 103
4.5 Performance evaluation of the PLI architecture in Tasks 1 and 2 111
4.6 Timelines of three example task execution by the iCub robot 114

5.1 A participant demonstrating a task to the iCub robot 125
5.2 Experimental setup and robot's camera view 127
5.3 Snapshots for the goal-highlighting activity 128
5.4 Abstract Tasks 1 and 2 used in the experiment 130
5.5 Abstract Task 3 used in the experiment 131
5.6 Task objects used for robot teaching and practising 132
5.7 Histogram of participants' age 142
5.8 Participants' gender and robotics experience 143
5.9 Average self-reported attention to visibility of objects 146
5.10 Average self-reported frequency of checking the RCV and its usefulness . . 147
5.11 Average self-reported attention to performing primitive actions 148
5.12 Occurrence of performance issues in each demonstration 152
5.13 Total and average occurrence of performance issues in demonstrations . . . 153
5.14 Average occurrence of sub-goal detection issues in demonstrations 156
5.15 Occurrence of sub-goal detection issues in each demonstration 157
5.16 Average demonstration variability scores 159
5.17 Average number of sequences in the generated programs 161
5.18 Average normalized program generation score 162
5.19 Different classes of sequences in generated programs 163
5.20 Average subjective evaluations of learning success 164
5.21 Average perceptions of robot's social attributes 165
5.22 Average perceptions of teaching workload 165
5.23 Time spent on providing demonstrations 173
5.24 Average task components and task objects disappearance rate 176

5.25 Average of demonstrations' characteristics for each participant 178
5.26 Correlation between sub-goal detection errors and objects disappearance . 179

B.1 The flyer for advertising the experiment 232
B.2 Overall setup in the experimental room 233
B.3 The material used for the cleaning tasks 234

C.1 The GHA about a robot learning how to make coffee 236
C.2 The GHA about a robot learning how to start a washing machine 243
C.3 The GHA about a robot learning how to declutter a kitchen space 250
C.4 The pre-session online module 257
C.5 The ending questionnaire module 261
C.6 The flyer for advertising the experiment 268
C.7 Overall setup in the experimental room 269
C.8 Another view of the experimental setup 270

List of Tables

3.1 LMM predicting the subjective estimates of success of demonstrations . . . 59
3.2 LMM predicting the time spent on providing demonstrations 61
3.3 LMM predicting the amount of gaze fixations on robot's hand 61
3.4 LMM predicting percentage of the time teaching with a particular strategy 64
3.5 LMM predicting NASA-TLX questionnaire items 67

4.1 List of sets identified by program generator 86
4.2 Robot sequence selection for the organization task 106

5.1 Experimental factors and their levels for each group 127
5.2 Items in the ending questionnaire 137
5.3 LMM predicting participants' self-reported attention to visibility of objects 145
5.4 Different types of performance issues in demonstrations 151
5.5 LMM predicting the occurrence of performance issues in demonstrations . 154
5.6 LMM predicting normalized sub-goal detection error rate 156
5.7 Different sources of sub-goal detection issues in demonstrations 158
5.8 LMM predicting demonstration variability scores 159
5.9 LMM predicting number of sequences in the generated programs 160
5.10 LMM predicting normalized program generation score 162
5.11 Participants' thoughts or suggestions on the teaching process 170
5.12 LMM predicting total demonstration time 174

5.13 LMM predicting total time spent on each sub-goal 174
5.14 LMM predicting normalized task components disappearance rate 176
5.15 LMM predicting normalized task objects disappearance rate 177

List of Abbreviations

AI Artificial Intelligence 1, 17, 25, 201

AIC Akaike's Information Criterion 55, 142

DOFs Degrees of Freedom 39, 41, 43, 45, 46, 76, 93

GHA Goal-Highlighting Activity 125–128, 134, 135, 137, 144–146, 148, 150, 154, 155, 159–161, 172–177, 179–181, 183, 185, 187, 189, 196, 235–256, 263

HRI Human-Robot Interaction 6, 11, 34, 93, 141, 142, 198, 200

k-NN k-Nearest Neighbours 31

LfD Learning from Demonstration 3, 32, 35, 76, 115, 118, 188, 194, 197, 198, 200

LLM Large Language Model 31, 35, 119, 200

LMM Linear Mixed-effects Model 54, 55, 59, 60, 62, 64, 69, 142, 144, 145, 150, 154, 155, 160, 161, 164, 166, 172, 173, 175

PLI Program-Level Imitation 4, 9–11, 21, 22, 35, 79, 80, 82, 84, 93, 99, 102, 111–113, 115–124, 132, 135, 138–141, 147, 149, 155, 166, 169, 172, 181, 183–186, 188, 189, 191, 194–196, 200, 201

RCV Robot's Camera View 124, 125, 127, 134, 135, 137, 144, 145, 147, 150, 154–156, 159, 160, 162, 164, 166, 175, 177, 179–181, 183–185, 187, 189, 196, 262, 269, 270

SLMs Social Learning Mechanisms 12, 15, 24, 32, 35

SLSs Social Learning Strategies 13, 19

TA Thematic Analysis 140, 167, 185

ToM Theory of Mind 33

VPT Visual Perspective-Taking 11, 33, 34, 36, 125, 179, 180, 189, 199

VR Virtual Reality 30

ZPD Zone of Proximal Development 24

Chapter 1

Introduction

Tools that assist our daily lives are valuable. For example, with the population aging in Canada and worldwide [21, 197], there is a growing demand for ways to help older adults perform their daily activities independently. Robots can be promising tools to increase people's independence by assisting them with routine tasks. Considering that people respond socially to the presence of robots, which makes possessing social skills essential for the acceptance of robots in real-life settings [55, 70], my research focuses on robotic systems that also interact socially with people, i.e., *social robots*. Social robots are designed to achieve goals beyond the specific domain of assisting older adults, as these robots have many other applications, such as in tasks including cleaning and organizing [187, 65], healthcare [34], and education [25].

1.1 Motivation

With recent advancements in Artificial Intelligence (AI) and robotics, today's robots excel at performing specialized actions such as accurate object manipulation [49, 130] and safe navigation [48]. Nevertheless, it is still challenging to use robots in real-world settings, for example, in a home environment to assist us with everyday activities [108]. In the first place, this is because our routine activities often involve multi-step tasks (e.g., cooking a dish, preparing coffee, and organizing a space) that consist of several distinct actions and may require high-level planning and coordination based on an in-depth understanding of task structures. Moreover, the habits, needs, and preferences of robot users, as well as environmental factors (e.g., lighting, layout of the homes, etc.) vary from one home to another and may even change over time. While advanced robots may be capable of

performing various specialized operations, it is not feasible that they are programmed by their designers in advance to know how to effectively carry out such multi-step and complex tasks 'out of the box' in new, dynamic environments and for users who have different preferences. In fact, robots need to acquire new task knowledge or adjust their skills, e.g., by learning from humans, to be successfully integrated into domestic environments as multi-functional tools rather than single-use devices [22]. To learn from humans, robots may engage with them in human-robot teaching scenarios or observe humans as they perform their own tasks. Adaptable, multi-functional robots can be more cost-effective, which may make robotic technology more accessible to a wide range of users.

For us as human beings, most of what we know and many of our skills have derived from *social learning* [206], a special kind of learning which is through observation of or interaction with others [89]. As an example, social learning occurs when children learn complex behaviours, such as language use and handling tools, by watching and interacting with their parents, minimizing the need for parents to explicitly teach them [191]. Utilizing social learning for teaching robots facilitates transferring skills to them via natural, intuitive, and interactive methods [27], and makes social robots more 'like us' [54]. Robots that can seamlessly learn from different individuals may successfully become part of our daily lives, e.g., to increase people's independence.

1.2 Challenges in Social Learning of Robots

The research I present in this thesis aims to enhance the social learning capabilities of robots. The main objective is to allow robots to learn how to perform everyday tasks from people who may have different backgrounds and individual preferences, in a manner that is familiar and natural to us. An inclusive learning approach can make robotic technology accessible for everyone, including older adults and those without technical knowledge. Drawing inspiration from how humans and other animals learn socially, this research addresses a number of challenges in the following key subjects within robots' social learning.

1.2.1 Non-expert Humans Teaching Robots

Having robots acquire new task knowledge from users presents a significant challenge, particularly to those who may greatly benefit from the assistance of robots, including older adults and their families. This challenge is because users may have little or no experience programming robots to perform novel tasks or adjusting them to their unique

home environments. To reduce the need for robot teachers to possess specialized technical knowledge and programming skills, several methods for robot learning have been developed to date, most notably *imitation learning* [27, 170] under Learning from Demonstration (LfD) techniques [22]. LfD techniques generally attempt to make the teaching process as natural as possible and, therefore, enable even non-experts to teach robots to use them more effectively in real-life environments [132, 22, 155]. *Kinesthetic teaching* is one example of a relatively natural technique for teaching a robot and is a technique whereby a human physically guides a robot through a task, e.g., by grasping onto it and moving its arms [45].

Even though the techniques noted above for teaching robots help make adaptable robots usable to everyone, in realistic settings, users who are not familiar with how to teach tasks to a robot may face various challenges while teaching robots, such as presenting non-smooth and inefficient trajectories [163] and failing to provide sufficient sub-task abstraction [77]. Thus, a variety of strategies have been proposed and tested to improve non-expert humans' ability to teach robots, including real-time generation of teaching guidance by the robot [44], or providing human teachers with instructional materials [43].

Nevertheless, in many real-world applications of robots, it may be difficult to easily access instructions on robot teaching. Non-expert users may find themselves in situations where the only available option to learn how to instruct a robot is simply exploring robot teaching on their own, without any additional resources or instructional tools. Considering such a scenario that might happen in the real world, Phase 1 of my research explores the feasibility of non-experts adapting their teaching strategies and learning how to teach robots through experiencing multiple rounds of teaching interaction, without any external help or resources.

1.2.2 Robots Learning Multi-step Tasks

Performing everyday activities often involves high-level planning and coordination based on an in-depth understanding of task structures. For example, operating a coffee machine to make coffee may involve grabbing coffee beans and a mug simultaneously to enhance task efficiency, grinding beans according to the specific type of coffee, optionally adding water based on its existing level, personalizing milk and sugar additions, etc. Moreover, in the case of learning a new activity, such as cooking an unfamiliar recipe or assembling a new piece of furniture, we often arrange skills we already have, such as mixing ingredients or screwing, into novel and complex sequences. Initially, we may exactly follow the given directions, but once we gain a better understanding of the sequence of stages and the structure of sub-goals, we may adapt to our unique situation to improve efficiency. For

instance, based on how our own kitchen is organized, we may pick up a spice required later while grabbing a mixing bowl needed right now, thus avoiding repeated trips to the same cabinet.

The high-level knowledge about achieving specific goals and how to structure events into discrete sub-goals is essential for successful task planning [71]. Humans and many other species of primates, therefore, do not always merely copy the chain of observed events. Instead, they identify and use hierarchical action structures, which include sequential regularities, bimanual coordination, and possible arrangements of sub-goals [203, 42, 99]. This imitation of overall task structures, rather than details of actions, is referred to as Program-Level Imitation (PLI) [42]. As will be detailed later in Section 2.1.3, PLI is a high-level mechanism used by primates and humans to learn complex skills efficiently.

Efficient learning and adaptation methods are essential for robots to assist us in daily tasks [155]. Due to the complex nature of human everyday activities, a robot may observe people carrying out a similar task with varying orders of actions and motions of the hands and use objects with unique appearances. In such cases, as with humans and other primates, robots could identify and use the hierarchical structure of actions to know what sub-goals are needed and in what order, and how irrelevant information can be discarded [185]. Knowing various ways of structuring task steps to achieve a specific goal may make the robot's actions more efficient. Therefore, in Phase 2 of my research, I propose a novel, biologically inspired PLI framework to enable robots to understand and perform complex actions by incorporating sequential regularities between the task steps that a robot can recognize and physically achieve.

1.2.3 Robots Identifying Most Suitable Teachers

In the near future, when robots are integrated into our daily lives and want to learn from humans, it is likely that a robot will sometimes find several people who can provide examples of how one particular task is performed. In such cases, having an answer to the question of 'who to imitate,' from the robot's own viewpoint, can assist the robot in utilizing the full potential of social learning [56]. Addressing the topic of automatic teacher selection in robots requires establishing metrics that robots can use to evaluate multiple teachers. To date, some effort has been made in this direction through approaches other than computational modelling. For example, the engagement of teachers has been proposed as a metric to define a good teacher for robots [141]. However, there is still no conclusive answer to this problem.

As humans, we may assess various factors in potential models, e.g., degree of expertise,

age, and familiarity, to find the best sources of information for our social learning [112, 152]. Social learning of primates also involves selecting which models to copy based on a variety of factors that are highlighted in Section 2.1.2, including the model's success, knowledge, prestige, and age [105].

The factors humans and primates use to identify the most suitable teachers may not always be available to robots or be meaningful enough to indicate a good teacher for them. For example, when learning how to organize a kitchen space, a robot may determine multiple members of the household to be 'expert' and 'successful,' because the task is fairly easy for humans, and all potential models are 'familiar,' because they are all members of the household.

My research, therefore, aims to identify factors in human teachers that can contribute to better robot learning outcomes, specifically in human-robot teaching. Taking a closer look at human-human teaching, high-achieving teachers are often well aware of the capabilities and existing knowledge of their students [151, 91]. Inspired by this particular aspect of human-human teaching, Phase 3 of this research examines whether learning from human teachers who are better aware of and pay more attention to the limitations, capabilities, and learning strategies of a learner robot leads to better robot learning outcomes. I also explore metrics that robots can potentially use to evaluate this aspect of human teachers' awareness to identify the most suitable teachers.

1.3 Research Objectives

To address the challenges noted above, the following main research objectives are defined. Note that Chapters 3, 4, and 5 each contain detailed research questions related to the main research objectives of the thesis.

O1 Understand how the behaviours of non-expert humans and characteristics of their demonstrations change when they teach tasks to a humanoid robot multiple times

It is important to determine whether or not non-expert humans who have no previous experience teaching robots can learn how to teach robots more effectively through repeated teaching interactions. To this aim, a research objective is defined to study trends in different aspects of the behaviours of human teachers for robots and what they teach the robot when they demonstrate multiple examples of how to perform basic household tasks to a humanoid robot, with no external training on how to teach a robot.

O2 Design, implement, and evaluate a learning framework for robots to allow them to understand the sequential structure of multi-step tasks and efficiently perform such tasks

One of the goals of this research is to design, implement, and evaluate a novel robot learning system for enabling robots to learn how to assemble their basic capabilities into high-level plans for multi-step everyday tasks. This system is expected to enable robots to acquire knowledge about task structures by learning from people with different styles in doing a particular task.

O3 Study how giving humans extra resources for understanding robot limitations and learning strategies when they teach a robot may impact robot learning and human-robot teaching

I propose to use the attention of human teachers to the limitations and capabilities of a learner robot as a criterion for identifying better teachers for the robot. To investigate the impact of this factor on robot learning, one of the research objectives is to study whether or not an increased awareness about a robot's visual capabilities and learning strategies in human teachers, by offering perspective-taking tools and extra teaching resources, improves robot learning outcomes.

O4 Detect patterns in non-experts' teaching that a robot can use to identify the most suitable teachers

Towards finding an answer to the question of 'who to imitate' in imitation learning of robots, one of the objectives of this research is to identify patterns in human teachers' behaviours and the characteristics of their provided demonstrations that can indicate higher learning outcomes for robots. Such indicators should be able to be assessed in teachers by robots themselves.

1.4 Contributions

Overall, this research contributes to developing adaptable robots that can operate autonomously and reliably in human environments to assist individuals in various daily activities by learning from diverse people via natural and interactive methods. The present thesis makes the following specific contributions to the body of knowledge in the fields of Human-Robot Interaction (HRI) and robot learning, through a three-phase research to achieve the research objectives noted above (summarized in Figure 1.1):

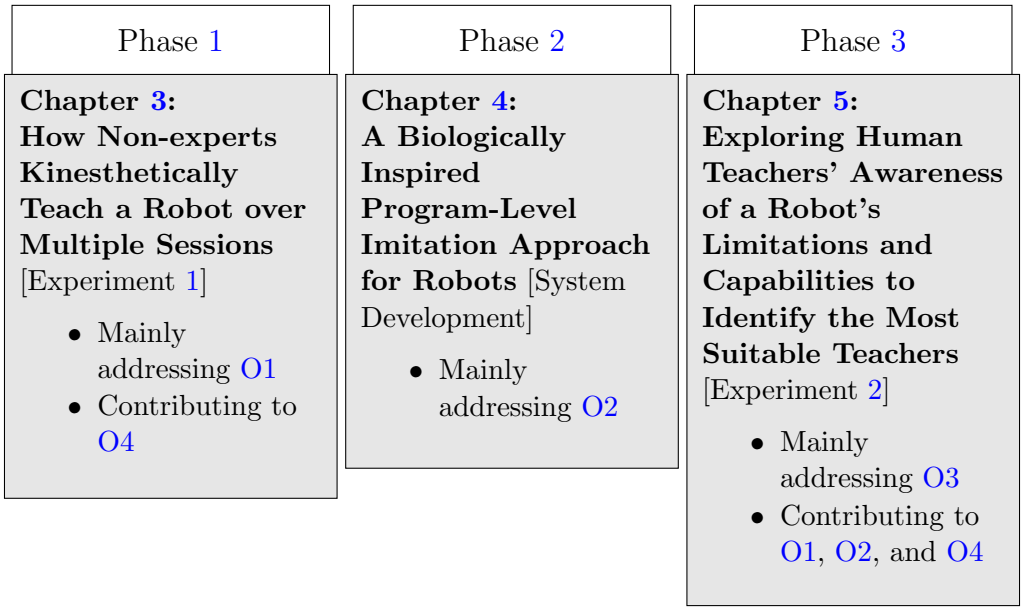


Figure 1.1: The phases of research presented in each chapter of this thesis and the research objectives addressed in each phase.

- **Studying the behaviour of non-expert humans when teaching physical tasks to a humanoid robot**

  In Phase 1 of this research, the experiment presented in Chapter 3 (Experiment 1) was conducted to investigate potential changes in the behaviours of inexperienced human teachers for robots and the success of teaching over time, when providing multiple (20 in total) kinesthetic task demonstrations without any external support. This in-person experiment involved 28 participants who were asked to kinesthetically teach a Pepper humanoid robot [5] different cleaning tasks over five repeated sessions, each session including four tasks. Overall, 140 experimental sessions were held with participants physically interacting with a humanoid robot that was functioning semi-autonomously.

  Unlike many other previous works, participants in this study were not formally instructed on how to perform the teaching or how to refine their demonstrations. Rather, over five sessions, they experienced kinesthetically teaching a humanoid robot multiple tasks in unsupervised interactions, i.e., without the presence of an experimenter, and without being given any detailed feedback on their performance. During the teaching process, measures including joint positions of the robot's arm, the gaze

direction of participants (as they were wearing eye tracking glasses), and videos of them teaching the robot were recorded and analyzed, mostly using computational tools I developed specifically for this study and also through manual annotation of videos from the 560 demonstrations. This investigation answered the questions of whether or not inexperienced and unsupervised human teachers can gradually become more proficient robot teachers through repeated teaching interactions with humanoid robots, and how this might happen, considering the physical properties of movements and behaviours of human teachers, contributing to addressing O1.

- **Designing, implementing, and evaluating a novel, biologically-inspired robot learning architecture**

  In Phase 2 of this research and to address O2 and also overcome some limitations of the first phase (e.g., the absence of a robot learning system), I proposed a novel, biologically inspired imitation learning framework that enables robots to understand and perform complex actions using high-level programs, which is presented in Chapter 4 of this thesis. Since humans tend to project their own mental model when a robot appears to be human-like [120], with a framework that is biologically inspired and enables robots to learn comparable to how humans and primates learn, if aspects of the learning process are visible to human teachers, it can be expected that non-expert users can more easily identify the capabilities of the learning system and teach robots more effectively.

  With the proposed imitation learning system, learner robots may use inconsistencies among the demonstrations to their advantage and discover various ways of performing a task based on limited examples, given that people may have different styles in completing a particular task. This framework gives robots the potential to receive task demonstrations from humans through various modalities (e.g., kinesthetic teaching, behavioural observation, or verbal instruction), with or without errors, and allows for customization in task performance according to the current user's expectations. A program learned by a robot may be used by different robots, each utilizing its own primitives to achieve the same sub-goals. By assessing sub-goals in real-time during task execution based on a program, robots can be interrupted and resume where they left off or recover from failed actions. I implemented the proposed learning architecture on a physical iCub humanoid robot [137] and evaluated the effectiveness of the system in multiple scenarios of learning abstract tasks, both in real settings and in simulation.

- **Exploring the role of human teachers' awareness of a robot's visual constraints and learning strategies in human-robot teaching outcomes**

To contribute to answering the question of 'who to imitate' from the point of view of learner robots, I proposed attributes in human teachers for robots that learn through visual observations, namely human teachers' awareness of and attention to the robot's visual capabilities and constraints, as well as to the learning strategy used by the robot, and explored how these attributes could impact robot learning outcomes. In Phase 3 of this research, I conducted another in-person experiment (Experiment 2), detailed in Chapter 5, to mainly address O3. This experiment involved 72 participants who taught three physical tasks to an iCub humanoid robot. Aspects of teachers' awareness were manipulated by offering extra resources for better understanding the limitations of the robot and/or its learning strategies to certain groups of participants. A dataset containing 1080 task demonstrations provided by non-expert humans was recorded as part of this experiment. This study presented experimental evidence suggesting a positive correlation between a proposed attribute in humans, i.e., their awareness of the visual constraints of a robot, and higher learning performance of the robot, which was quantified by establishing several subjective and objective metrics. The results of this experiment also contributed to O1, as the changes in the characteristics of 15 demonstrations provided by each participant over a session were analyzed, and also to O2, as the effectiveness of the developed PLI framework was evaluated with non-expert human teachers.

- **Proposing metrics for robots to identify the most suitable human teachers**

  Data collected in Experiments 1 and 2 suggested a number of indicators for high-achieving human teachers for robots. Because the first experiment mostly focused on the behaviours of humans when teaching a robot, the identified indicators were the characteristics of human teachers' behaviours. An advanced robot could collect those measures to make inferences and potentially identify better human teachers. The second experiment led to the discovery of a correlation between a defined characteristic of the demonstrations, which was recorded by the robot, and an indicator of high learning success. These patterns observed in the collected data addressed O4.

1.5 Thesis Outline

The following chapters make up the rest of this thesis:

Chapter 2 provides a literature review of topics related to the thesis, including social learning in animals, humans, and robots, how non-expert humans teach robots with and without external help, learning task structures by imitation in robots, human teacher's

awareness of the capabilities of students, and perspective taking in humans and robots. This chapter draws partially from the background and related work sections of my published papers [18, 19, 20].

Chapter 3 presents Experiment 1, studying how non-experts kinesthetically teach a humanoid robot over multiple sessions. This chapter is based on my research published in [18].

Chapter 4 details the design, implementation, and evaluation of the proposed PLI learning framework, based on my research published in [19].

Chapter 5 reports Experiment 2, which explores the role of human teachers' awareness of a robot's visual constraints and learning strategies in improving robot learning outcomes. A part of this chapter is based on my publication in [20].

Chapter 6 concludes the thesis by summarizing the findings concerning each research objective, listing the contributions to knowledge, and discussing the overall limitations of the presented work and proposing future research directions.

Appendix A provides the ethics clearance certificates for the conducted experiments.

Appendix B presents additional photos from the setup of Experiment 1.

Appendix C presents screenshots from all pages of virtual activities and questionnaires used in Experiment 2, as well as some additional photos of the setup.

Chapter 2

Background Literature

Throughout this chapter, I present background literature about topics related to social learning in animals, humans, and robots. Social learning of animals, in particular primates, inspired many aspects of the present thesis. Therefore, Section 2.1 introduces the basic concepts in social learning of animals, describes the mechanisms involved (Section 2.1.1), and outlines the strategies that guide this kind of learning (Section 2.1.2). Section 2.1.3 details a special form of social learning, i.e., Program-Level Imitation (PLI), which has served as an inspiration for addressing O2. Lastly, to consider alongside the model-based biases observed in the social learning of primates (discussed in Section 2.1.1) when addressing O4, Section 2.1.4 briefly discusses the quality of high-achieving teachers in human-human teaching scenarios.

I then present literature related to the social learning of robots in Section 2.2, from both human-centric and robot-centric points of view. In particular, in Sections 2.2.1 and 2.2.2, I start by discussing the existing literature on how non-expert humans teach robots and how their behaviours while teaching robots are studied, which are topics directly related to O1. Next, in Section 2.2.3, I discuss the state-of-the-art in robot learning systems for learning task structures, a topic that is also the focus in O2. This discussion is followed by a detailed example of a robot learning system designed with biological inspiration in Section 2.2.4, similar to my overall approach in addressing O2. Finally, in Section 2.2.5, I present research that has been focused on enabling robots to identify the most suitable human teachers, a problem that this thesis also aims at addressing under O4.

Before the end of this chapter, in Section 2.3, I briefly discuss Visual Perspective-Taking (VPT) from a cognitive and developmental point of view, and also mention its relevance to HRI. This background information relates to one of the resources provided

to the participants in the experiment that was conducted to address O3. I conclude this chapter by situating the work presented in this thesis in the reviewed literature.

2.1 Social Learning in Animals and Humans

As noted before, many animals, including primates and humans, gain valuable life skills and knowledge from copying others, through a special kind of learning referred to as *social learning* [93]. Social learning is a form of learning that occurs via observation of or interaction with other members of the same species [89]. The development of culture relies on the repeated occurrence of social learning to spread behaviours across a group and/or generations [205]. Social learning stands in contrast to *asocial learning*, or *individual learning*, i.e., learning without social interactions, e.g., by reading a book [89]. Throughout this chapter, I often refer to the learner in the context of social learning as the 'observer' and the teacher as the 'demonstrator' or the 'model.'

A variety of Social Learning Mechanisms (SLMs) are involved in the social learning of animals, as will be detailed in Section 2.1.1. To summarize some important mechanisms, *imitation* is copying another individual, which can include reproducing novel body actions and mimicking tool use or object handling [206]. *Emulation* involves reproducing only the desired outcomes of the actions, to be achieved using either the same method as the model or the observer's own method [204]. *Local and stimulus enhancement* occurs in a social learning interaction when the observer's attention is directed to a relevant location or object by the actions of another individual [206].

In the lifetime of most primates, social learning occurs in three phases:

1. **Learning from the mother or another primary attachment figure**: In many species of primates, mothers care for and carry the infants, and therefore, the mother is the main and often the only model for social learning from the beginning of the lifetime of an infant. In this stage, young individuals attentively watch their mothers' techniques, with a focused and close visual attention called *peering* [173]. In this case, the learner faces the demonstrator and may nod as a sign that he understands the actions.

2. **Selective learning in an expanding social world**: Individuals later start to learn from those identified as better models than their parents for more complex tasks. This stage often involves the use of selective learning heuristics that are referred to



Figure 2.1: An example of a social animal, i.e., Mountain Gorillas, where social learning has been documented [42]. Image: © 2019 iStock, credit: MartaTari [4].

as Social Learning Strategies (SLSs) [105]. A detailed discussion on this topic is provided later in Section 2.1.2.

3. **Learning from residents after the migration**: The third stage of social learning occurs when animals migrate and encounter new social and ecological circumstances [206]. In this stage, since animals are unfamiliar with the local resources, social learning from residents can be helpful. It is common for species to give up their existing personal preferences or behaviours in order to adapt to the new local norms.

2.1.1 Social Learning Mechanisms

Social learning of animals is mediated by several mechanisms, individually or in combination. While different authors categorize these mechanisms differently, I refer to the classification proposed by Hoppitt [94]:

- **Stimulus enhancement**: As described by Heyes (1994) [89], stimulus enhancement happens when observing a demonstrator (or its actions) exposes an observer to a stimulus, and that exposure induces a change in the observer's behaviour. This means that an animal may be more likely to acquire and perform a behaviour related

to a particular object if it watches another animal perform an action with that object. This way, stimulus enhancement can lead to an increased rate of interaction with certain types of objects. To direct the attention of the observer to an object in stimulus enhancement, the demonstrator may act in such a way to make an object visible to the observer or use social cues such as eye gaze.

- **Local enhancement**: Local enhancement occurs in situations when the presence of a demonstrator or its actions done at a particular location increases the likelihood for an observer to visit or interact with objects at that location [93]. Local enhancement refers to the effects of 'locations' rather than 'objects,' as is the case in stimulus enhancement.

- **Observational conditioning**: According to Heyes (1994) [89], observational conditioning happens in case the observation of a demonstrator exposes an observer to a relationship, and that exposure results in a change in the observer's behaviour or response. Through observational conditioning, an observer can match their response to an object or an event with the demonstrator. This is learning how to evaluate objects or events based on their positive or negative values.

- **Response facilitation**: Response facilitation, as defined by Byrne (1994) [38], takes place when watching a demonstrator performing an action, which leads to obtaining a reward in a certain situation, increases the chance that the observer performs the same action. Response facilitation can inform the observer when and where to perform a particular task to be productive or useful.

- **Social facilitation**: Zajonc (1965) [218] defined social facilitation as situations in which an observer's behaviour is affected by the mere presence of a demonstrator. It may be easier for the observer to engage in exploratory or risky behaviours when other individuals are present, as it may feel safer. This way, learning about novel objects or actions can be facilitated.

- **Contextual imitation**: According to Byrne (2002) [40], contextual imitation occurs if watching a demonstrator performing some actions in a particular context makes an observer more likely to perform the same in a similar context. As a result of contextual imitation, the observer can learn to perform a particular action in a specific place, time, or in response to certain stimuli.

- **Production imitation**: Based on a definition by Byrne (2002) [40], production imitation happens when, after watching an observer doing actions that are novel to

an observer's repertoire, the observer becomes more likely to do the same actions later.

- **Observational R-S learning**: Contextual or production imitation can happen in situations when the observer watches the demonstrator 'being rewarded' by doing the actions. In such a case, the observational R-S learning mechanism is involved [89]. This process can lead to learning new behaviours that result in rewards.

- **Emulation**: Emulation occurs in case observing a demonstrator interacting with objects in an environment makes an observer more likely to take some actions that produce the same effect in the environment [53, 190]. In other words, emulation is copying the 'outcomes' of actions done by a demonstrator, rather than copying the 'details' of actions. Emulation can also be used as a more diverse term applied to methods including *object movement re-enactment* (copying how the model makes objects move) and *affordance learning* (understanding properties of the objects or other aspects of the environment to be able to utilize them with their own strategies) [204].

- **Opportunity providing**: In opportunity providing, a demonstrator acts in such a way that an observer will be engaged in learning that could not happen otherwise [93]. Opportunity providing can be accomplished by offering an easier, less dangerous, or more accessible form of the task to the observer, in order to practise and gradually learn. *Scaffolding* is another term that is often used and refers to the same concept, including modifying environmental conditions to assist a learner [39]. This process is typically employed by a mother to facilitate a task for her child, particularly when the child is at a developmental stage.

- **Inadvertent coaching**: Coaching happens when a demonstrator provides a social response to the behaviour of an observer to encourage or discourage that behaviour [93]. In this case, learning may occur in the form of a socially guided trial-and-error.

Theories explaining how social transmission occurs using SLMs have been developed from various perspectives. Hoppitt and Laland (2008) [93] argue that only a few features of a social learning situation determine which mechanisms will be at work, and therefore, propose the classification tree illustrated in Figure 2.2 to identify the set of possible mechanisms in different situations. The first feature that separates sets of social learning mechanisms is whether learning is specific to a certain location, to a particular class of stimuli, or neither (*context specificity*). The second feature dividing possible SLMs is whether or not learning is affected by the results of the demonstrator's actions (*outcome*

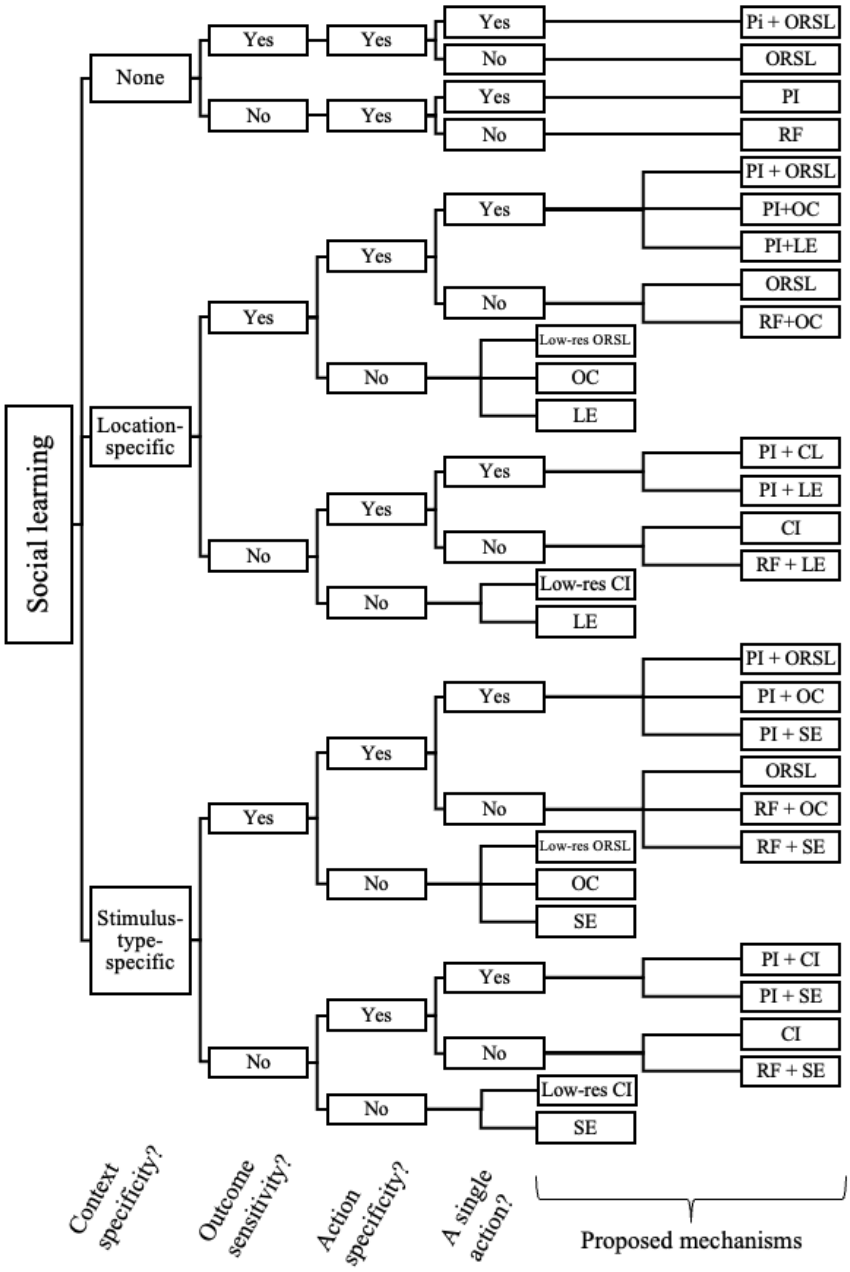


Figure 2.2: A classification tree for social learning mechanisms or a combination of mechanisms involved based on properties of learning, reproduced based on [93]. SE = Stimulus Enhancement, LE = Local Enhancement, OC = Observational Conditioning, RF = Response Facilitation, CI = Contextual Imitation, PI = Production Imitation, ORSL = Observational R-S Learning, Low-res = Low resolution. Zoom in for more visible text.

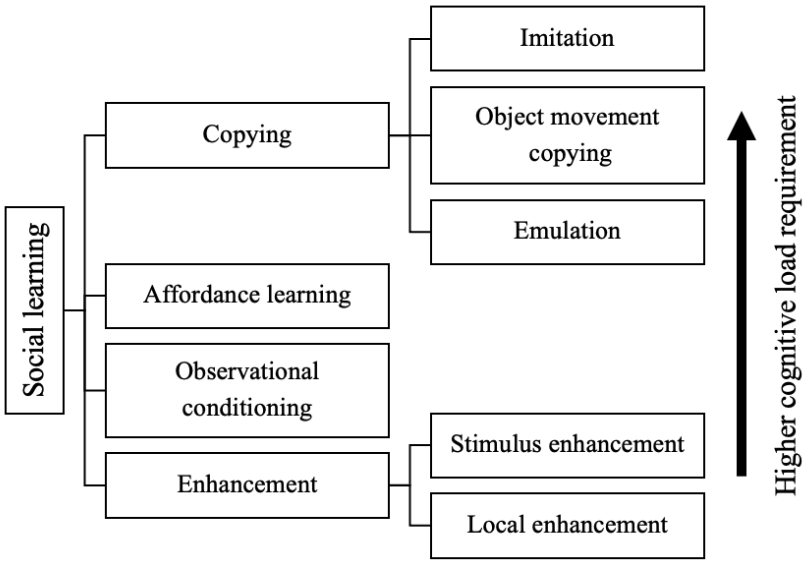


Figure 2.3: A taxonomy of social learning mechanisms based on their required cognitive load, reproduced based on [202]. Note, the mechanisms included in this diagram are not exclusively derived from the classification presented in the thesis.

sensitivity). For example, does learning happen only if the demonstrator is rewarded? For the third division, one needs to determine whether learning is about a specific action or not (*action specificity*). In the case of action-specific learning, the decision tree also considers whether that action is a single novel action or a sequence of actions. Note that the social learning model proposed by Hoppitt and Laland (2008) [93] is intended to help researchers identify the active mechanisms in particular social learning situations, rather than explaining animals' internal learning processes.

Whiten (2021) [202] organizes social processes based on their cognitive load requirement. Figure 2.3 shows an array of social learning processes with increasing cognitive load from bottom to top. At the most elementary level, to achieve local enhancement, only a simple form of associative learning, which is common throughout the animal kingdom, will be required. On the highest level of this scale is copying the complex forms of actions, which can be a cognitively demanding task.

Social learning in animals is often identified and studied via approaches including social learning experiments and mapping of behavioural variation across groups of animals within the same geographical area [205]. Moreover, various computational models, by incorporating ideas from learning theory, cognitive science, and AI have been proposed to describe

how social learning occurs in humans and other animals, e.g., [81, 64, 73]. To provide an example, one view of modelling social learning is offered by Byrne (1998) [42, 31]. This viewpoint assumes that the behaviours of animals are governed by three types of 'brain records': *stimuli*, *responses*, and *goals*. A stimulus is a configuration in the environment that, by learning, can trigger other stimuli or appropriate responses. Actions in an animal's repertoire are considered responses that may follow goals, and there exists an individual *activity level* corresponding to each brain record, which controls the actions of the animal. As a result of watching a demonstrator gaining something useful for survival, social learning mechanisms will likely be activated by activity levels of some brain records being 'primed' (i.e., adjusted) [42, 31]. For example, a location or an object can automatically grab an observer's attention and therefore, prime some brain records. This makes the observer more likely to focus on that location or object, thereby helping it learn. Byrne outlines that this process will lead to the engagement of local or stimulus enhancement mechanisms. The function of the emulation mechanism can be another example: observing a demonstrator achieving some results can prime the brain record of that result as a goal, thereby making the observer more likely to work toward that particular goal by any means. Furthermore, brain records corresponding to a particular response can be primed in an observer's brain and therefore, make it more likely to perform the same type of behaviour. In this case, the response facilitation mechanism has been involved.

A large number of the proposed models for social learning have focused on reinforcement learning, where an observer learns values associated with different actions by using reward prediction errors (see [142] for a review). These models often assume that: (1) values of choices are learned incrementally through iterative trial-and-error, (2) monetary loss or gain reward functions guide the behaviours, and (3) learning happens in a context-insensitive manner [64, 142]. According to FeldmanHall and Nassar (2021) [64], many of the proposed models fail to capture the complexity of social learning scenarios. In their view, modelling social learning is a challenging task since the motives and intentions of a model are highly unpredictable from an observer's perspective, social rewards depend on the needs and motivations of observers and models, and one person's actions can affect the motives and intentions of another. The authors propose a model in which learning outcomes depend on a large number of variables that form internal states of both the observer and the model [64].

2.1.2 Social Learning Strategies

As noted earlier in Section 2.1, animals are selective during their second phase of social learning, e.g., what they copy, when they copy, and from whom they copy, by following

several Social Learning Strategies (SLSs) [105, 31]. Figure 2.4 summarizes a subset of SLSs used by animals. Among all the categories of SLSs, context-dependent SLSs, in which what is learned depends on a given social context, are more relevant to my research and are therefore discussed in this section. Context-dependent SLSs include three main sub-categories, detailed in the following: *when* strategies, describing when individuals copy others, *who* strategies, identifying from whom individuals copy, and *frequency* strategies, explaining the adoption of the most common variants of behaviours.

- ***When* strategies**: Several factors can affect animals' decisions of when to use social information [105]. Theoretically, as the cost of obtaining accurate but expensive personal information through asocial learning increases, the reliance upon gaining less accurate but cheap social information through social learning rises. Individuals will use social information when they are uncertain, either because they have no relevant information as a result of prior personal information being unproductive, unreliable, or outdated, or because the accumulated knowledge of others is more reliable.

- ***Who* strategies**: The evidence for model-based biases in humans and other animals, i.e., a tendency to learn more for some demonstrators, is substantial [105]. Through studying a small-scale society of *Fijian villagers*, Henrich and Broesch (2011) [88] observed some learning biases in the selective phase of social learning. Within the studied population, perceived within-domain success and knowledge were the most important biases. In domains such as fishing, where a potential model's achievements could be directly observed, 'success' seemed far more important than 'knowledge.' When there was no direct evidence of success, 'age' was used as an indirect predictor of the model to learn from. Moreover, a 'cross-domain success' effect was observed, meaning that perceived success in one domain could affect learning in another domain, mostly when two domains were related. It was also found that learners tend to pick models for specific domains from a particular 'sex,' for example, males as fishing models [88].

  Other studies have suggested that primates hold even more biases for selective social learning. For example, some species of brown capuchin monkeys (*Cebus apella*) were found to be understanding the relative proficiency of group-mates and observing the more skilled ones to help maximize their social learning opportunities [144]. In another study about chimpanzees, Kendal et al. (2015) [104] have found that species are also inclined to copy dominant and knowledgeable individuals, who have been trained to succeed. Although these target models may be as successful as the others,

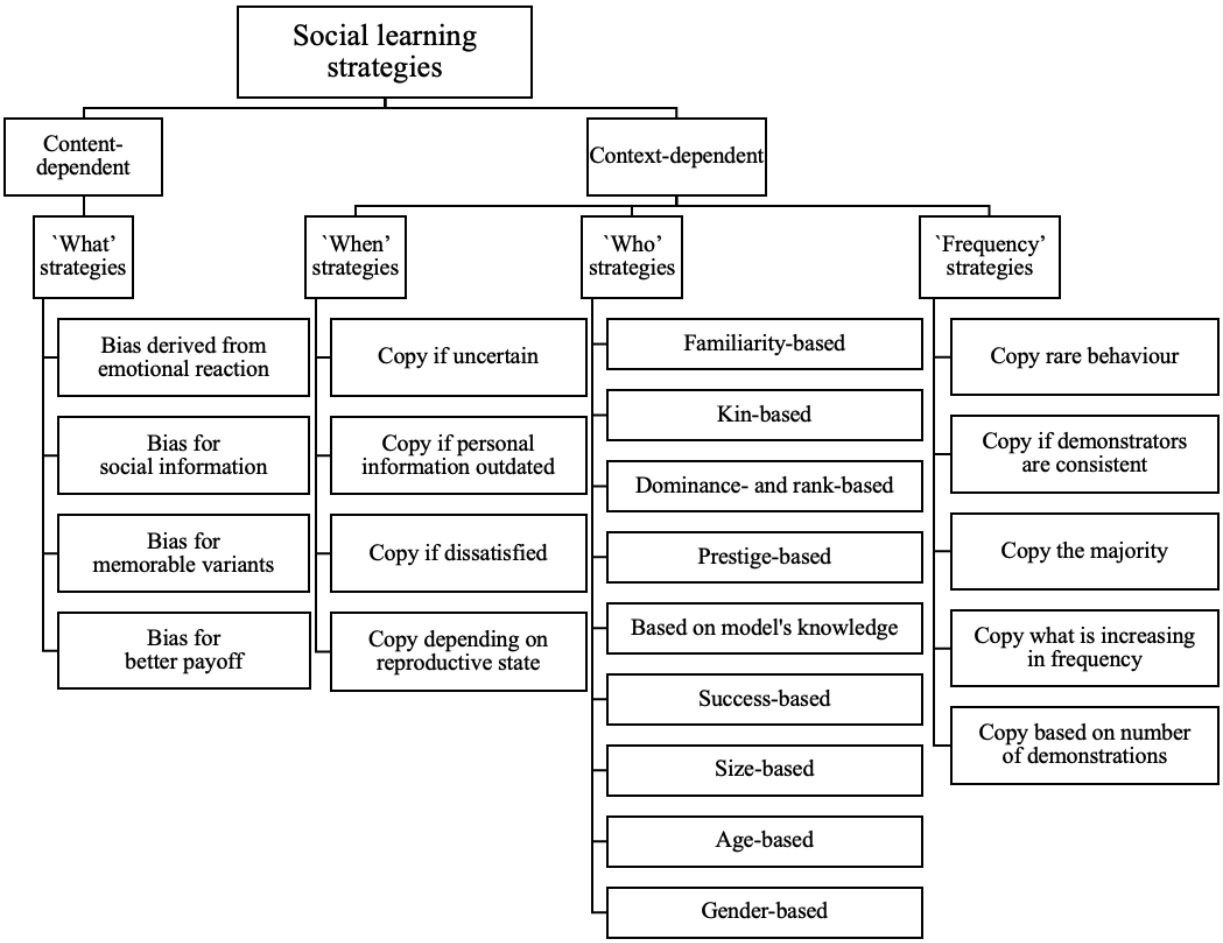


Figure 2.4: A subset of social learning strategies categorization and their uses, reproduced based on [31]. Animals are often selective in what they copy, when they copy, and from whom they copy, and take into account the frequency of the behaviours they observe.

their confident and purposeful way of approaching the tasks is what leads chimpanzees to prefer them.

Research on human children has also shown that infants learn selectively from different adults. Even eight-month-old infants can monitor the reliability of social information sources to adjust their social learning behaviours [193]. Harris and Corriveau (2011) [83] have found that two major heuristics are often used by children to identify which models are more trustworthy than others. First, children exhibit a bias to endorse the claims and imitate the demonstrations of those whom they identify as attachment figures, and/or who have previously proven to be reliable sources of information. To do this, children keep track of the history of potential models. When no recorded experience with another person is available, the second heuristic is used by children. In this case, children are more likely to imitate those who seem to belong to their cultural group. In a study by Birch et al. (2010) [28], two- and three-year-old children's preference for imitation was examined when adult models showed nonverbal cues such as facial expressions and body postures for indicating their level of certainty. It was observed that when both a confident and an uncertain model were available, children selected to learn about objects from the confident one.

Overall, the observer's familiarity with the demonstrator, its social rank, proven knowledge and success, as well as age, gender, and size of the model, can influence social learning. Individuals may copy any aspect of the behaviour of an individual who is of high status, whether or not that trait helped the model attain high status.

- ***Frequency* strategies**: As noted in Figure 2.4, frequency-dependent strategies explain how the most common variant of a given behaviour is more likely to be adopted by others [105]. Animals may become more attentive to behaviours that they frequently observe, and therefore, employ social learning processes to acquire the dominant skill that might be useful for survival.

2.1.3 Program-Level Imitation in Primates

Byrne and Russon (1998) [42] propose that imitation in primates occurs at *action-level* and *program-level*. Basic action details are exactly copied in action-level imitation, whereas in PLI, the overall structures of more complex tasks are implicitly reproduced, and basic actions are arranged into complex behaviours when the animal interacts with the world. In the context of PLI, 'task structures' refer to the hierarchical organization of complex behaviours formed by small basic action units familiar to the animal, whose details will be learned via action-level imitation or trial-and-error. These structures include sequential

regularities in task steps (as reported to be copied by chimpanzees [201] and gorillas [42], see Figure 2.1), coordination of two hands, organization of subroutines, etc. The novelty of PLI is in the arrangement of basic actions into complex behaviours, not in matching the style and details of basic actions.

To give an example, Mountain Gorillas prepare *Galium ruwenzoriense* plants, which are covered in small hooks that can cling to nearby surfaces, to eat as food via a hierarchical process. A gorilla executes a program consisting of the following actions: repetitively picking green strips of galium with one hand, folding loose strands in with the other hand, then holding the bundle loosely and removing debris with the other hand, and finally eating the prepared food. Another example program learned by Mountain Gorillas is processing nettle (*Laportea alatipes*) plants, which have painful stings, to also eat as food. The hierarchical process of this behaviour is presented in Figure 2.5(A). Byrne and Russon [42] believe that such behaviours are learned, *"because these techniques are found only in a small population of gorillas, and indeed are valueless outside the very limited altitudinal zones in which temperate plants like nettle and bedstraw grow in Africa."*

Animals parse the behaviours they repeatedly observe to achieve PLI. When a mountain gorilla processes *Galium ruwenzoriense* plants, an observer will identify important sub-goals within the action sequences. These sub-goals are states that consistently appear across all successful attempts and, therefore, appear to be essential for the ultimate task goal. However, the specific actions used to reach these sub-goals may vary each time based on the situation and the initial state, since each plant is unique in form (see Figure 2.5(B)). Animals repeatedly watching models identify sub-goals in the process by learning from statistical regularities [41]. For example, a gorilla might notice that his mother 'always' folds a bundle of leaf blades that protrudes from one hand with another hand and holds down this folded bundle with her thumb. Similarly, the gorilla may observe that debris removal 'always' occurs in the same location within the sequence after certain sub-goals. By recognizing such consistent patterns, the animal learns which actions are crucial to success, which actions are not, and in what order they must be performed.

Humans also imitate hierarchical structures in complex actions. Whiten et al. (2006) [203] examined this by demonstrating two approaches to opening a puzzle box to three-year-old children. Children incorporated a new object that was absent in demonstrations but present at the test into their actions, based on the specific hierarchical structure each child had observed. Moreover, children tended not to follow the specific left-to-right or right-to-left sequence of actions, which supported the idea that they did not merely follow the serial order of observed actions. Studies of the human brain, e.g. [78, 99, 10], also support the existence of hierarchies for action planning and execution in human goal-oriented activity.

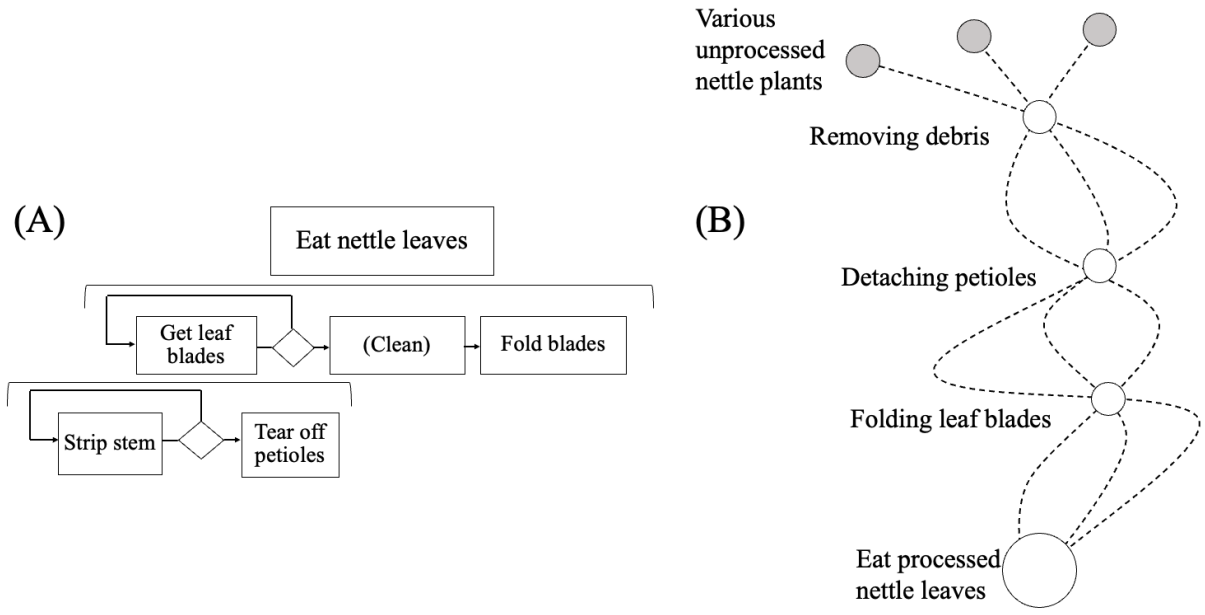


Figure 2.5: (A) Hierarchical organization of processing nettle leaves (*Laportea alatipes*) learned by mountain gorillas. There are two independent modules of 'getting leaf blades' and 'folding blades,' occasionally separated by the third optional action of 'cleaning' in this program. 'Getting leaf blades' may be iterated and is made up of two components: 'striping stem,' which may also be iterated, and 'tearing off petioles.' (B) Processing three nettle plants can be represented as paths (shown with dashed lines) in a two-dimensional space. Because plants are different in form, each process has a unique starting point. All the processes lead to the same goal, i.e., eating processed nettle leaves. Some sub-goals are necessary for the eventual success of this process and occur in all examples, as noted where dashed lines converge. Figures are reproduced based on [41].

2.1.4 Teacher's Awareness of the Capabilities of Students

The explicit presentation of generalizable knowledge by a model (i.e., teacher), followed by an interpretation by a learner (i.e., student), is referred to as *pedagogy* [51]. According to Csibra and Gergely (2006) [51] transfer of knowledge from one human to another through explicit teaching is a specific type of social learning which takes place using a specific form of communication. Similar to other previously discussed SLMs, pedagogy conveys knowledge or skills that are generalizable beyond the present context. However, unlike most of SLMs, pedagogy does not rely primarily on observation. Teaching requires active participation and effort by the teacher, often involving cognitive, temporal, and energetic costs, since the act of teaching may not benefit the teacher themselves.

Due to the importance of teacher quality in the learning outcomes of students [160], several characteristics have been proposed to describe a good teacher. Effective and constructive behaviours, sufficient knowledge of the subject matter, a constructivist view of teaching, a strong and coherent professional identity, effective communication skills, offering constructive feedback, commitment, creativity, courage, care, flexibility, kindness, fairness, empathy, and tolerance are all among the qualities of a good teacher [115, 23].

What inspired the research presented in this thesis is that research on human-human teaching has highlighted the importance of teachers' knowledge and awareness of students, including knowledge about misconceptions, their interests, and easy and difficult topics for students, for improving their learning outcomes [91]. Such awareness can greatly benefit instructors, e.g., by letting them design instruction to target common student confusions and mistakes [183, 90]. For instance, in teaching mathematical concepts, the knowledge of teachers about their students' problem-solving knowledge, obtained via regular assessments of students, has been found to be positively correlated with the development of problem-solving abilities in the learners [151].

The concept of Zone of Proximal Development (ZPD) emphasizes that effective teaching requires that teachers understand the gap between what learners can do independently and what they can achieve with guidance so that the given instructions and tasks go not too far beyond the learner's current competence, and can build on learners' existing abilities [176]. ZPD implies that an adult must know what the child is attempting to achieve and be sensitive to the child's abilities and signals to help a child learn effectively [128].

2.2 Social Learning and Robots

Among the different definitions and categories for social robots, my research centres on *socially intelligent robots* (in the form of humanoid robots), which are robots with the ability to exhibit aspects of human-style intelligence based on models of human cognition and social competence, and can have adaptive functionalities [55]. With the advancement of technology in recent years, robots and software agents also use social learning to allow non-expert users to teach new skills to such systems for their personalized use [37]. For example, as noted earlier in the thesis, imitation learning is one of the robot social learning mechanisms that makes the transfer of skill from humans to robots more efficient [36, 27], especially for robots with many degrees of freedom and when the state-action space is large. By observing a demonstration or experiencing a sequence of actions, imitation learning provides a quick and easy way for people without any extensive knowledge of programming to teach a robot how to do complex actions.

To let robots dynamically acquire new task skills, several robot learning frameworks have been developed, e.g. [98, 45, 46, 196, 50, 219]. Recent reviews on robot learning and adaptation, including the usage of modern imitation learning methods, can be found in [119, 212, 155, 221, 66], which may be useful for readers seeking context beyond the focus of this thesis. The ability of the robots to learn new skills is often considered a major factor in cognitive architectures such as HAMMER [58], because achieving human-level AI requires systems that can adapt to new situations and tasks [116].

Robot learning frameworks often attempt to make the teaching process as natural as possible [155]. *Kinesthetic teaching*, a technique whereby a human physically guides a robot through a task, is one of the natural and intuitive modalities used for teaching robots [45], e.g., in [60, 13, 207, 182]. In some other robot learning systems, e.g., [156, 102, 177], teachers' behaviours can be remotely recognized by the robot, e.g., using vision sensory inputs, motion-capturing devices, or even videos. Note that learning by observing humans is important for robots and is also challenging, because potential obstructions, fast movements, unfamiliar actions, and sensor noise may prevent demonstrations from being visible to robots' vision systems [155]. Natural language has also been used as a modality for providing task examples to robots, e.g., in [184, 194, 178].

Imitation learning requires a correspondence between the observer and the demonstrator [140]. If both the demonstrator and the observer share similar embodiments, this correspondence involves only mapping the equal parts of the bodies. Nevertheless, in human-robot teaching, a human teacher and a robot are likely to differ in the number and/or type of joints, the length of the links, and the amount of force/torque they can apply, and therefore, exact behaviours cannot be copied. Solving the correspondence problem

is a challenge in the social learning of robots. Kinesthetic teaching, which allows learning via internal observation, is an example of an effective solution to this problem [140].

Studies have shown that interacting with others is a crucial component of lifelong social learning in humans and several species of primates [206]. Mechanisms to support teaching robots through social feedback and interaction are therefore crucial for successful social learning of robots [209, 113]. In a human-robot teaching interaction, a robot must be able to interpret the social cues in human teachers' behaviours and instructions, and express itself in a manner that the teacher understands [121]. With a two-way flow of information between the human teacher and the learner robot, feedback from the learner robot can increase teachers' awareness of the learning state and improve learning outcomes [33]. A recent study proposed and tested a metric called Mental Model Mismatch Score as a feedback mechanism to improve the alignment between human teaching intentions and what the robot actually learns about how to solve a puzzle [157]. The usage of a robot's gaze has also been found effective in improving teachers' understanding of a learner robot's progress and capability when learning a decluttering task [96].

2.2.1 Non-expert Humans Teaching Robots

As noted in Chapter 1, human demonstrations for teaching robots, particularly those provided by non-experts, may not always be optimal. There are several possible causes of sub-optimal performance in teaching robots by demonstration. To begin with, non-expert users may exhibit unnecessary, superfluous actions that do not contribute to achieving the ultimate goal of the task when providing demonstrations to a robot [101]. Such actions, if learned, may lead to over-imitation (see [92] for a review[1]). Second, incorrect or unsuccessful teaching attempts may happen that could impact the effectiveness of the demonstrations [101]. As a side note, a system proposed by Grollman and Billard (2011) [79] can use those failed demonstrations to determine what went wrong and to avoid repeating the teacher's mistake. Moreover, there may be unmotivated actions in the demonstration that cannot be learned by the robot due to the lack of sensors for recording relevant data [101]. Lastly, a demonstrated trajectory might not be safe for the robot, e.g., it could cause self-collisions (i.e., different parts of the robot colliding with each other during the demonstration), or the trajectory could contain singularity points (e.g., when a few links of a robotic arm are aligned so that the robot is unable to move any further) [67]. Another

[1]Hoehl et al. (2019) [92] define over-imitation as *"imitation of perceivably causally unnecessary actions in relation to the goal of an action sequence performed by a model,"* or in other words, adding extra actions when copying observed actions or action sequences for achieving a given goal.

example of an unsafe demonstration is applying too much pressure when holding or moving robot parts that might harm the robot [67].

Focusing on the kinesthetic teaching of a robotic arm, Sakr et al. (2020) [163] tested three training methods intended to improve non-expert human demonstrations: (1) *Discovery* method, in which participants could learn to teach the robot through trial-and-error followed by evaluating their previous demonstrations, (2) *Observational* method, when participants watched an expert teaching the robot, and (3) *Kinesthetic* method, which involved participants holding the robot's gripper passively while a trajectory taught by an expert is being performed by the robot. By comparing the performance of participants teaching a robotic arm how to press a few buttons placed in holes, prior to and immediately after training, the researchers found that all three tested training methods improved human teachers' performance [163]. In particular, the observational method was the most effective approach for reducing the teaching time [163]. In another study, Orendt et al. (2016) [143] evaluated a kinesthetic teaching system in terms of ease of use and robustness in task completion. They compared brief verbal instructions with a graphical tutorial when instructing two groups of users, consisting of both robotics experts and non-experts. These two types of instruction, however, did not affect teaching success [143].

Some computational approaches have also been used to improve humans' teaching of robots and facilitate their learning from non-expert users [111, 210, 44, 101]. For instance, Cakmak and Thomaz (2014) [44] developed a system that enabled a robot to give real-time instructions to participants when teaching it a classification task, and to guide the participants toward providing more effective examples. The framework developed by Schrum et al. (2023) [172] enabled a driving simulator to produce feedback from the system to improve the quality of sub-optimal human teachers' demonstrations. Sakr et al. (2025) [165] proposed and tested a method using a measure of information entropy to help non-expert human teachers, who needed to wear an augmented reality headset, identify areas of a trajectory where a robot is uncertain. This method of guiding teachers helped participants provide better kinesthetic demonstrations, which enhanced the generalization efficiency of a robotic arm. The researchers also found that human teachers who had undergone training and gained experience providing demonstrations related to a specific domain could later demonstrate more useful examples in a new task domain [165, 138].

The performance of participants as teachers for robots has been evaluated in the literature reviewed above with a number of metrics. These include the success of guiding the robot through the task [43], time spent on providing the demonstrations [44, 163, 43], the accuracy of learning, if the system was capable of actually learning from the participants [44], path-length and smoothness of the demonstrations [163], g motion-related features such as manipulability and joint-space jerk [26], how strongly teachers grasp ob-

jects and how consistently they perform teaching [145], and, in case participants could ask for help, the number of help requests [44, 43]. Note, there has been some work specifically on quantifying teaching behaviour in robot learning scenarios, but mainly focused on the use of machine learning (e.g., [175, 164]). Sakr et al. (2022) [164] conducted a two-session experiment to validate their proposed computational framework for assessing the quality of kinesthetic demonstrations. In their experiment, 12 out of 27 participants (called 'fast-adapters') consistently provided high-quality demonstrations from the start. However, it took some time and practice for the rest of the participants (referred to as 'slow-adapters') to adapt to robot teaching so that they provided high-quality demonstrations only in the second session [164].

2.2.2 Human Teachers' Behaviour When Teaching Robots

A number of human-centred studies have explored aspects of participants' behaviours and mental models when teaching robots. For example, Kaochar et al. (2011) [103] studied teaching styles and behavioural patterns of human teachers for robots. Participants used a computer-based interactive teaching framework for teaching an unmanned aerial vehicle in a simulated environment by giving demonstrations of actions. The majority of human teachers used all three available modes for teaching the robot: labelling objects, defining procedures, and providing feedback. Researchers then identified different patterns used by participants to switch between teaching and testing of the robot, as well as different teaching styles, depending on how organized the teacher's instructions were [103]. Khan et al. (2011) [107] also investigated human teaching strategies, in the context of teaching a robot whether an object is graspable or not, by assigning binary labels to pictures of the objects. The participants in their study were found to employ three major teaching approaches: (a) starting from extremely graspable or non-graspable objects and moving towards marginal examples, (b) teaching objects as per their original order, or (c) focusing only on graspable examples [107]. Strategies of human participants when teaching a simulated agent how to sort some blocks, particularly in terms of the communication of intent and execution of actions, have also been analyzed by Ramaraj et al. (2021) [154]. This study found that participants were teaching the system new concepts by instantiating and/or describing them, expanding or correcting previous knowledge of the system, evaluating its learning progress, and revising their own instructions. In another work, researchers have studied participants' emotional experience and their perception of the robot and compared teaching some basic arm movement tasks visually to a robotic character with watching the robot self-learning [62]. Compared to watching the robot self-learning, teaching the robot boosted participants' perception of trust, but did not affect their perception of the robot's

competency. Hedlund et al. (2021) [87] investigated human teachers' reactions to robot failure, mainly in terms of trust and workload, when participants employed kinesthetic teaching, tele-operation, or motion capture for robot teaching. Participants rated motion capture as the least difficult of the three instructional methods and tele-operation as the most difficult, in terms of workload [87]. Kim et al. (2009) [109] have also studied how non-experts provide voice inputs to teach a robot.

Still, there exist only a few studies that have taken into account the actual kinesthetic teaching of a robot for exploring human teachers' behaviours. In social interactions, eye gaze is an important nonverbal cue that can be used to infer others' intentions [63]. Saran et al. (2019) [168] have studied gaze patterns of novice and expert human participants during kinesthetic demonstration of two tasks on a robotic manipulator: pouring pasta from a cup into an empty bowl and placing a spoon on a table with respect to other objects. According to their results, both groups of users tended to fixate almost entirely on objects being manipulated or objects with which manipulation occurred during teaching. However, the gaze pattern of users with varying robot experiences showed some differences: novice robot users attended more to the robot's gripper than expert users, possibly because they often struggled to manipulate the robot's arm [168]. Saran et al. (2022) [167] later investigated the speech patterns of human teachers during kinesthetic or video-based teaching of multi-step manipulation tasks to a robotic manipulator, namely box-opening and fruit-cutting tasks. Participants were asked to include speech in their instructions, either explicitly or implicitly. It was found that users conveyed similar semantic concepts through spoken words regardless of the modality of teaching. However, participants as teachers were more expressive during kinesthetic teaching, but talked less often when performing actions themselves for video-based teaching [167].

2.2.3 Robots Learning Task Structures

A growing body of research has focused on enabling robots to obtain a model of the underlying concepts of complex tasks and understand task structures instead of just mimicking demonstrations [47, 131, 171, 138]. Previous work used decision trees to infer action sequences to perform tasks in a structured way [170, 166]. Lázaro-Gredilla et al. (2019) [134] developed a probabilistic approach for robots to generate task programs using spatial primitives to arrange objects on a tabletop to a desired state, e.g., in a circle. Xu et al. (2018) [213] proposed a framework to break down robot manipulation tasks into smaller sub-tasks. Inferring hierarchies and structured representations has also gained attention in learning trajectories, e.g., in [118, 188, 224].

Uhde et al. (2020) [195] developed a method for identifying dependencies in human actions, particularly if one action should be performed before a second action. Using a dataset of three simple household activities performed in various settings in Virtual Reality (VR), sub-goals of 'putting objects on the table' and 'opening or closing doors and drawers' were detected in the performed actions, by identifying pre-defined patterns in the observed movements in the VR environment. To determine if sub-goal e_1 is 'instrumental' for another sub-goal e_2, the system identified candidate pairs in which e_1 has occurred before e_2 more frequently than a certain threshold as opposed to the reversed order in the dataset. To validate candidates, a robot in a VR setting mentally simulated performing the two sub-goals to ensure: (1) e_2 is not possible without e_1 executed before and (2) e_1 is possible with e_2 executed before (note, other applications of mental simulation for robots include e.g. posture selection [139] and perspective taking [69]). A simpler approach for understanding constraints in ordering sub-goals through human demonstrations was implemented for robots to learn and perform pick-and-place tasks [61]. In this approach, all sub-goal pairs were considered in the order they were demonstrated, and then contradicting pairs, i.e., when an opposite order was observed, were removed.

Summers-Stay et al. (2012) [185] designed a system that uses human demonstrations to generate tree structures describing important aspects of an activity. This system operates based on 'action grammars' [147], where activities are viewed as either bringing two objects together with a tool and creating a new object or tool, or using a tool with an object and creating a transformed object. Computer vision techniques were used for hand state determination and object recognition from videos of participants performing In cooking and crafting, an activity could be outlined by considering what objects or tools were grasped by each hand and also when objects were brought together and began to be treated as a single object. When one of the hands of a demonstrator (or tools held in the hands) came into contact with a new object, a new leaf node appeared in the tree. When objects were brought together, a new node was created with each of the original object nodes as its children.

More recently, a similar approach has been tested with robots for understanding and execution of actions [215, 214]. A method called 'Functional Object-Oriented Network (FOON)' is also proposed to model the connectivity of objects and actions in manipulative tasks observed in videos [149], FOON creates task tree graphs containing structured information about the relationship between the object states and available actions. Nodes in such graphs are either actions or objects having states. Objects interacting with actions are connected through edges. and has been applied to robots for task tree extraction and action planning for cooking tasks [148, 162].

Lee et al. (2013) [123] have proposed a method that enables robots to capture high-level,

reusable structures of tasks using probabilities of meaningful patterns occurring within action components. Such knowledge could help a robot accurately process and imitate sequences of more complex actions that share the same structures, in case demonstrations contain errors due to environmental noise or teaching mistakes. This approach was tested in real-world settings when an iCub robot imitated solving a Towers of Hanoi puzzle and imitating dance movements [123].

Several recent studies have focused on utilizing the information generated by a Large Language Model (LLM) to help agents plan and perform tasks, e.g., [181, 122, 200, 97, 12, 125, 126]. To elaborate on one example, Ahn et al. (2022) [12] trained an LLM via prompt engineering to produce sequences of low-level skills already in a robot's repertoire, to perform physical tasks. When a robot is asked to perform a task, the LLM indicates step by step how likely each skill contributes to completing the task. These probabilities will then be combined with the likelihood of each skill's success in the current situation, derived from an 'affordance function' trained using reinforcement learning. This mechanism ensures that LLM-generated task plans are appropriate for the physical environment and the robot's capabilities.

2.2.4 An Example Biologically-Inspired Robot Learning System

Saunders et al. (2007) [170] have designed a robot learning system that uses the principles of *self-imitation* and *environmental scaffolding*, derived from animal and human psychology. Self-imitation, i.e., matching behaviour that is observed internally rather than externally, helps to avoid the correspondence problem. The process of scaffolding allows the learning system to increase learning outcomes once the teacher simplifies the environment, making it easier for the robot to learn.

For the learning mechanism, k-Nearest Neighbours (k-NN) algorithm was used that utilized weighted Manhattan distances between vectors containing robot sensor readings (i.e., state vectors, including e.g., joint positions, target location in the camera, sonar values, etc.) and primitives of the robot (can also be combined into sequences, tasks, and behaviours as will be described). By remotely controlling the robot and initiating simple robot primitives using a computer interface, a teacher could 'put the robot through' actions and perform teaching. At the same time, that data could be recorded and saved in the robot's memory as a new competency [170].

The designed scaffolding mechanism allows the robot to reduce the impact of irrelevant features in the recorded state vectors when the learning environment is simplified by the teacher. In that case, an information-gain approach would adjust the weights of

unimportant features. This aspect of the learning system was directly inspired by animal behaviours. As noted earlier under 'opportunity providing' as one of the SLMs, the concept of scaffolding, in which environmental conditions are modified in order to assist a learner, has been well observed in animals.

In learning mode, new competencies could be saved as sequences of primitives with no references to the environment, or goal-directed tasks, by taking internal and external states into account. A set of tasks, sequences, and primitives could be collected as a behaviour, with or without a goal. The system was also capable of predicting actions and matching them with the provided demonstrations, thereby indicating to the teacher during teaching that the robot has already mastered the task. The authors have examined the robot learning system by asking a teacher to teach the robot how to follow an object and point to it. The researcher used a robot that had an arm installed on a Pioneer mobile platform, equipped with a camera. Although the training was long (more than 3 hours), the robot successfully mastered the skill [170]. This learning framework has also been tested on the Khepera mobile robot [169].

Another example of a biologically-inspired robot learning system can be found in [57, 58].

2.2.5 Robots Identifying Suitable Human Teachers

To date, only a few studies have attempted to address the question of 'who to imitate' in LfD. Considering that in a normal classroom, teacher *engagement* can be a determining factor leading to a more positive teacher-student relationship and to more fruitful student achievements, Novanda (2017) [141] proposed using engagement as a factor to find better teachers for robots. A scale for teacher engagement has been developed, which includes the factors of cognitive, emotional, and social engagement [110]. Moreover, the concepts of attention, involvement, interest, immersion, rapport, empathy, and stance have been widely mentioned related to engagement, in the human-agent interaction literature [75]. Attention, as a vital aspect of engagement, controls the orientation of the senses towards incentives that are relevant to the engagement. Because the metrics related to engagement are mostly subjective and hard to measure from a robot's perspective, Novanda (2017) [141] has proposed measuring verbal and nonverbal immediacy in human teachers for robots as an alternative to engagement. In human-human interaction, a number of behavioural cues, including vocal cues, which themselves consist of pitch and tone of the voice, tempo, and volume, are often used by people to increase listener engagement. Increased volume improves instructor clarity. Teacher eye contact with students is also another indicator of

immediacy in classrooms, and finally, head nodding of the instructor gives the students a sense of understanding and can have a positive effect on their reactions. As evidence of their commitment to students' success, engaged teachers show a willingness to devote more time to their duties, pay extra attention to students, and stay more patient [161].

To study how the above-mentioned factors can be used to recognize a qualified teacher, the research group conducted a set of experiments using a Kaspar humanoid robot [141]. In those experiments, participants once acted as teachers and then acted as human evaluators of the taught material. Evaluators rated the videos recorded from the demonstrations made by teachers in terms of teacher engagement in real time. Then, the researcher attempted to establish relations between the rated engagement and the measured behavioural data by the robot. General conclusions were hard to make, as it was difficult to discover which factors contributed the most to the dynamics of the engagement rating.

Cho and Jo (2013) [50] have developed an online learning framework in which the robot autonomously decides to accept or neglect the demonstration in each new teaching trial. In that algorithm, the robot judges the appropriateness of the new demonstrations and learns the behaviour only on suitable teaching trials based on the projection of motion data to a latent space and the description of motion data in a Gaussian mixture model. More recently, Sena and Howard (2019) [175] have introduced a model related to the domain of algorithmic machine learning for learning from demonstration, which includes the teacher's understanding of, and also influence on, the learner. They added a new space called "The teacher's belief space" to the standard model to consider the teacher's contribution to learning and also give the teachers feedback to guide them toward better teaching practice.

2.3 Visual Perspective-Taking in Humans and Robots

Visual Perspective-Taking (VPT) is a feature of human cognition for understanding a visual scene from one's own perspective and from other locations in the physical space [135]. VPT can be described as an aspect of Theory of Mind (ToM), which refers to our ability to understand that others have their own mental states, including thoughts, goals, emotions, and intentions [124, 153]. VPT relies on several cognitive mechanisms, including the activation of left lateral prefrontal cortices for inhibiting one's own irrelevant perspective and bilateral parietal and temporo-parietal areas for the representation of different perspectives [174]. Two levels of VPT are developed sequentially in human children: Level 1, understanding whether or not an object is in another person's line of sight, which is usually developed in children by around the age of 2; and Level 2, the ability to know

that an object may visually appear different for another person compared to one's own perspective, which is a more complex ability seen in children over the age of 4 or 5 [82].

To date, VPT has been employed in several HRI studies, e.g., [192, 35, 208, 69, 68, 220, 11, 100]. Breazeal et al. (2006) [35] have utilized the ability of VPT in a learner robot to facilitate social learning by enabling the robot to focus on areas of demonstrations that appear more important to the teacher. In their proposed learning architecture, a robot could learn by inferring a human teacher's actual intent based on their visual perspective, even if the teacher's demonstrations were not optimal [35]. Researchers have also developed games with a Kaspar robot in which children with autism show objects to the robot and receive feedback to help children explore elements that are important in developing VPT skills [208, 220]. In those games, a screen displaying the robot's field of view was placed next to the robot to let children explore what would happen if components of the game became invisible to the robot. Fischer and Demiris (2016) [68] have created a computational framework to allow VPT at both levels in a humanoid robot. Using this framework, the robot tracks the line of sight of the user to determine objects' visibility from their perspective (VPT Level 1) and also mentally rotates objects to estimate how objects would appear from the visual perspective of the users (VPT Level 2) [68].

A number of studies have tested the ability of VPT in human interaction partners in HRI scenarios through experimental approaches that often involved asking participants to interpret letters (e.g., [52]) or numbers (e.g., [223]), or count the number of visible dots (e.g., [199, 211]), and then determining whether their response matched their own perspective or the robot's perspective. For instance, research has shown that humans naturally consider taking the visual perspective of a robot when interpreting a number placed between the observer and the robot as either a 6 or a 9, improving with increased human-likeness of the robot and the presence of goal-directed actions towards that number [223, 222]. Humans were also found to take the perspective of a robot more often, in case the robot exhibits more interactive behaviours, including eye gaze shifts and hand gestures [216].

2.4 Positioning My Work Within the Literature

The experiment presented in Chapter 3 studies potential changes in the kinesthetic teaching of non-expert participants as they gradually become more experienced robot teachers, with no external or formal training provided. This approach is comparable to the discovery method used in [163] in terms of allowing for trial-and-error; however, my study includes repeated interactions over multiple sessions and uses a naturalistic unsupervised setup

without external intervention into how teaching should occur, without even the presence of an experimenter. This experiment did not include asking participants to self-reflect on their performance, as done in [163], or undergo training as done in the recent study that highlighted the positive impact of experience on user demonstrations [165].

My research also offers a more comprehensive analysis of different aspects of participants' behaviour during kinesthetic robot teaching, compared to the literature reviewed in Section 2.2.2, and identifies different clusters of non-expert teachers, to help address the unresolved problem of 'who to imitate' from a robot's perspective. This includes an analysis of strategies used for manipulating the robot's arm (specifically, the placement of participants' hands on the arm of the robot when physically guiding the motions) and human teachers' workload, as well as an investigation of their gaze behaviour during kinesthetic teaching tasks, which was also conducted in [168]. Compared to [164], my research focuses specifically on non-expert participants and studies broader aspects of their performance and behaviours over more human-robot teaching sessions. The general approach to evaluating the success of participants guiding the robot through the tasks and time spent on providing the demonstrations is based on [44, 163, 43, 138].

The system that I present as part of research presented in Chapter 4 is biologically inspired, as was the case in the example presented in Section 2.2.4, but this research is motivated by imitation learning and PLI in primates among all different SLMs. The proposed system functions based on inputs of human teachers to a physical humanoid robot in the real world, which may include errors and contradictory examples, as opposed to learning offline from a large dataset, e.g., in [195]. My approach takes into consideration more complex relationships between sub-goals, e.g., one sub-goal may block or obviate the need for another sub-goal, compared to temporal relationships discussed in [195, 61]. It was noted that LLM-based approaches to high-level task planning can provide substantial benefits for robots (e.g., minimizing additional training as language models may use their own knowledge to plan tasks, avoiding sub-goal detection in demonstrations, etc.). Compared to LLM-based approaches, e.g. [12], and also methods that learn and use task structures, e.g., [123], the key aspect of my proposed LfD approach is that the robot identifies 'multiple' possible sequences of actions to achieve an ultimate goal, and applies this knowledge to effectively select and execute the most suitable action sequence considering user-defined criteria and the present physical task setting.

As noted earlier, some effort has been made to date to address the question of 'who to imitate' in imitation learning of robots. For example, a teacher's engagement has been proposed as a metric to define a good teacher for robots [141]. However, there is still no conclusive answer to the problem. To contribute to this problem, I first considered '*who* strategies' used by animals, noted in Section 2.1.2, but found that those strategies may

not be relevant to robots as explained in Section 1.2.3. I therefore looked deeper into what makes someone a good teacher in human-human teaching interactions, described in Section 2.1.4. To the best of my knowledge, no previous work has focused on the impact of human teachers' awareness of 'inherent limitations' of learner robots on the learning performance of a robot in the physical world, as is studied in Chapter 5 of this thesis. This experiment provided human teachers for a robot with an opportunity to take the visual perspective of the learner robot, by placing a screen showing the robot's view next to the robot, as was done in [208, 220]; however, to the best of my knowledge, no previous study has utilized VPT in humans in the context of teaching robots physical household tasks.

Chapter 3

Experiment 1: How Non-experts Kinesthetically Teach a Robot over Multiple Sessions

In real-world applications, robots should adapt to users and environments by learning from humans; however, users may not know how to teach new tasks to a robot. The experiment reported in this chapter (Experiment 1) investigates whether participants without any experience in teaching a robot would become more proficient robot teachers through repeated kinesthetic human-robot teaching interactions. In this experiment, 28 non-expert participants were asked to kinesthetically teach a humanoid robot different cleaning tasks in five repeated sessions, each session including four tasks. Participants in this study are not formally instructed on how to perform the teaching or how to refine their demonstrations. Throughout the experimental sessions, participants' gaze patterns, methods of manipulating the robot's arm, their perceived workload, and some physical properties of the demonstrated actions were measured and then analyzed. As will be discussed in this chapter, the data analyses revealed a diversity in non-experts' human-robot teaching styles in repeated interactions.

The current chapter outlines the research questions explored in this experiment (Section 3.1), details the methodology when conducting the experiment and analyzing the data (Section 3.2), presents the findings (Section 3.3), discusses how the results contribute to answering the research questions (Section 3.4), explains the limitations specific to this experiment (Section 3.5), and proposes future research directions to address them (Section 3.6). A summary of the research findings is also presented in Section 3.7 together with some concluding remarks.

3.1 Research Questions

The experiment presented in this chapter explores the following research questions related to two main research objectives (O1 and O4) in this thesis. Each of the sub-questions focuses on a certain aspect of participants' behaviour or a particular characteristic of their demonstrations when non-expert humans kinesthetically teach a robot.

RQ1 Over multiple sessions of non-experts kinesthetically teaching cleaning tasks to a humanoid robot,

- **RQ1.1** how does the **success of teaching** change over time?
- **RQ1.2** how does the **teaching duration** change over time?
- **RQ1.3** how do **participants' gaze patterns** change over time?
- **RQ1.4** how does the **method used for manipulating the robot's arm** change over time?
- **RQ1.5** how do **different aspects of perceived workload** change over time?

RQ2 Does data from non-experts kinesthetically teaching a robot indicate any pattern to find better teachers for the robot?

The study is exploratory and does not test any formal hypotheses based on previous literature with regard to the RQs, due to lack of studies, except for RQ1.3. For RQ1.3, we[2] hypothesize that as participants become more experienced in robot teaching, they gaze less often at the robot's hand (H1), in line with the study of Saran et al. (2019) [168], as discussed in Chapter 2.

RQ1 is related to O1 as it focuses on understanding how multiple aspects of non-expert humans' behaviours and their demonstrations change when they teach some tasks to a humanoid robot multiple times. RQ2, which asks for patterns in non-experts' teaching that a robot can use to identify the most suitable teachers for its kinesthetic learning, is also related to O4.

[2]Using "we/our/us" in this thesis as a matter of academic style, however, the work presented in this thesis is my own.

3.2 Methodology

The experiment reported in this article involved repeated sessions where study participants attended five in-person sessions each. In all sessions, participants kinesthetically guided a Pepper humanoid robot [5] through four cleaning tasks by grasping onto and physically guiding the motions of its left arm. 'Cleaning' tasks were selected among other types of tasks because they are part of almost everyone's daily routine, so most people are familiar with them. In Tasks 1 and 2, Pepper asked the participants to teach it how to clear different pieces/particles from cutting boards fixed onto a table. The next two tasks, Tasks 3 and 4, were to teach Pepper how to clear different whiteboard marker patterns from specific areas of a whiteboard. These tasks are illustrated in Figure 3.1. A sponge was attached to Pepper's left hand for cleaning the surfaces (see Figure 3.2). The areas to be cleaned were expanded in the second task of cleaning the table and the whiteboard, i.e., in Tasks 2 and 4, to introduce additional challenges. The size and the positioning of the cutting boards and the areas in the whiteboard were chosen to be within the workspace of the robot's left arm and were fixed for all participants. In Tasks 2 and 4, compared to Tasks 1 and 3, those areas were extended to the right side and included areas that were harder to reach for the robot's left hand. This required participants to use both hands when teaching Tasks 2 and 4 in order to adjust at least two Degrees of Freedom (DOFs) in the robot's arm for more effectively covering the whole area.

To accommodate individual differences, participants in this study were allowed to move the robot's arm and teach it using any approach they wanted, with no time limit given. They ended teaching when they felt they were not able to guide the robot through the task any further because the robot always asked the participants to teach it how to perform the task 'as much as it could.' Participants received no direct instructions about the trajectories and how to provide the demonstrations. More details about the experimental design and procedure are presented in Section 3.2.2.

The purpose of our experiment was to understand how people's teaching strategies and their behaviours and attitudes change when they repeatedly get involved in the kinesthetic teaching of a humanoid robot. Therefore, the robot in our experiment was only observing the demonstrations and 'giving the impression of learning,' without having any learning capabilities (similar to [103, 107, 154]). However, we designed certain aspects of the experiment to make the human-robot teaching experience more realistic and to give participants the impression that the robot had learned something in the sessions, in order to keep participants motivated:

- After completing all four tasks, and before the end of each session, the robot cleaned

the same particles as in Task 1 located on a different table (see *Session conclusion* in Figure 3.1), just to show that it could perform a similar task by itself after being taught the four tasks by the participant. The performance of the robot depended on the participants' success in teaching Tasks 1 and 2. This was done by executing one of three pre-recorded motion patterns (since the robot was not actually learning from participants' demonstrations), decided by the experimenter during the session based on the outcome of the participant's table cleaning demonstrations in Tasks 1 and 2. More detail is presented in Section 3.2.2, under *Session conclusion*. Note that the robot performance was not intended to give examples of potentially effective ways of executing the task, and it was not meant to reflect precisely on how the participant had taught the robot in tasks 1 and 2. Instead, coarse-level feedback was given in the sense that the robot's performance was at most as successful in accomplishing the task as each participant's individual demonstrations in Tasks 1 and 2. Participants only watched the feedback and were not given any additional insight into their teaching. Note, employing real-time learning of robot movements from the two demonstrations in Tasks 1 and 2 goes beyond the scope of this study, which takes a human-centred approach.

- The material on the cutting boards and patterns drawn on the whiteboard were different in each of the five sessions for each participant, to make participants feel that each session was dedicated to teaching something new to the robot. The materials and patterns for each session of each participant were counterbalanced with a Latin square design. We used pieces of paper, paper clips, safety pins, binder clips, and buttons distributed on the cutting boards, as well as a set of horizontal or vertical lines, cross-hatching patterns, small circles, and random handwritten letters drawn on the whiteboard.

Additional photos from the experimental material and setup are provided in Appendix B.

3.2.1 Robotic Implementation

A Pepper humanoid robot (by Aldebaran Robotics) [5] was used in this experiment. This robot was selected as it was able to move around the experimental room on its own using its wheels and was tall enough to stand next to participants to receive kinesthetic demonstrations of different tasks. Pepper was sufficiently capable and safe to be used in an unsupervised setting, i.e., without the presence of an experimenter in the room, with

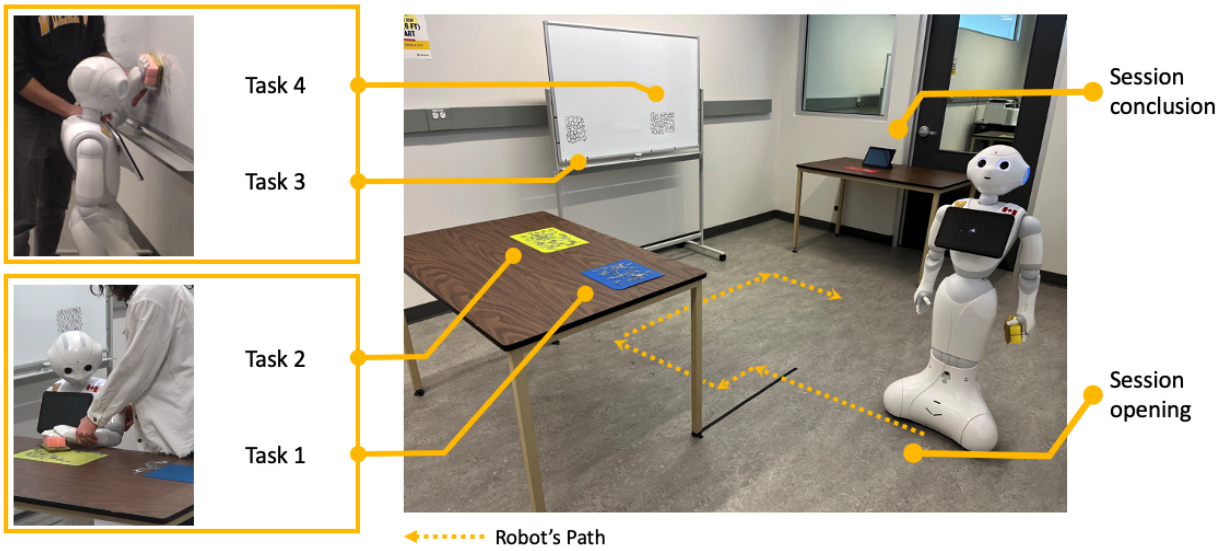


Figure 3.1: Layout and sequence of the experimental setup (right) and two examples of participants teaching cleaning tasks to the robot (left). At the location indicated as *Session opening*, Pepper greeted the participants and gave them information about the teaching session. The robot followed the path marked on the picture with dashed lines to interact with participants and receive demonstrations of four tasks at four different locations (*Tasks 1-4*). In *Session conclusion*, the robot showed the participant how it could clear the table on its own, and then asked them to pick up the tablet located on the table to complete the ending questionnaire. One-way mirrors to the adjacent room, where the experimenter was observing the scenario, are visible in the top right corner of the image.

non-expert participants. The robot has been used in multiple settings and is safe for human participants. Moreover, as a humanoid and social robot with 20 DOFs, Pepper seemed suitable for human-robot teaching tasks simulating home or office environments. The robot's behaviours in the sessions were semi-autonomous: Pepper greeted the participant, provided the necessary instructions, and asked participants to carry out teaching in different locations, as well as to complete a questionnaire, by using scripted actions. The experimenter remained in a room adjacent to the experimental room to oversee the session through one-way mirrors, and remotely issued additional commands for the robot when needed. These commands were intended for relocating the robot after its own autonomous navigation or selecting a behaviour among different options. The experimenter was not visible to the participants, who carried out the session on their own in the experimental room.

During the experimental sessions, the robot needed to position itself at specific locations for learning in front of a table and a whiteboard, and also perform a task in front of another table. To ensure consistency of recorded data and the played trajectories, only a few millimetres of error in the position of the robot were acceptable after navigating into these positions. Although all the manoeuvres of the robot were pre-planned with respect to the fixed setup in the experimental room, the robot's navigation abilities failed to meet that expectation. To address the issue, the experimenter needed to remotely 'nudge' the robot after every autonomous navigation, i.e., correct the positioning error of the robot. This was necessary in case of a large navigation error or when a participant shifted the position of the robot's base during teaching. However, visually verifying the precise location of the robot by the experimenter looking through one-way mirrors was challenging because the bottom part of the Pepper robot had a rounded shape and a high ground clearance. To overcome that obstacle, we attached three pieces of 3D printed white cylindrical shapes to the bottom of the robot, which stayed only 1-2 mm above the ground, and also marked circular signs on the floor corresponding to the position of those attachments when the robot was precisely at a given location. To adjust the position of the robot, the experimenter issued slow movement commands with the goal of matching the marks on the floor with two of the cylinders that were visible for each task (as pictured in Figure 3.2).

To control the robot, two pieces of software were developed. We used Choregraphe, an application for creating and testing behaviours for NAO and Pepper robots [2], to program behaviours needed for running the sessions. These behaviours had slight variations in the robot's dialogues to account for different materials to be cleaned and also to skip some conversations in later sessions. Note that the descriptions of the tasks and the instructions stated by the robot became shorter after the third session to avoid fatigue, as the participants were already familiar with the procedure. To let the experimenter send additional commands to the robot when one of the Choregraphe behaviours was running, another program was developed using Python. A block of Python code to receive external commands was put in multiple places within the Choregraphe behaviours, e.g., before and after every movement of the robot, to allow the experimenter to correct navigation errors. These blocks, when triggered, could pause the behaviour and read a character from a text file stored in the robot's memory to move the robot accordingly in four directions, rotate it clockwise or counter-clockwise, or choose between any set of behavioural options. Another Python code running on the controlling computer was utilized to modify that text file, based on keyboard input, through an SSH connection with the robot. The resulting outcome was that after starting a session, the robot was autonomously operating and interacting with the participant until one of the above-noted blocks was triggered, i.e., when there was a need to choose an action based on the experimenter's confirmation

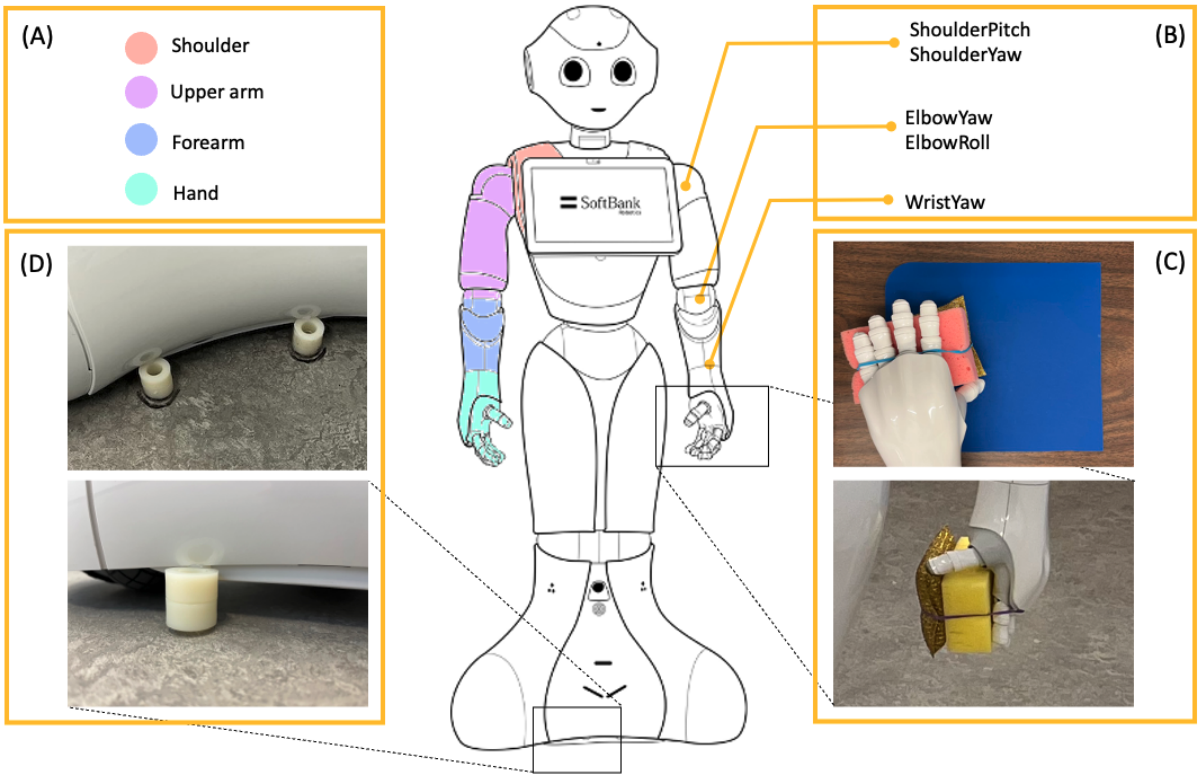


Figure 3.2: The Pepper robot used in this experiment. (A) Areas on the robot's arm parts for encoding participants' hand placement during teaching, (B) five main joints (DOFs) of the robot's arm used for kinesthetic teaching are marked, (C) the sponge attached to the left hand of the robot for performing the cleaning tasks is shown, (D) the attachments used for precisely locating the robot in the experimental room are shown, showing the situation where they are aligned with the marks on the ground meaning the robot is at the correct location.

or the robot needed to be moved. As soon as the necessary commands were sent and executed, the robot continued autonomously. It is worthwhile mentioning that we tried to allow approximately 15 minutes between the sessions of different participants to let the robot cool down, to minimize Pepper's faults and breakdowns.

3.2.2 Experimental Procedure

Ahead of their first visit, participants completed an online module running on our lab server to give consent, confirm their participation eligibility (based on criteria listed in Section 3.2.4), and answer a few demographic questions. Participants also received instructions on how to safely access our lab and follow COVID-19 regulations.

Upon arrival for a session, in a room different from the study room where Pepper was located, the participants put on eye-tracking glasses with the experimenter's assistance. To ensure an ideal fit of the glasses, the most suitable nose tip for every participant was selected before the first session. Once the eye tracker had been calibrated, participants entered the experimental room together with the experimenter. After the participant was located in a suitable location in front of the robot, the experimenter left the room and observed the experiment from an adjacent room through one-way mirrors. As visualized in Figure 3.1, the procedure used for the experiment consisted of the following steps:

Session opening: Each experimental session was initiated by the robot greeting the participant and explaining their role as a robot teacher. The robot also informed participants of two potential issues, harmful to the robot, that could come up in such kinesthetic demonstrations, by asking them to avoid excessive arm extension and to avoid hitting any of the robot's body parts with its arm. More specifically, the robot said: *"I need you to teach me how to perform some cleaning tasks by grabbing my arm and showing movements that I need to do. ... Before we move on, I want to remind you a few important things. Although I have multiple degrees of freedom in my arm, I have certain limitations. I may not be able to reach for objects that are too far or too close. Therefore, to provide useful instructions for me, please avoid stretching my arm too much. Also, be careful to perform actions in such a way that my arm doesn't collide with any part of my body."* Only in the first session, the robot asked the participant to freely move its left elbow and then, its whole left arm, after turning the corresponding motors off, to let them experience the feeling of physical interaction with a robot before their first kinesthetic teaching task to minimize the novelty effect.

Tasks 1 and 2 [Cleaning the table]: To receive the demonstration of the first task group (i.e., cleaning the table, Tasks 1 and 2), the robot followed the path indicated in Figure 3.1 to position itself in front of a table with its left arm on it. Note that the paths for the robot navigation were defined in a way that the robot usually moved away from the participant, so the participant could follow the robot safely around the experimental room. After arriving at the desired position, the robot explained more about the current teaching task: "*I want to clear the [paper clips] from this blue area of this table, as much as I can. After the beep, I will let you grab and move my left arm to show me how should I*

do this. Please try your best not to shift my position in the entire experiment." The robot then played a beep sound and turned off the motors in its left arm to allow kinesthetic teaching. While participants were teaching the robot, the robot was actively gazing at its left hand to appear attentive to the task, as suggested in [15]. Once the participant said "done," as they were instructed, and gave up the robot arm (the experimenter also had to remotely approve this), the robot turned on its motors again, thanked them, slightly lifted its left arm, and moved to the second task. For Task 2, the same procedure was followed for receiving another kinesthetic instruction of cleaning an area of a table.

Tasks 3 and 4 [Clearing the whiteboard]: The robot moved itself to the front of the whiteboard for receiving kinesthetic instructions for the next group of tasks (i.e., cleaning the board; Tasks 3 and 4). It rested its left hand on the whiteboard tray and then asked the participant to come closer and teach it how to clear the area in front of it, similar to Tasks 1 and 2. All other details of the procedure were the same as above.

Session conclusion: After the participant provided all the demonstrations, the robot approached another table, placed its left arm on it, and said: "*I will show you how I can now clear the [paper clips] from the table!*" At that time, the experimenter selected a pre-defined cleaning behaviour of {no cleaning at all, partially cleaning, or full cleaning}, that came closest to the participant's maximum success in teaching Tasks 1 and 2 during that session, in order for the robot to perform cleaning on its own, simulating robot learning. The behaviour selection was based on the experimenter's observation of each participant's behaviour in teaching Tasks 1 and 2, specifically their level of success of demonstrations. For those whose performance fell between two categories, the one with lower success in cleaning was selected. For *no cleaning at all* behaviour, the robot repeatedly and slightly moved its left hand, by rotating only a single elbow joint, to rub the sponge against the table. For *partially cleaning* the table, the robot utilized the same behaviour, yet exhibiting a wider range of motion, to effectively move some particles to the side of the cutting board. *Full cleaning* behaviour was programmed with movements in three DOFs of the robot, to cover the full surface of the cutting board, including areas closer to the right side of its tablet. After showing its own performance, the robot asked the participants to answer the ending questionnaire, asking about the workload of each teaching session (NASA-TLX [85], discussed in Section 3.2.3), on a tablet that was already located on the table. Following the completion of the questionnaire, the robot thanked and said goodbye to the participant. The experimenter entered the room at that point and helped the participant remove the eye-tracking glasses. In the case of the fifth, final session, the experimenter debriefed the participant about the real purpose of the study. Otherwise, the experimenter scheduled their next visit. Participants did not have two sessions on the same day[3]. Note, in order

[3]On average, there were 4.11 (SD=2.75) days between successive sessions for a participant.

to prevent participants from becoming too self-conscious about their behaviour, it was not stated until the debriefing at the end of the study that the purpose of the experiment was to understand the 'behaviours of human teachers.' The study had the rather general title of "Human-Robot Teaching Study."

3.2.3 Data Collection and Analysis

We recorded and analyzed the (a) time-stamped positions of the joints in the robot's left arm, (b) participants' gaze data, (c) videos of participants teaching the robot, as well as (d) initial and ending questionnaires.

Time-stamped Positions of the Joints in the Robot's Left Arm

To capture the physical properties of the demonstrated cleaning behaviours, position encoder values of all five DOFs in the left arm of the robot were saved while participants were teaching the robot, at a frequency of 10 Hz. The recording of this data for each task began after the beep and ended when the participant verbally indicated that their teaching was finished. Using programmed Python blocks in Choregraphe, that data was saved as text files in the robot's computer so they could be downloaded later. The data was then computationally processed to calculate the *time spent on providing the demonstrations* and the *robot's joints range of motions in demonstrations*, and to estimate the amount of task space covered in the participants' demonstrations. This last component is an estimate of how successful the cleaning demonstrations were and is therefore called *success of demonstrations* (estimated computationally).

To estimate the extent to which the provided demonstrations covered the task spaces, the following automatic post-processing procedure was performed:

- The joint positions were first converted to the trajectory of the end-effector, i.e., the left hand of the robot, by computing forward kinematics relative to a fixed frame attached to the robot's base, using the NAOqi Cartesian control API [1]. In this frame, the x-axis points to the robot's front, the y-axis points to the robot's left, and the z-axis points upwards.
- Trajectories of the robot's left hand were examined along axes orthogonal to the surfaces being cleaned, i.e., the z-axis for clearing the table and the x-axis for cleaning the whiteboard, allowing the distance between the robot's hand and the surfaces

being estimated, to identify sponge-surface touching points. To obtain a smooth and continuous estimation of the distance between the sponge and the surface over time, a spline of degree 5 with smoothing factor s was fit on the orthogonal component versus time.

- Touching points were found based on two parameters: (a) the second derivative of the spline was larger than a threshold $h \geq 0$ at those points (which could reflect contact and separation), and (b) points were closer to the surface than the starting point plus a small tolerance $t > 0$ (to account for sponge compression during contact). An illustration of this procedure is shown in Figure 3.3 (B) for an example demonstration. In this case, the sponge appears to be periodically in contact with the table to push the particles.

- Next, the algorithm approximated the sponge's orientation at each touching point based on the first derivative of the trajectory projected on the cleaning surfaces, determined by its two adjacent neighbour points.

- Lastly, points on rectangular grids (corresponding to cutting boards or writing areas of the whiteboard presented in the inverse kinematic frame) that were covered by placing the sponge at the touching point areas were marked as clean.

Figure 3.3 (C) shows the trajectory and the estimated cleaned area of the previous example. We set the parameters $s = 0.003$, $h = 0.01$, and $t = 5mm$ for Tasks 1 and 2, and $s = 0.002$, $h = 0$, and $t = 5mm$ for Tasks 3 and 4, in order to obtain accurate estimations by testing multiple combinations. This computational approach does not require any vision system, which could, in future studies, allow the robot to automatically track human teachers' success in real-world teaching situations (as discussed in Section 3.4.5).

Participants' Gaze Data

Participants' eye gaze was recorded using Tobii Pro Glasses 2 [9] at a frequency of 120 Hz during all the sessions. Gaze data were represented by points mapped on the glasses 'scene camera video,' i.e., the video of what was seen by the participants, taken using the camera embedded in the glasses. To process the data, the moments when each teaching task began and ended were first manually annotated on the scene camera videos. Then, the pixel coordinates of participants' eye gaze at each recording time step and the type of eye movement (Fixation or Saccade) were exported from the Tobii Pro Lab software for our analysis. We created a Python code to help measure the amount of participants'

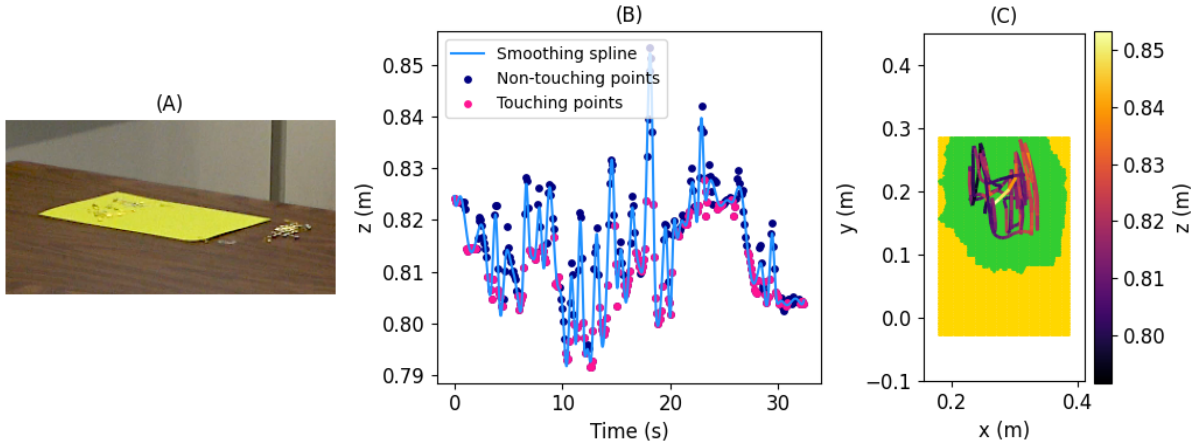


Figure 3.3: Illustration of the computational method used to estimate the success of demonstrations, showing an example from Task 2. (A) Picture of the partially cleaned task space, i.e., the yellow cutting board on the table with binder clips partially cleared, (B) The vertical z-component of the robot's end-effector positions versus demonstration time, as well as the fit spline for determining touching points, (C) The demonstrated trajectory from the top (projected on the x-y plane) and the yellow board on the table. The computational algorithm estimated that 53.4% of the cutting board, highlighted in green, has been cleaned. The subjective estimate (i.e., rated by an experimenter and verified by a second rater) of the cleaned task space in this example, based on the picture of the task space, is 60%.

gaze fixations on important regions related to the teaching tasks within their field of view, namely the robot's hand, for all the tasks and cutting boards for Tasks 1 and 2. Ultimately, each fixation gaze data point was determined to be inside or outside of the defined regions.

By applying basic image processing techniques on the scene camera videos of Tasks 1 and 2, our algorithm automatically found dynamic quadrilateral regions corresponding to the cutting boards. To do so, the program carried out colour thresholding based on a given range of pixel values in the HSV colour space, using OpenCV [30], and found a mask corresponding to the blue or the green cutting boards. An opening morphological transformation with an 8×8 elliptical/circular kernel was then applied to the mask to remove noise. In the end, the cutting board was represented by a quadrilateral contour with four edges.

In addition, for all the tasks, dynamic regions corresponding to the sponge attached to the robot's hand were identified. The process was semi-automated, with a human (myself)

involved to correct any tracking errors. The program displayed the first frame of each teaching task from the scene camera videos, randomly and without any label. A circle had to be drawn by the researcher around the sponge, on top of the robot's hand, in that frame of the video in order for the CSRT Tracking API [8] of OpenCV (which uses the discriminative correlation filter from [133]) to start tracking the sponge within an initial bounding box set based on the drawn circle. Afterwards, the video was played. The researcher could pause the video and draw another circle around the sponge to define a new reference bounding box in case poor tracking accuracy was detected. Figure 3.4 shows two snapshots of the processed scene camera videos of participants teaching Task 1, with the regions corresponding to the robot's hand and the cutting board, both marked on the pictures.

Considering the cutting boards in the analyses of participants' gaze was meant to help examine whether their visual attention to task-related areas, apart from the robot's hand, showed any change after they became more experienced. In fact, the cutting boards were fixed to the table and therefore, no teaching action needed to be done on them directly. In a preliminary evaluation of our data, we observed no patterns in terms of the amount of fixations on the cutting boards or cutting boards combined with the robot's hand, which were only available for Tasks 1 and 2 throughout the sessions. Thus, the analysis of human teachers' gaze data presented in this work only includes the *amount of gaze fixations on the robot's hand*, to test our hypothesis outlined in Section 3.1.

Videos of Participants Teaching the Robot

In order to investigate the methods and strategies used by participants to manipulate the robot's arm for kinesthetic teaching, all the sessions were video-recorded using two cameras. It was necessary to have two cameras as our experimental design required participants to face different directions to teach multiple tasks. To minimize distractions, the camera in the experimental room was on a small tripod inside a closet, and the other one was behind one of the one-way mirrors in the adjacent room.

Using BORIS event-logging software [72], I reviewed all the videos to encode the participants' hand positions during teaching as well as different participant behaviours, related to the tasks but not contributing to the movements, observed during demonstrations (e.g., anthropomorphic behaviours and comparisons between themselves and the robot). The position of each hand of participants was labelled as lying {nowhere, on the table, on the robot's hand, on the robot's forearm, robot's upper arm, or robot's shoulder}. The exact location of different robot arm parts, i.e., hand, forearm, upper arm, and shoulder is shown in Figure 3.2. These markers were then used to calculate the amount of *the usage of both*



Figure 3.4: Two snapshots of the processed scene camera video obtained from the Tobii Glasses 2 gaze recording device. The images show two participants' own views while teaching the robot Task 1. Black circles drawn on the images locate the robot's hand, purple quadratic regions represent the cutting board, and green dots indicate participants' gaze fixation targets. Each fixation gaze data point was determined to be inside or outside of the defined regions to estimate the number of gaze fixations at each region, including the robot's hand.

hands, *guiding the robot's upper arm* with at least one of their hands, and *not guiding the robot's hand* while teaching each task, as percentages showing the duration of those events compared over the entire teaching duration. The number of *changes in teaching strategy* during teaching each task, counting any change in the placement of each hand of a participant during teaching, was also derived from the annotated videos. For example, if a participant was guiding the robot's upper arm with their right hand and then used that hand to move the robot's forearm, one change in their teaching strategy was counted.

The four features noted above were selected because we noticed some trends in our preliminary evaluation of the data and identified those as worth checking in detail: teaching with both hands and holding the robot's upper arm could imply adapting a teaching strategy by non-experts for safer/easier robot arm manipulation (based on what experts may do while performing same tasks). Furthermore, instances when teachers were not guiding the robot's hand while demonstrating tasks could mean adjusting the joint of the robot's arm, without contributing to the task directly, and finally, changes in teaching strategy could help understand to what extent non-experts try new methods for robot teaching.

With the videos of participants teaching the robot, I also evaluated the *success of demonstrations* (estimated subjectively), i.e., the percentage of cleaned areas in the demonstrations based on screenshots showing the task spaces after the sessions. Using a Python script, I was presented with cropped images containing only the teaching workspaces, in random order and with no label, to minimize biases. The percentage of subjective success was entered with a precision of 5%. These estimates were equivalent to the success of demonstrations estimated computationally with the method described in Section 3.2.3.

Initial and Ending Questionnaire

The initial questionnaire, completed prior to participants coming to the first session gathered data on participants' demographics. This demographic questionnaire asked participants about their *(a) age*, *(b) gender*, *(c) occupation*, and *(d) handedness*, as well as *(e) whether they had ever programmed any robot, interacted with a human-like robot, or a Pepper robot in particular*, and *(f) if they have ever participated in a study related to human-robot teaching*, to control for potential confounds. At the end of each session, we asked participants to answer a NASA-TLX workload questionnaire. The NASA-TLX is a widely-used assessment tool that allows participants to perform subjective workload assessments, on six scales of *mental demand*, *physical demand*, *temporal demand*, *performance*, *effort*, and *frustration* [85]. Our approach in this study, which is a common variation to NASA-TLX [84], was to capture and analyze sub-scale ratings of workload rather than a

single overall workload score. This way, we studied workload from a number of different aspects, e.g., mental and physical. This questionnaire was answered on a tablet that could locally store the data. Answers to each scale of the questionnaire were recorded on a continuous scale ranging from 0 to 1000, with 1000 being the highest workload score on the NASA-TLX sub-scales.

3.2.4 Participants

In total, 28 volunteers completed our study. As outlined in the study recruitment flyer, the eligibility criteria were: *(a) not being a roboticist or an HRI researcher*, *(b) being able to see close distances without glasses* (so that eye-tracking glasses could be worn without lenses attached), and *(c) meeting requirements regarding participants' potential safety risks*[4]. The participants were all University of Waterloo students, between the age of 18 and 31 (M=21.54, SD=2.55). 12 self-identified as female, 15 as male, and one who did not disclose their gender. All participants except one were right-handed. Seven participants had previously programmed a robot, mainly during high school activities, but none had ever interacted with a human-like robot. The study received full ethics clearance from the University of Waterloo Human Research Ethics Board (included in Appendix A).

3.2.5 Data Reliability

All subjective data obtained from videos of participants teaching the robot, i.e., the success of demonstrations and their hand positioning during teaching, was verified by another researcher coding a small portion of the data. For the subjective estimates of the success of demonstrations derived from images of the tasks, I coded all and another researcher coded 25% of the data (140 picture evaluations). We then performed an inter-rater reliability analysis. Ratings of the random subset of data that had been coded twice produced an intra-class correlation coefficient of 0.95 (95% CI: 0.88 to 0.97, $F(139, 18.6) = 45.4, p < .0014$), confirming good to excellent reliability of the subjective approach for reporting the success of demonstrations [114].

Positions of participants' hands during teaching were similarly coded, and 10% of the videos of the sessions (54 tasks) were also coded by another researcher to verify the calculated measures. Intra-class correlation coefficients of 1.00 (95% CI: 0.99 to 1.00,

[4]People who used medical devices that were susceptible to disturbance by infrared light and/or radiation, people with photosensitive epilepsy, and people who had implanted medical devices were excluded in order to eliminate participants' safety risks due to proximity to the robot and the measuring devices.

$F(53, 53.7) = 410, p < .0014$), 0.99 (95% CI: 0.98 to 1.00, $F(53, 53) = 152, p < .0014$), 0.73 (95% CI: 0.56 to 0.83, $F(53, 45.9) = 6.69, p < .0014$), and 0.75 (95% CI: 0.59 to 0.84, $F(53, 53.1) = 6.64, p < .0014$), for the usage of both hands, guiding robot's upper arm, not guiding robot's hand, and changes in teaching strategy, respectively. These numbers confirm the excellent reliability of the first two measures and the moderate to good reliability of the last two [114].

To remove noise from the data collected for all 560 tasks over the 140 sessions, the following procedure was carefully followed:

(a) Time spent on providing the demonstrations was adjusted for a total of 20 tasks due to special circumstances occurring during teaching that increased the recorded time. These events included re-alignment of the sponge in the robot's hand (13 instances), eye-tracking glasses re-calibration (2 instances), and participant asking a question during teaching (1 instance).

(b) To ensure consistency, the success of demonstrations data for 12 tasks (estimated both computationally and subjectively) were removed, because the robot was either not at the right position when receiving demonstrations or was moved considerably by the participant during teaching, as detected in the videos of participants teaching the robot.

(c) Gaze data were not included in case fewer than 70% of all the possible data points when teaching each task had been recorded, to ensure that the calculated fractions of fixations represent almost the entire duration of each task. This factor led to the omission of gaze data for 66 tasks. In most of the cases that led to data loss in gaze recording, Tobii glasses could not detect the eyes properly when a participant was looking down as a result of standing close to the robot and teaching a task on a short table.

Due to technical errors, there was one experimental session during which position data of the joints in the robot's left arm was not recorded and another session during which video recording was not saved. In two of the experimental sessions, NASA-TLX questionnaires were not recorded. Errors in recording or processing of participants' gaze data also led to missing gaze data for 13 tasks.

3.2.6 Data Analysis and Statistical Modelling

It should be first noted that to address RQ2, we looked deeper into participants' individual characteristics and potentially patterns that can indicate more suitable teachers for robots based on their achievements in providing physical demonstrations. In particular, we examined the average success of Tasks 1 and 2, i.e., cleaning on the table, in each session. The success of Tasks 3 and 4, i.e., clearing areas of the whiteboard, was not considered as almost all participants provided highly successful demonstrations (see Figure 3.5). Therefore, it could make it difficult to distinguish between participants' performances. Taking the five success averages as the features, we applied a k-means clustering algorithm to categorize the participants into different clusters [150, 117]. $k = 3$ produced the most reasonable clustering. Accordingly, we introduced a factor named *cluster*, assigned to every participant (more details are presented in Section 3.3.1).

To answer different elements of RQ1, we used Linear Mixed-effects Model (LMM) [24] and analyzed the changes in obtained measures as participants taught the robot over multiple sessions with *participant* and *task nature* (the robot being taught at the table or at the whiteboard) as random effects: a random intercept for each participant, and a random slope for task nature. The considered factors for the analysis of different categories of data to study each research question are as follows:

- For **success of the demonstrations (RQ1.1)**:
  - *Session number* and *task order*[5] as numeric variables ranging from 1 to 5 and 1 to 4, respectively, to account for between-sessions and within-sessions progress of participants in teaching the robot.
  - *Age* and *gender* of participants as possible confounding factors.
  - *Amount of gaze fixations on the robot's hand, usage of both hands, guiding the robot's upper arm, not guiding the robot's hand,* and *changes in teaching strategy* as confounding behavioural variables that could have influenced teaching time and success.
- For **time spent on providing the demonstrations (RQ1.2)**:
  - *Session number* and *task order*, plus *amount of gaze fixations on the robot's hand, usage of both hands, guiding the robot's upper arm, not guiding the robot's hand,* and *changes in teaching strategy* for the same reasons as above.

[5] *Task nature* and *task order* together uniquely define the task. In each level of task nature, i.e., cleaning the table or the board, the one with the larger task order is cleaning the bigger surface.

  - Participants' assigned *cluster*, as well as *age* and *gender* of participants as possible confounding factors.

- For **amount of gaze fixations on the robot's hand (RQ1.3)**:
  - *Session number* and *task order*, plus *age* and *gender* of participants, for the same reasons as above.
  - Participants' assigned *cluster*, to identify potential differences in their behaviours and see if they are correlated with high/low achievements.
  - *Usage of both hands, guiding robot's upper arm, not guiding robot's hand,* and *changes in teaching strategy* for taking into account the possibility of the methods used for manipulating the robot's arm influence the gaze patterns.

- For the behavioural data on manipulation and teaching strategy, i.e., **amount of usage of both hands, guiding robot's upper arm, not guiding robot's hand, and changes in teaching strategy (RQ1.4)**:
  - *Session number* and *task order*, plus *age* and *gender* of participants, for the same reasons mentioned earlier. Measures concerning RQ1.4 did not follow a normal distribution for our participant pool. The LMM analysis is therefore limited to improving participants only.

- For **NASA-TLX workload items (RQ1.5)**:
  - *Session number*, and *age* and *gender* of participants, for the same reasons mentioned earlier.
  - *Average time spent on providing the demonstrations*, *average success of the demonstrations*, *average gaze fixations on robot's hand*, *average usage of both hands*, and *average amount of guiding robot's upper arm* as possible confounding factors, since NASA-TLX workload questionnaires had been collected after each session (not for each task).

The factors listed above were initially included in the models. We then kept only a subset of those to minimize the Akaike's Information Criterion (AIC) [29]. After checking the interaction effects, we gained the final models presented in the Results section. We also visually checked if the assumptions for the usage of LMMs were met, namely the residuals having constant variance and are independent and normally distributed, using histogram and Q-Q (quantile-quantile) plot for the residuals and scatter plot of residuals

against the fitted values. If not, we applied a transformation to our data (the only case was the analysis with regards to RQ1.1, Section 3.3.1). Note, the baseline category for participants' assigned cluster is 'high-achievers' and for gender is 'female,' in all presented models.

3.3 Results

In this section, we explain how *success of the demonstrations*, *time spent on providing the demonstrations*, *amount of gaze fixations on robot's hand*, *behavioural data on manipulation and teaching strategy*, and *NASA-TLX workload items* were impacted by participants' repeated teaching interactions with the robot, both between the sessions and within the sessions, to answer RQ1.1-5. We also report how the confounding factors taken into account affected these measures and explore the characteristics of participants who achieved different success rates in teaching, to answer RQ2.

3.3.1 [RQ1.1] Success of Demonstrations

We initially observed that participants were able to move some joints of the robot within a wider range overall. For example, in teaching Task 1, the ElbowYaw joint in the left arm of Pepper was moved on average within 40.57 degrees (SD=30.19 degrees) in Session 1, 57.58 degrees (SD=37.81 degrees) in Session 2, and 65.66 degrees (SD=38.61 degrees) in Session 3. These trends suggested that the overall area covered by the demonstrations may have increased over time.

The percentage of covered areas (i.e., cleaned areas) in the demonstrations, indicating the success of demonstrations, was estimated using two approaches: subjective (described in section 3.2.3) and computational (described in section 3.2.3). Figure 3.5 shows the average success of demonstrations over the sessions, estimated using both methods. As seen in the figure, computational evaluations of success were noticeably lower than subjective evaluations in Tasks 1 and 2. However, computational estimates had high positive correlations with the subjective estimates in all the tasks according to Pearson correlation analyses: $r(136) = .744, p < .001$ in Task 1, $r(137) = .837, p < .001$ in Task 2, $r(136) = .861, p < .001$ in Task 3, and $r(136) = .907, p < .001$ in Task 4. The trends observed in Figure 3.5 also show that both subjective and computational approaches undergo changes in a correlated manner. All these findings support the validity of the computational method for revealing variations in the success of demonstrations.

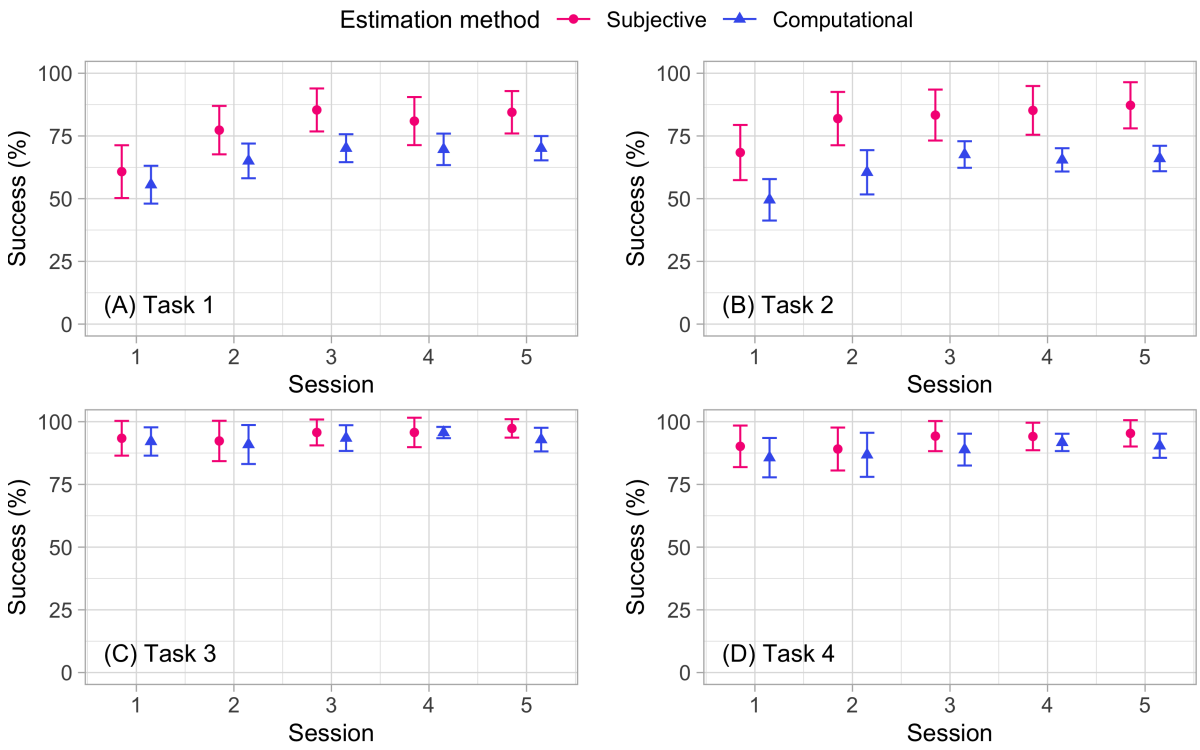


Figure 3.5: Success of demonstration of each task estimated using two methods of subjective and computational over five sessions. Error bars indicate 95% confidence intervals.

For the rest of the statistical analyses presented in this article, only the 'subjective' estimates of demonstrations' success are taken into account, since they are more accurate estimates, in the sense that they are derived from a human observer's ratings that could be considered as a ground truth. As shown in Figure 3.5 and in line with the increase in range of motion in individual robot's joints, the provided demonstrations in later sessions covered larger areas overall. The average success rate of participants, across all four tasks, was 78.5% (SD=27.21%) in Session 1, 85.22% (SD=23.98%) in Session 2, 89.72% (SD=20.02%) in Session 3, 89.12% (SD=20.51%) in Session 4, and 91.18% (SD=18.30%) in Session 5. In Tasks 1 and 2, participants, on average, improved their teaching success over the few first sessions. The success of providing the demonstrations of Tasks 3 and 4 was generally higher than that for Tasks 1 and 2.

As described earlier in Section 3.2.6, we looked at the teaching success of participants in Tasks 1 and 2 and categorized non-expert teachers into three clusters. As shown in Figure 3.6, participants in each of the three clusters shared a unique characteristic in

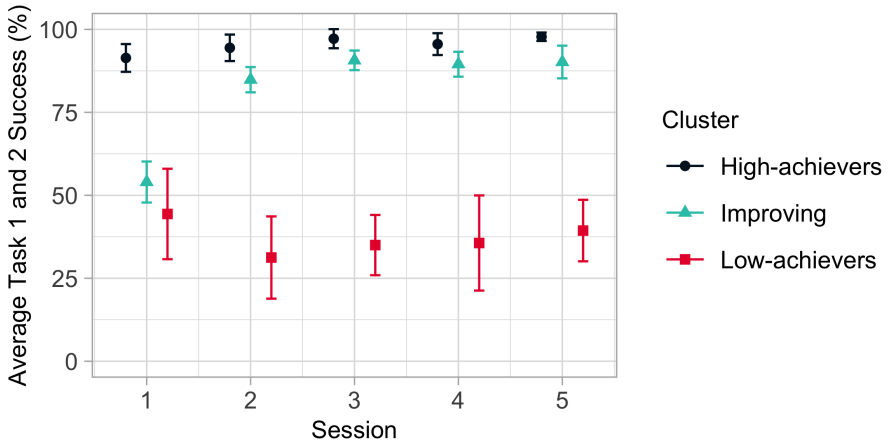


Figure 3.6: Average success of demonstration of Tasks 1 and 2 for each cluster of participants over five sessions. Error bars indicate 95% confidence intervals.

terms of average success in providing demonstrations for Tasks 1 and 2 over the sessions. We therefore called them *high-achievers*, *improving*, and *low-achievers*, as explained below:

- The 'high-achievers' were 9 participants who showed high teaching success from the first session. In this cluster, the average success of Tasks 1 and 2 for each participant was at least 80% in every session. 44 out of 90 total provided demonstrations of Tasks 1 and 2 were fully successful. According to Figure 3.6, high-achievers also improved slightly over the sessions. The small confidence interval observed for the average demonstration success in Session 5 implies that high-achievers became consistently successful eventually.

- The 'improving' participants were the biggest cluster consisting of 15 participants whose average teaching success of Tasks 1 and 2 was not high at the beginning, but later improved. As seen in Figure 3.6, the improvement was more prominent over the first three sessions. Each participant in this cluster provided table cleaning demonstrations of at least 87.5% successful in a session after Session 1.

- The 'low-achievers' cluster consisted of 4 participants who less successfully demonstrated the table cleaning tasks without showing any noticeable improvement over the sessions. For these participants, the most effective demonstration of table cleaning tasks in a session was only 67.5% successful.

Table 3.1: Linear Mixed-effects Model predicting exponentially transformed subjective estimates of the success of demonstrations of participants

Covariate	**Estimate**	**SE**	**t**	**p**
Session:cluster *High-achievers*	7.511	1.667	4.506	***
Session:cluster *Improving*	8.260	1.309	6.310	***
Session:cluster *Low-achievers*	-2.657	2.189	-1.214	
Task order	8.494	2.217	3.831	***
Gender *Male*	19.414	8.349	2.325	*
Gaze at hand	0.398	0.124	3.216	**
Usage of both hands	0.179	0.072	2.494	*

Female gender is baselines. * = $p < .05$; ** = $p < .01$; *** = $p < .001$.

An LMM fit on the transformed data showing the success of demonstration is presented in Table 3.1. To fit the LMM, subjective estimates of the success of participants' demonstrations in each task are divided by 20 and transformed exponentially, so that the model met the assumptions. In this model, the effect of Session is studied within each cluster. Both high-achievers and improving participants demonstrated the tasks more successfully as they moved forward in the sessions ($se = 1.667, t = 4.506, p < .001$ and $se = 1.309, t = 6.310, p < .001$, respectively). Overall, as participants proceeded with tasks within the sessions, i.e., from tasks 1 to 4, the success of their demonstration was improved too ($se = 2.217, t = 3.831, p < .001$). Male participants, on average, provided significantly more successful demonstrations than female participants ($se = 8.349, t = 2.325, p < .05$). The percentage of gaze fixations at the robot's hand was another significant factor positively affecting the success of demonstrations ($se = 0.124, t = 3.216, p < .01$). The model also revealed that increased usage of both hands for teaching led to more successful demonstrations ($se = 0.072, t = 2.494, p < .05$).

3.3.2 [RQ1.2] Time Spent on Providing the Demonstrations

Figure 3.7 shows the average time spent on providing demonstrations of each task over the five sessions. The results indicate a steady decrease in demonstration time for Tasks 1 and 2 over the five sessions. There is also a decrease in Tasks 3 and 4 demonstration times, but most noticeably when comparing only Session 1 with other sessions. Table 3.2 shows the results of an LMM fit on the data. According to this model, as participants proceeded through the sessions, they spent less time teaching the robot ($se = 0.388, t =$

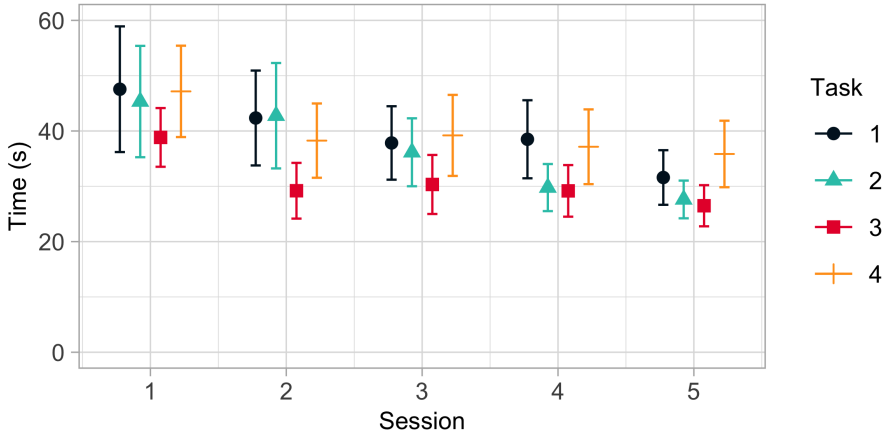


Figure 3.7: Average time spent on teaching each task to the robot over five sessions. Error bars indicate 95% confidence intervals.

$-8.405, p < .001$). Using both hands for teaching slightly increased the teaching time ($se = 0.027, t = 3.097, p < .01$). Moreover, teaching time was positively affected by the number of changes in teaching strategy ($se = 0.351, t = 7.371, p < .001$).

3.3.3 [RQ1.3] Amount of Gaze Fixations on Robot's Hand

The average percentage of gaze fixations on the robot's hand over all gaze fixations while teaching each task is illustrated in Figure 3.8 for the five sessions. As can be seen in the figure, participants spent less time looking at the robot's hand when teaching Task 2 compared with other tasks. An overall decreasing trend is also visible in Figure 3.8, mostly when looking at Tasks 3 and 4. As indicated by the LMM fit on this data (see Table 3.3), the fixations on the robot's hand decreased as the participants proceeded through the sessions ($se = 0.371, t = -3.919, p < .001$), which confirms H1, which predicted as participants become more experienced in robot teaching, they gaze less often at the robot's hand. There was also a decrease in the fixations on the robot's hand in successive tasks within the same session ($se = 0.912, t = -4.824, p < .001$). Participants' age was found to be significantly affected by the fixations on the robot's hand: younger participants looked more often at the robot's hand when teaching ($se = 0.463, t = -3.074, p < .01$, keeping in mind that participants were all university students and therefore had a narrow age range). The final model did not include any behavioural characteristics of participants for the robot's arm manipulation, meaning that none of the usage of both hands, guiding the

Table 3.2: Linear Mixed-effects Model predicting the time spent on providing the demonstrations

Covariate	**Estimate**	**SE**	**t**	**p**
cluster *Improving*	2.848	4.248	0.670	
cluster *Low-achievers*	-11.409	6.012	-1.898	
Session	-3.259	0.388	-8.405	***
Gender *Male*	-1.354	3.760	-0.360	
Usage of both hands	0.083	0.027	3.097	**
Changes in teaching strategy	2.590	0.351	7.371	***

High-achievers cluster and *Female* gender are baselines. ** = $p < .01$; *** = $p < .001$.

Table 3.3: Linear Mixed-effects Model predicting the amount of gaze fixations on the robot's hand during teaching

Covariate	**Estimate**	**SE**	**t**	**p**
Session	-1.453	0.371	-3.919	***
Age	-1.423	0.463	-3.074	**
Task order	-4.398	0.912	-4.824	***

** = $p < .01$; *** = $p < .001$.

robot's upper arm, guiding the robot's hand, and changes in teaching strategy were found to be affected by participants' gaze allocation patterns.

3.3.4 [RQ1.4] Behavioural Data on Manipulation and Teaching Strategy

Figure 3.9 illustrates the trends in the data on the robot's arm manipulation and teaching strategy derived from the analysis of video recordings of the sessions. This figure includes the average percentage of duration when participants used both hands, participants held the robot's upper arm, and the robot's hand was not guided during kinesthetic teaching of each task over five sessions, as well as the average number of changes in the teaching strategies (i.e., placement of participants' hands on robot's arm) for each participant cluster. As can be seen in Figure 3.9 (A) and (B), there was an increase in the average time teaching using both hands and guiding the robot's upper arm over the sessions, for improving participants. This increase was most pronounced in the early sessions. For

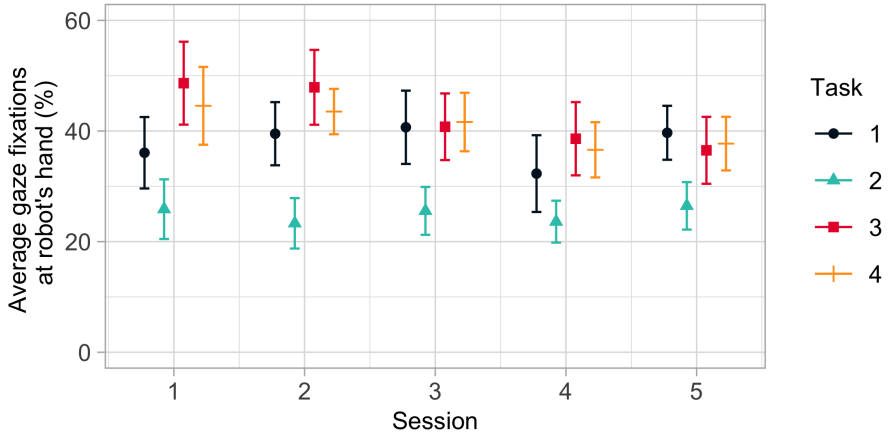


Figure 3.8: Average percentage of participants' gaze fixations on the robot hand during teaching each task (compared to all gaze fixations during teaching) over the five sessions. Error bars indicate 95% confidence intervals.

high-achievers, there is also an increasing trend of guiding the robot's upper arm when teaching over the sessions. Low-achievers used both hands for teaching around three times less than the other two clusters. Low-achievers barely guided the robot's upper arm when providing their kinesthetic demonstrations. In fact, no participant in that cluster ever touched the robot's upper arm in Session 3 when teaching any task to the robot. As seen in Figure 3.9 (C), compared to other clusters, low-achievers spent more time teaching without guiding the robot's hand. From Figure 3.9 (D), it can be seen that in later sessions the majority of participants were more consistent in the number of changes in their teaching strategies (smaller confidence intervals). In every session, high-achievers made the highest average number of changes in their way of guiding the robot's arm when teaching it. Improving participants made slightly fewer changes than high-achievers. Low-achievers made considerably fewer changes compared to the other two clusters, especially in the first four sessions.

The LMMs investigating the behavioural data on manipulation and teaching strategy of improving participants are presented in Table 3.4. Improving participants used both hands at the same time for teaching and guided the robot's upper arm at least with one hand more frequently as they gained more experience over the sessions ($se = 0.789, t = 6.030, p < .001$ and $se = 0.907, t = 7.465, p < .001$, respectively), and within the sessions ($se = 1.028, t = 4.068, p < .001$ and $se = 1.376, t = 3.562, p < .01$, respectively). The models also suggest that the percentages of times when the improving participants were not guiding the robot's

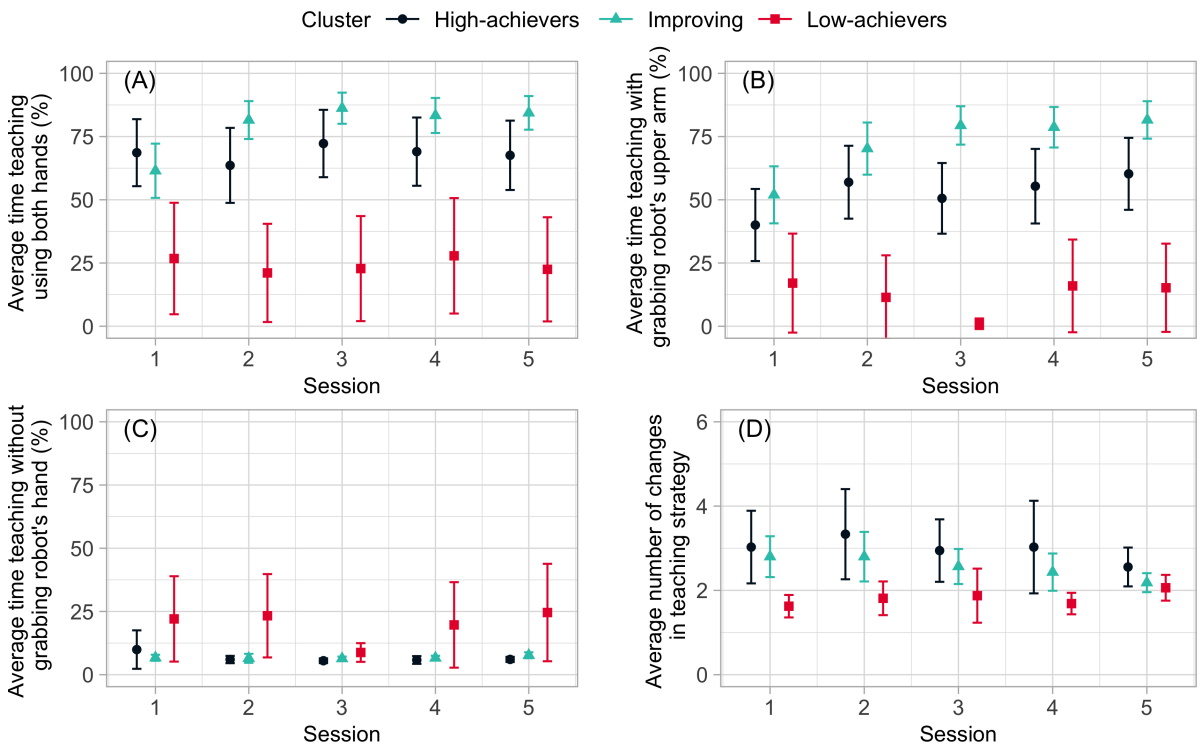


Figure 3.9: Average percentage of the time when (A) participants used both hands, (B) the robot's upper arm was guided, and (C) the robot's hand was not guided, as well as (D) the average number of changes in the teaching strategies, during kinesthetic teaching over five sessions, for each cluster of participants. Error bars indicate 95% confidence intervals.

hand decreased with their progress within sessions ($se = 0.282, t = -5.945, p < .001$). Improving participants changed their teaching strategy less often in later sessions in the experiment ($se = 0.056, t = -2.851, p < .01$) and when progressing within the sessions ($se = 0.103, t = -3.186, p < .01$). As noted before, we do not provide further statistical analyses concerning the other two clusters due to a lack of data (too few participants fell into these clusters) and their data did not follow a normal distribution.

3.3.5 [RQ1.5] NASA-TLX Workload Items

Figure 3.10 shows results for the NASA-TLX workload questionnaire items answered after teaching all the tasks in each session. While the average perceived mental demand and physical demand seem to remain unchanged over the different sessions (Figure 3.10 (A) and

Table 3.4: Linear Mixed-effects Model predicting improving participants' average percentage of the time when both hands were used, robot's upper arm was guided, robot's hand was not guided, during teaching each task, as well as the average number of changes in the teaching strategies

	Usage of both hands				**Guiding robot's upper arm**			
Covariate	**Estimate**	**SE**	**t**	**p**	**Estimate**	**SE**	**t**	**p**
Session	4.759	0.789	6.030	***	6.768	0.907	7.465	***
Gender	10.622	11.591	-0.916		10.920	18.448	1.487	
Task order	4.183	1.028	4.068	***	4.902	1.376	3.562	**

	Not guiding robot's hand				**Changes in teaching strategy**			
Covariate	**Estimate**	**SE**	**t**	**p**	**Estimate**	**SE**	**t**	**p**
Session	0.212	0.162	1.310		-1.600	0.056	-2.851	**
Gender	0.565	0.858	0.658		-0.044	0.312	-0.140	
Task order	-1.678	0.282	-5.945	***	-0.327	0.103	-3.186	**

** = $p < .01$; *** = $p < .001$.

(B)), the average perceived effort appears to decline over the sessions (Figure 3.10 (E)). There are also drops in the average perceived temporal demand and frustration in the final Session 5 (Figure 3.10 (C) and (F)). The average perceived performance seems to increase after Session 1 (Figure 3.10 (D)).

According to the LMMs fit on the NASA-TLX questionnaire data (see Table 3.5), session number only affected the perceived effort of participants, which significantly decreased with repeated experience in multiple sessions ($se = 7.382, t = -2.508, p < .05$). Spending more time on providing the demonstrations was also correlated with a higher perceived mental demand and effort ($se = 0.768, t = 2.361, p < .05$ and $se = 0.989, t = 3.335, p < .01$, respectively), but less positive evaluation of performance ($se = 1.137, t = -2.700, p < .01$). With more successful demonstrations, participants rated their own performance more positively ($se = 1.187, t = 3.851, p < .001$).

3.3.6 Other Observations

When reviewing videos of the teaching sessions, we noticed five participants who exhibited interesting behaviours toward the robot during or after teaching tasks (they will be called

Participants A, B, C, D, and E here). While these observations are anecdotal, we are documenting them here since they might inspire future work. Participant A communicated their instructions verbally during kinesthetic teaching of some tasks in the first three sessions. Participant B verbally repeated the robot's instructions, e.g., about not moving its position, while teaching it. Participant C put a hand on the robot's back after finishing teaching and Participant D gave thumbs-up and petted the robot, following the teaching. Participant E tested the movements on their own body before teaching the robot, to see how their arm joints may move to achieve a certain motion. All five participants referred to here belonged to the improving cluster, except for participant A, who was among the low-achievers.

The average age of the high-achiever participants was 20.78 (SD=2.10) and three of them had programmed a robot before. Participants in the improving cluster had an average age of 21.93 (SD=2.89). Among them, four had previously programmed a robot. The low-achiever participants were on average 21.75 years old (SD=0.83) and none had programmed a robot before.

3.4 Discussion

This study was conducted to understand whether and how inexperienced human teachers may gradually become more skilled robot teachers through repeated unsupervised kinesthetic teaching of a humanoid robot, with only limited instructions and with no detailed, direct feedback provided on their kinesthetic teaching approach (to address O1). Participants only received, after each teaching session, coarse-level feedback on their overall teaching performance in Tasks 1 and 2, via the robot providing one out of three pre-defined demonstrations of a similar task. When participants experienced kinesthetic teaching of a robot in a set of cleaning tasks in multiple sessions, we studied the success of demonstrations, changes in the time it takes them to provide their demonstrations, their gaze pattern, aspects of how they manipulate the robot's arm, and their perceived workload.

3.4.1 [RQ1.1-1.2] Teaching Success and Duration

According to the patterns observed in the success of demonstrations of table-cleaning tasks throughout the sessions, we discovered that our non-expert participants could be categorized into three clusters. For 15 out of 28 participants, i.e., improving participants, after they taught the robot more tasks over the five sessions, their demonstrations covered

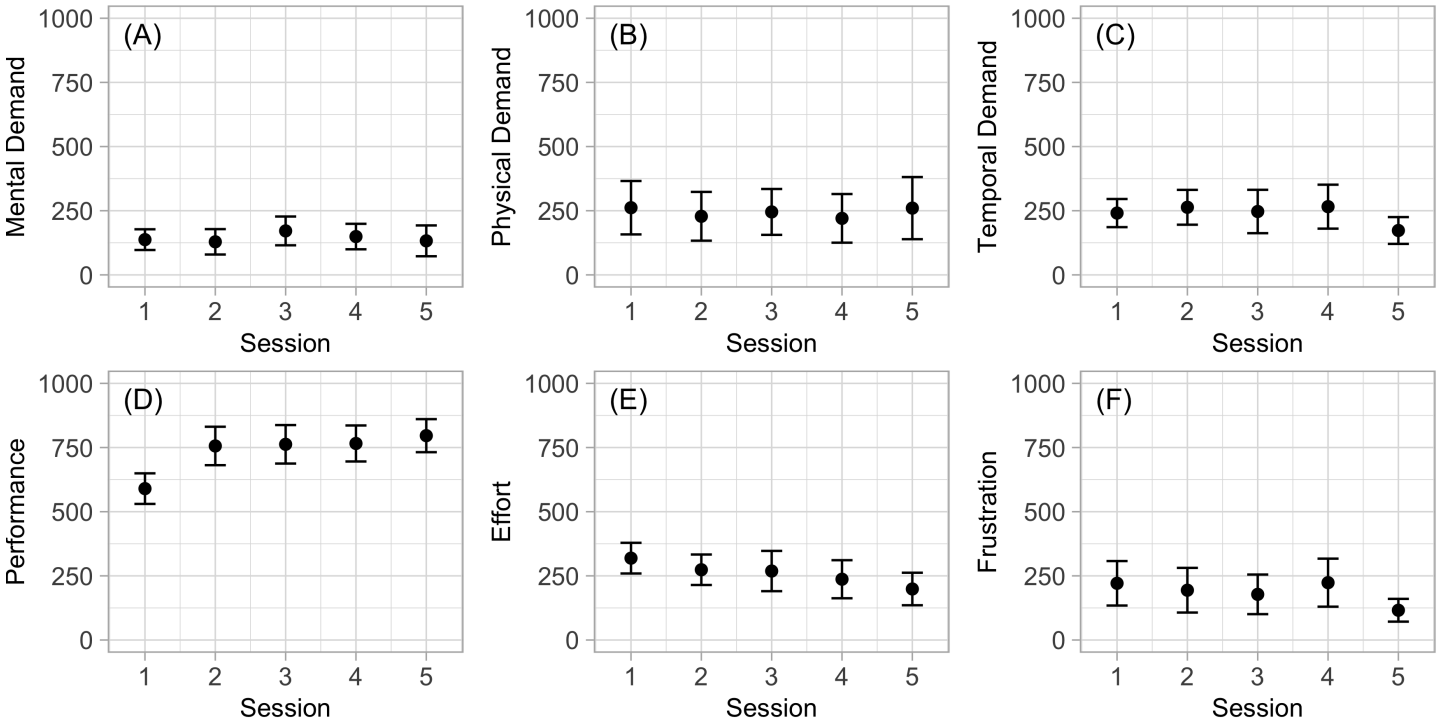


Figure 3.10: Average ratings of NASA-TLX questionnaire items over the five sessions. Error bars indicate 95% confidence intervals

Table 3.5: Linear Mixed-effects Model predicting NASA-TLX questionnaire items

	Mental Demand				**Physical Demand**				**Temporal Demand**			
Covariate	**Estimate**	**SE**	**t**	**p**	**Estimate**	**SE**	**t**	**p**	**Estimate**	**SE**	**t**	**p**
Age	-15.090	7.501	-2.012									
Gender	8.509	37.156	0.229		1.948	71.486	0.027		78.513	43.626	1.800	
Avg. time	1.813	0.768	2.361	*	3.019	1.557	1.940		1.092	1.146	0.953	
Avg. gaze hand	-0.315	1.056	-0.298		-1.588	2.092	-0.759		0.874	1.568	0.557	
	Performance				**Effort**				**Frustration**			
Covariate	**Estimate**	**SE**	**t**	**p**	**Estimate**	**SE**	**t**	**p**	**Estimate**	**SE**	**t**	**p**
Session					-18.512	7.382	-2.508	*				
Gender	-16.418	51.585	-0.318		1.327	54.389	0.024		0.010	50.664	0.000	
Avg. time	-3.069	1.137	-2.700	**	3.299	0.989	3.335	**	2.276	1.202	1.895	
Avg. success	4.685	1.159	4.044	***								
Avg. gaze hand	-1.240	1.534	-0.809		0.6447	1.2203	0.528		-0.894	1.627	-0.549	

* = $p < .05$; ** = $p < .01$; *** = $p < .001$.

significantly larger areas, which was more effective, as a result of them using a larger range of the robot arm's movements. 9 out of 28 participants, i.e., high-achievers, could provide successful demonstrations from the start and so showed only small improvements over the sessions. The presence of 'improving' and 'high-achiever' participants in kinesthetic teaching based on certain measures of teaching success was also observed in the study of Sakr et al. (2022) [164], referred to as 'slow-adapters' and 'fast-adapters,' respectively. In our study, 4 out of 28 of the non-expert participants, i.e., low-achievers, consistently failed to provide highly successful kinesthetic demonstrations during the five sessions. Such a cluster was not found in [164]. Note that in our study, all the participants were non-expert robot users (although some had experienced basic robot programming, e.g., in high-school), whereas in [164] participants had different levels of robotics expertise. Here, all the low-achievers were from the participants who did not have any experience with robots.

Overall, the majority of participants in this study achieved higher success in demonstrating the tasks while they spent less time teaching the tasks. The success of teaching was improved even within the sessions when teaching the robot four tasks. According to these findings, which address RQ1.1 and RQ1.2, the majority of participants in our study were able to teach robots more efficiently in terms of success and duration as they gained more experience. A high level of adaptation to robot teaching was observed in 15 out of 19 participants who were not successful in teaching the robot at first (both low-achievers and improving participants), despite the fact that teaching the robot was done in an unsupervised setting and participants did not receive any detailed, direct feedback on each teaching task. It should be noted that those conclusions could be limited to the sample we studied, as discussed later in Section 3.5.

Regarding the considered confounding variables, teaching time was positively affected by the number of changes in the manner in which participants manipulated the robot's arm and how much they used both their hands for teaching. The first effect could be expected, since every change in the teaching strategy may take time for participants to reflect upon, not only for repositioning the arms but also for thinking of a new manipulation strategy. Yet, one might expect that using both hands at the same time should lead to easier robot arm manipulation and therefore shorter demonstration times, contrary to what we observed. A possible explanation could be that after participants used both their hands more, they were more careful in moving the robot's arm and, therefore spent more time providing the demonstrations. This is in line with another finding of ours that suggested more usage of both hands led to more successful demonstrations. An alternative explanation could be that, as participants might have been moving multiple joints when using both hands, and since the robot had earlier asked them to be careful about not colliding its body parts, they performed the task more carefully, thus increasing the time

of demonstration. Future studies may uncover how and why the speed of providing physical demonstrations of tasks may be influenced by the usage of both hands for teaching.

Note, given that the usage of both hands was found to positively affect teaching time, it might seem contradictory that over the sessions, participants spent less time teaching, when they used both hands more for teaching. In fact, all the significant factors included in the LMMs jointly affect teaching time. In this case, because the negative effect of session number on teaching time reported in Table 3.2 has a much greater magnitude compared to the positive effect of the usage of both hands, affected negatively by session number and positively by the usage of both hands combined, teaching time decreased in later sessions.

3.4.2 [RQ1.3] Human Gaze Patterns During Teaching

To test our hypothesis with regards to RQ1.3, we analyzed the percentage of participants' gaze fixations that were directed at the robot's hand. The data shows that participants gazed less often at the robot's hand overall, with gaining more experience in teaching the cleaning tasks to the robot, both over different sessions and within single sessions. These findings are in line with [168], where expert users spend less time fixating on the robot's gripper when kinesthetically teaching a robot how to pour pasta from a cup into a bowl. Our hypothesis H1 is therefore confirmed.

A possible explanation for the observed decline in the percentage of fixations on the robot's hand with repeated teaching can be that participants paid more visual attention to other parts of the robot, e.g., to the robot's joints to monitor its motions to produce more useful trajectories, or to the robot body's parts in order not to cause self-collisions, after gaining more experience. In the present study, the gaze target of the participants was not examined further to determine where exactly human teachers looked when they paid less attention to the robot's hand, as this highly time-consuming analysis was not a focus of the study. For a more in-depth gaze analysis, the main challenge was to localize more regions in view in dynamic scene camera videos. Here, we implemented a method to track the robot's hand, which was easier due to its special shape that was not changing during teaching. Other parts of the robot's arm, e.g., forearm, upper arm, and shoulder (see Figure 3.2), were all white, may have been occluded by the participant's own arm during teaching, and may have looked very different from various angles observed by participants during teaching, thereby making them difficult to track computationally. In alternative studies, one could attach trackable labels to different robot parts to use with computer vision systems, but that could introduce bias by shifting participants' attention and, therefore, was avoided in our study. Another possible explanation for the decline in the percentage

of fixations on the robot's hand could be that participants looked ahead to predict the outcomes of the actions and decide on other aspects of their demonstrations. In light of all these, the gaze data collected in our study may be further examined using more advanced computer vision techniques in future studies.

There were also other interesting observations with regards to the gaze behaviour of human teachers. As illustrated in Figure 3.8, a lower average gaze at the robot's hand was detected during teaching Task 2 as compared to the other tasks. If the above-mentioned assumption is true, which suggests participants paying less attention to the robot's hand implies checking the joints' motions more to avoid self-collisions, then our observation might be justified. Task 2 was one of the challenging tasks, as can be noticed by the lower average success of teaching compared to Tasks 3 and 4 (see Figure 3.5) and its task space was extended compared to Task 1. Those challenges and also the extent of the task may have required participants to look more at the joints or at the body parts of the robot, and not at the robot's hand, during teaching. This could also suggest that the participants were able to adjust their strategies themselves based on the nature of the tasks, without any instruction or feedback on their performance.

3.4.3 [RQ1.4] Behavioural Data on Manipulation and Teaching Strategy

The analyses of data extracted from videos of participants teaching the robot (addressing RQ1.4) suggested that improving participants were more likely to use both hands for teaching and more often guided the robot's upper arm as they got more robot teaching experience. These findings suggest that participants whose success was substantially improved over the sessions changed their strategies and adapted to robot teaching in an unsupervised setting in terms of their method of physically guiding the arm of the robot. Manipulating the robot's arm with both hands and moving the robot's upper arm itself could possibly improve teaching performance as it enables teachers to cover a larger area in their physical demonstrations, which seems to have been discovered by the participants after they gained experience, as noted in Section 3.4.1.

It was also observed that as improving participants gained more experience in kinesthetic robot teaching within the sessions, they were more likely to hold the robot's hand during teaching. Some instances when teachers were not guiding the robot's hand while demonstrating how to perform a physical task are when improving participants might have simply been adjusting the joint of the robot's arm, without contributing to the task directly. Therefore, as not guiding the robot's arm decreased in the sessions, improving

participants might have spent less time merely repositioning joints in the robot's arm as they taught the robot more and more tasks in a short period of time.

In our experiment, participants employed many different strategies to physically manipulate the robot's arm, and we therefore could not cluster human physical behaviours while teaching into a few types (as was done in [103, 107]). However, in our study, significantly fewer changes in the teaching strategy of improving participants, in terms of how they manipulated the robot's arm, were observed in later sessions, and in later tasks within the sessions. This implies that after teaching more tasks, improving participants came up with more consistent teaching strategies and employed them from that point onwards in future robot teaching tasks in the experiment.

In this study, we found certain differences among the three clusters and observed patterns in the behaviours of high-achievers and improving participants that might have contributed to the success and effectiveness of teaching the robot, which also suggested that the majority of participants could learn how to adapt their strategies after exploring robot teaching without any detailed, external feedback on their behaviour. Low-achievers mostly used only one hand for kinesthetic teaching and did not guide the robot's upper arm as much as the other participants. Plus, they seemed to have changed their teaching strategy less often than others in the initial teaching tasks. We need to raise some caution with the interpretation of the data regarding some of the clusters, especially low-achievers, since fewer participants fell into this cluster compared to improving participants (see Section 3.5 for discussion on this limitation).

Note that some characteristics of the robot, e.g., its anthropomorphic arm design, may have influenced participants' strategies for manipulating the robot's arm. This is because participants were adapting to the forward model of the robot's arm, with the kinematics being comparable to that of humans. The size of the robot, the noise generated while moving its arm, and the weight and the force response of the robot's arm could all impact different aspects of the results. Even the positioning of the robot relative to the participants could impact the suitability of different methods for grabbing and moving the robot's arm. In certain cases, e.g., standing on the right side of the robot while guiding its left arm, the usage of both hands for teaching could have been more difficult.

3.4.4 [RQ1.5] Human Teacher's Workload

To study RQ1.5, we asked participants to self-evaluate their workload for each of the five teaching sessions. We used the NASA-TLX workload index for this purpose. The only scale that was found to be affected by the experience of participants over the sessions was the

perceived effort. Our results show that by experiencing robot teaching in multiple sessions, the tasks were perceived as less effortful for the participants: participants felt that it was easier to accomplish the tasks in later sessions. This could be expected because when we practise a certain type of task repeatedly over time, we tend to perceive less effort doing it, so the task will require a lower cognitive load [186]. It is also noteworthy that temporal demand was not affected by the average time spent on providing the demonstrations, thereby supporting that participants actually did not perceive any time limit was imposed on their demonstration. This was expected as the participants were not given any time limits, they could spend as much time as they wished on their teaching.

There were a few confounding factors that affected certain dimensions of human teachers' workload when teaching tasks to the robot. Achieving a higher success rate in demonstrations led to a higher performance rating by the participants. This shows that participants' own evaluation of their performance in the teaching tasks was consistent with external evaluations (i.e., evaluation of success through pictures).

Moreover, by spending more time teaching the robot, participants reported higher perceived effort and mental demand. The same effect on physical demand was also close to being statistically significant ($p = .055$). All these trends could be expected, as time spent on performing a task may influence perceived effort and demand, e.g., by increasing fatigue. We also observed that with using both hands more often for teaching, demonstration times increased. Using both hands, which can be more mentally and physically demanding, might have led to an increase in the above-mentioned scales of participants' workload. Contrary to those trends, we found that longer demonstrations led to a lower evaluation of performance, perhaps because it indicates less efficiency in performing a task. An effect approaching significance ($p = .061$) of teaching time on perceived frustration was also observed which might suggest that participants became more frustrated as they spent more time teaching. All these observations can be investigated further in future experiments with greater statistical power, as they may suggest trade-offs between time and success of demonstrations and perceived workload and effort. The effect of participants' fatigue on the reported workload measures can be further examined in future research.

3.4.5 Other Reflections

[RQ2] The Question of Who to Imitate

Another research objective pressed in this thesis is to detect patterns in non-experts' teaching that a robot can use to identify the most suitable teachers (O4). To contribute to this

objective, we defined RQ2 and investigated our data at a more individual level and identified three clusters of high-achievers, improving, and low-achievers, based on participants' success in teaching two of the challenging tasks in our experiment, i.e., how to clean areas of a table. According to this diversity in teaching styles for our cleaning tasks, a human teacher who uses both their hands for teaching, guides the robot's upper arm (and even more frequently after repeated interactions), and dynamically adjusts their way of holding a robot's arm is more likely to belong to a high-achiever cluster and thus provide successful task demonstrations. As discussed in Section 3.4.5, a robot may be able to capture such data to make inferences and potentially identify better human teachers.

Measuring Teaching Success Through Robot's Joint Data

The success of demonstrations included in our analyses was estimated subjectively, based on my evaluation of pictures of the tasks after teaching was completed, to determine how much of the surfaces were cleared by participants' kinesthetic demonstrations. However, we explored the possibility of estimating demonstrations' success obtained from a computational approach which was based on automatically processing the robot's joint motion data. Although the computational estimates were highly correlated with the subjective success of the demonstrations, they were noticeably lower than the subjective evaluations in Tasks 1 and 2. It is possible that this difference is due to the computational estimates not accounting for particle interactions (interactions among those particles involved in Tasks 1 and 2) and the three-dimensional nature of the tasks, thus affecting only those tasks. For example, a few pieces of paper could be pushed aside easily at once if they were close together. Because there were no particles involved in cleaning the whiteboard, the sponge could only clean the areas that came into contact with the whiteboard when enough force was applied. Despite this shortcoming, which suggests that the accuracy of the proposed computational method could be affected by the nature of the task, the computational method still allows tracking teachers' progress and success automatically and reliably by the robot itself, despite the possibility of underestimating the success of demonstrations. Thus, this method could be useful in future experiments, applications, and/or field studies, as it does not require manual coding of data. This method can be useful, particularly in realistic settings that lack an accurate vision system, where varying and unpredictable lighting conditions exist (e.g, in a real home, healthcare, or work setting) or when a vision system is not desirable due to privacy considerations, such as in a home.

Data Collection by the Robot Itself

Data analyzed in this experiment (excluding the questionnaires) were obtained from two primary sources aside from the robot: eye-tracking glasses to record participants' gaze targets and videos of the sessions to identify the location of participants' hands on the arm of the robot, and also the success of demonstrations. However, a remote gaze-estimation system [179], if installed on a robot, e.g. [180], could replace eye tracking glasses to evaluate the gaze direction of a human teacher. The robot may then compare that gaze direction with the location of its hand, to detect how long the teacher has fixated on the robot's hand during teaching. A robot equipped with touch-sensitive skin [95] on its arm, e.g., an updated iCub robot [137], could detect if two areas of its arm are touched, to track the usage of both hands, and sense if its upper arm is guided, or its hand is not guided, while being kinesthetically taught by a human teacher. This would eliminate the need to have a camera and manually annotate videos to obtain data on manipulation strategies, while also gathering information that can be informative for understanding performance and behaviour change of non-expert teachers. For the success of demonstrations, we offered here a computational alternative to subjective evaluations of snapshots of the tasks. Sensing and analyzing the data presented in this work by a robot itself would enable it to track the progress of its teachers during teaching, and this information might be useful for learning selectively from available teachers.

Cleaning Tasks Used in this Experiment

Some of the reported results suggest specific differences between the tasks that involved cleaning the whiteboard (Tasks 3 and 4) and those that involved cleaning an area of a table (Tasks 1 and 2), as reflected in (a) the higher success rates of Tasks 3 and 4, even in the early sessions, (b) a lower percentage of fixations on the robot's hand when teaching Tasks 1 and 2 in a majority of sessions, and (c) more frequent changes in the teaching strategies when teaching Tasks 1 and 2 compared to Tasks 3 and 4 in early sessions. This implies that participants might have found it harder to come up with a successful way of teaching Tasks 1 and 2. Teaching cleaning tasks on the table might have been more challenging than on the whiteboard, possibly due to differences between the same tasks done on horizontal versus vertical surfaces. Because the order of the tasks was fixed, another possibility for easier teaching of Tasks 3 and 4 could be that participants got used to teaching the robot the first two tasks before moving on to the other two tasks.

3.5 Limitations

The presented work had several limitations. To begin with, and most importantly, the robot in our experiment had no actual learning mechanism. As a result, the robot was 'pre-programmed' to provide coarse-level feedback to participants by executing a single task at the end of every session with the purpose of giving the impression that their demonstrations had an impact on the learning of the robot. In this situation, the robot's behaviour did not match exactly each participant's demonstration skills during the session, despite being selected by the experimenter based on the participants' performance. Note that if the robot had not given any feedback after participants taught the 4 tasks, the experiment could have been boring for participants, especially since they had to participate in five sessions.

After being told in the debriefing stage that the robot was not actually learning, a few participants mentioned they had suspicions about the learning capabilities of the robot. According to them, this was mainly due to the fact that the robot's performance was not exactly the same as their demonstrations. However, this impression might be the case even with a real robot learning system: the resulting learned robot behaviour would likely be based on a set of demonstrations and therefore, the learned behaviour would be different from one's own instructions. Still, if the robot in our experiment had been capable of actual learning in between sessions and providing feedback based on what was learned from participants' teaching in the previous session, this feedback would have affected how participants would teach the robot in subsequent teaching sessions. Such a system could help participants gain a better understanding of how their instructions help the robot improve its performance.

This study also had limitations regarding data analysis. There might be a risk of bias in the self-reported data used for assessing participants' workload. Moreover, as noted in Section 3.2.3, the gaze recordings of participants were only processed according to the amount of time they fixated on the hand of the robot. Localizing other areas of interest, e.g., the arm and body of the robot, could allow a more comprehensive analysis of human teachers' gaze during robot teaching. The behavioural data on manipulation and teaching strategy could also be analyzed more precisely, by considering multiple states for the placement of the human teacher's hand. To improve these aspects in the future, computational data processing using artificial intelligence techniques may be required instead of manual coding of the data, which presented its own challenges (e.g., being a lengthy process, ensuring sufficient agreement between two raters, etc.).

3.6 Future Directions

In this study, we discovered that the majority of our inexperienced and unsupervised participants could gradually become more proficient robot teachers by repeatedly teaching a humanoid robot, and we explained how this might happen, based on the physical characteristics of the actions and behaviours of the human teachers. Future studies may recruit participants with various levels of experience with robots to study how exposure to robots and different participant backgrounds could affect teaching behaviours. Moreover, the results of the present study did not indicate why certain participants (i.e., low-achievers) did not exhibit high demonstration success. The potential role of factors such as motivation in human teachers for robots can be investigated by future research, e.g., through interviews with participants.

Including actual robot learning in such a work is an important topic to study in the future, which was not the focus of our present experiment (but is later considered in Phase 3 of our research). Involving a robot equipped with a LfD system could further explore the relationship between a robot learning while being taught skills by human teachers, and how the human teachers' teaching changes over time. When using a real learning robot, it may be necessary to reconsider some aspects of the present study. For instance, some other metrics for evaluating the success of demonstrations, mostly concerning machine learning aspects, e.g., [175, 50], could be used. Moreover, due to a lack of precision in sensing and actuation, the Pepper robot might not be the best choice for kinesthetically learning fine motor skills, including forces. However, there is a trade-off between using a robot, such as Pepper, suitable to be used safely in realistic, unsupervised scenarios, as compared to more sophisticated and capable robots, possibly with more DOFs or non-human-like configuration of the joints, that cannot be used safely and reliably in settings other than laboratory and/or highly supervised environments. The adaptation of non-experts when teaching different types of robots can be explored in future research.

The robot in our study only provided minimal instructions to participants on how to teach it, and the behavioural feedback it gave included prerecorded actions (although they were matched with the participant's performance by the experimenter). In future work, such feedback can be generated based on what has been learned, in case a robot is capable of actually learning. In the absence of robot learning, prerecorded actions for the behavioural feedback could be selected by the robot itself, e.g., by automatically assessing the number of particles cleaned from the tasks. Efforts can also be made in the future to improve humans teaching robots by utilizing certain strategies, e.g., by means of the robot providing social and communicative feedback to the teachers using verbal and/or non-verbal cues such as force feedback, or enabling social learning through feedback that

communicates the robot's learning process or internal states [33]. A robot could adjust its strategies according to the type of user among the identified clusters it believes it is interacting with.

In this chapter, we presented experimental findings from a sample of 28 human teachers in the specific context of teaching a robot how to clean surfaces. In the future, context and domain should be considered when determining whether a human teacher is a good choice for a robot to learn from. People might have strengths and weaknesses in certain domains, and therefore, show varied performances when teaching robots different tasks. Personality, education, and other individual characteristics might play a role, too.

3.7 Summary and Conclusion

Phase 1 of this research targeted human participants who had no previous experience teaching robots and examined how repeated human-robot interactions in physically teaching a robot can affect the speed and success of teaching, human teachers' gaze behaviour, manipulative strategies, and perceived workload; mainly to address O1. In other words, we asked if non-experts can adapt their strategies and learn how to teach robots through repeated kinesthetically teaching of a robot. Based on the results, the answer was: 'the majority of non-experts can.' While nearly one-third of non-expert participants were good at teaching the tasks to the robot from the start, the results showed that it was not too difficult for a majority of the other participants to adapt to robot teaching and improve their teaching without being given any formal training, but rather by experiencing robot teaching on their own over a few sessions of unsupervised teaching.

Although there is evidence from the literature that it is possible to train human teachers explicitly in order to improve robot teaching [43, 163], this work demonstrated that allowing participants to explore the teaching task and adapt to a robot over time is effective as an implicit, unsupervised strategy of training human teachers. Allowing users to explore robot teaching, rather than being trained on how to do it, might be a good option in certain settings and applications, particularly where safety risks for both humans and robots are low. The results further illustrated how exactly changes in multiple aspects of human teachers' teaching behaviour might occur as they became more experienced robot teachers. These insights could be used by robots in the future to track their human teachers' progress and experience.

It was found that non-expert robot users react in diverse ways when kinesthetically teaching a robot over multiple sessions. Some may be good from the start, some may

improve, and some may struggle throughout. The results showed that some aspects of human teachers' behaviour might be useful for identifying the most suitable teachers for robots, which contributed to addressing O4. Roboticists could use the proposed methods and findings to enable a robot to track human teachers' progress and level of experience, so that the robot can identify better human teachers (i.e., high-achievers) to potentially result in better robot learning outcomes.

To take one step further towards achieving the research objectives, in Phase 2 of the research, presented in the next chapter, I designed a robot learning framework that can be used by non-experts to teach household tasks to a humanoid robot, which was not included in the experiment presented in this chapter.

Chapter 4

System Development: A Biologically Inspired Program-Level Imitation Approach for Robots

In this chapter, I report the design of a novel Program-Level Imitation (PLI) learning architecture for robots, to address O2. This approach enables robots to receive diverse task demonstrations from human teachers, identify possible sequences of known actions in the demonstrations, organize sub-goals into sequential structures, and apply this knowledge for effective task execution based on user-defined requirements in new situations. Since this learning architecture focuses on high-level action plans, the robot in this chapter is programmed to have pre-existing low-level visuomotor control and sensing capabilities to perform specific sub-goals and detect the effects related to these sub-goals on objects within its view. To learn a new task, the PLI system analyzes the observed sequences of sub-goals and synthesizes feasible arrangements of robot primitives, along with useful relationships between the sub-goals, in the form of a 'program.' For task execution, the robot evaluates possible arrangements of sub-goals through mental simulation and selects the most suitable sequence of actions for the present situation and user preferences (e.g., prioritizing speed, safety, etc.). The robot then uses information contained in the program to accomplish sub-goals in the selected sequence to achieve the overall task goal.

The present chapter outlines the core principles of PLI architecture (Section 4.2), details its implementation on an iCub humanoid robot (Section 4.3), and then evaluates and discusses the robot's performance when learning and executing two abstract real-world tasks (Sections 4.4 and 4.5) using both qualitative observations and quantitative measures.

At the end of this chapter, limitations of our system as well as proposed future research directions are noted in Sections 4.6 and 4.7, followed by a summary and concluding remarks (Sections 4.8)[6].

4.1 Research Questions

The goal of the research presented in this chapter is to address O2. We therefore ask:

RQ1 How can we design, implement, and evaluate a learning framework for robots to enable them to understand the sequential structure of multi-step tasks and efficiently perform such tasks?

4.2 Program-level Imitation Architecture

This section describes our proposed PLI architecture for robots, which consists of modules illustrated in Figure 4.1. In summary, for learning a new task, *demonstrations observer* processes human-provided demonstrations and identifies the occurrence of sub-goals (explained in Section 4.2.1). *Program generator* then analyzes the demonstrations to determine how different arrangements of sub-goals can lead to the task goal (explained in Section 4.2.2). This module obtains a program that includes multiple possible sequences of sub-goals to complete the task and relations between them.

When a user asks the robot to carry out a learned task, *sequence selector* evaluates the available sequences based on user-selected functions that impose their expectations (described in Section 4.2.4), considering the current task state obtained from *task progress observer* (outlined in Section 4.2.3). The most suitable sequence in the program is then executed by *action planner* (discussed in Section 4.2.4).

This architecture relies on a robot's pre-programmed abilities to achieve certain basic 'sub-goals' with respect to objects and places, as well as the robot's ability to detect the 'effect' associated with the completion of each known sub-goal in the environment using its own sensors. Only these basic actions, called 'primitives' of the robot, will be used for learning and performing new multi-step behaviours. Objects and places involved in each task need to be introduced to the robot before learning. Each primitive is a goal-directed

[6]A short video summarizing the entire work presented in this chapter can be found at: https://ieeexplore.ieee.org/document/11005968/media

small action involving a set of objects and places. Designers of a robot may program functionalities of acting to perform specific sub-goals and detecting their achievement so that a robot comes with such capabilities out-of-the-box.

In this work, given the abilities of an iCub robot [137], we programmed three primitives of *grasping an object and putting it inside a box*, *pressing a button*, and *sliding a box to a different place*. While the details of our robotic implementation are given in Section 4.3, the primitives will be referred to in this section to explain the architecture. We use the following notation to indicate the sub-goal associated with the achievement of each primitive, using IDs assigned to known objects and places:

- $O^{X,Y}$: Grasping object X and putting it inside box Y
- B^{X}: Pressing button X
- $P^{X,Y}$: Sliding box X to place Y

4.2.1 Demonstration Observer

The *demonstrations observer* module identifies and records the occurrence of effects associated with each known primitive, i.e., achievement of sub-goals, in demonstrations of a given task by processing robot sensory data. The method by which the sub-goals are extracted is specific to the particular robotic implementation, according to the desired modalities for teaching and the robot's capabilities for sensing. A robot may detect sub-goals by using its vision system, by monitoring its own body state while undergoing kinesthetic teaching [16], or even by receiving verbal instructions. A robot may utilize more than one modality at the same time for sub-goal detection, potentially with a mechanism for resolving disagreements between the modalities. Sub-goals may also include information about characteristics of objects for identifying them, e.g., their colour or size. Details of our present implementation of *demonstrations observer* on an iCub robot are presented in Section 4.3.2.

To facilitate data processing, *demonstrations observer* records each demonstration in the form of multiple lines of text, each line listing certain sub-goals, and writes it to a text file containing all previous demonstrations of a given task. Sub-goals are indicated with the notation introduced earlier in this section. The first line of a demonstration lists the achieved sub-goals in the initial state of the tasks (i.e., sub-goals that exist before the demonstration is started), derived from *task progress observer* module, plus the first detected sub-goal in the demonstration. Every new line in a demonstration is for the

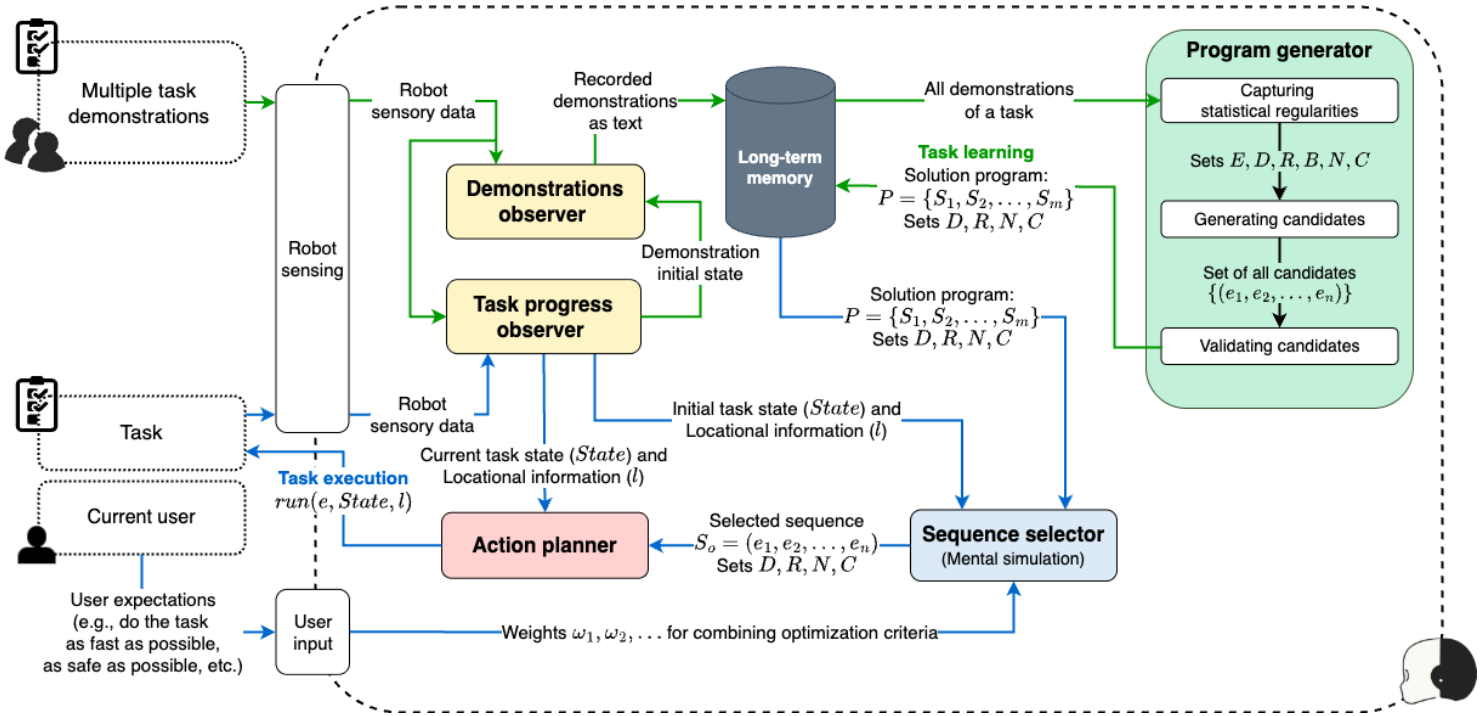


Figure 4.1: The overall architecture of our proposed PLI architecture. For learning a new task (follow green lines), multiple task demonstrations are recorded by *demonstrations observer* in the robot's long-term memory, with the initial task state obtained via *task progress observer*. *Program generator* then processes the collection of recorded demonstrations and generates multiple sequences of actions leading to the task along with useful relations between sub-goals, i.e., solution program. When the robot's current user asks for the execution of a task (follow blue lines), the most suitable sequence will be selected by *sequence selector*, based on the current user's expectations from the robot and the present state of the real task obtained from *task progress observer*. *Action planner* then initiates robot primitives to achieve the task goal by following sub-goals according to the sequence in the program. © 2025 IEEE.

addition of a sub-goal or cancellation of a previous sub-goal (Figure 4.3 and 4.4 present examples). A sub-goal in a demonstration is said to be 'newly-achieved' when it is at the right-most place of a line. In every line, all other sub-goals are said to be 'already-achieved.' Accordingly, the list of already-achieved sub-goals in the first line of each demonstration is referred to as its 'initial state.' The list of all sub-goals in the last line of a demonstration is referred to as the 'final state' of a demonstration. After a demonstration is finished, the teacher should indicate whether the task goal was successfully achieved or not, and this will be recorded in an extra line dividing the demonstrations in the text file containing all demonstrations of a given task. This module also considers:

- *Cancellation of a previous sub-goal:* Effect related to an already-achieved sub-goal may be undone when another sub-goal is achieved. For example, when a box is moved, the effect of the box being in its original place will be undone. In this particular case, as coded in the *demonstrations observer*, the new sub-goal will be added in one line, and the omission of the already-achieved sub-goal will be recorded in another line.

- *Repeated achievements of a sub-goal:* A sub-goal that has previously appeared in a demonstration (as a newly-achieved sub-goal or in the initial state) may be detected again later on. In this case, the repeated sub-goal is recorded using an index. Repetitions of a sub-goal will be treated as unique sub-goals for program generation, as different rules may be applied to each of them.

- *Missing information:* Some recorded sub-goals may have missing information. For example, *demonstrations observer* may detect that an object has been moved to a box, but the module is uncertain which box. In this case, the missing target will be noted by the letter X instead of the ID. It is up to the implementation in which situations sub-goals with missing information should be recorded (e.g., what degree of uncertainty is acceptable).

- *Programs as sub-goals:* A teacher may instruct the robot to include a previously learned program in the task demonstration. In this case, a single sub-goal denoting a program will be recorded (e.g., Pro^{boil_potato}). As needed, the teacher may either perform the actions themselves, ask the robot to perform the actions, or skip that part.

For demonstrating a task to the robot, a human teacher can choose to activate one specific learning modality or let the robot acquire information from multiple modalities. With only one active learning modality, sub-goals will be added to demonstrations immediately.

However, with multiple modalities involved at the same time, the agreement between the modalities will be checked first. Upon detection of any sub-goal using any modality, the sub-goal, time of detection, and modality of detection will be appended to a list. Every member of this list will be periodically checked to see if there is no other sub-goal detected by another modality within a few seconds that contains conflicting information regarding the objects and places. If there is no conflicting sub-goal, the sub-goal will be added to the demonstration. Otherwise, a mechanism will be activated to resolve disagreements among the modalities (e.g., verbally asking the teacher to repeat that part), or there will be missing information in the sub-goal.

Here is an example demonstration. Assume button 1 is already pressed before the demonstration starts. A human teacher moves box 1 to place 2 as the first step in the demonstration. Next, they grab object 2 and put it in box 1. Then, they move object 1 to another box, but *demonstrations observer* can not detect exactly which box. Once done, they move box 1 to place 1, which will undo the effect of that box being in place 2. They then press button 2. Lastly, they press button 1 one more time and indicate to the robot that the demonstration is successfully done. In this case, *demonstrations observer* writes the following eight lines to the file containing all demonstrations of the task:

$B^1, P^{1,2}$
$B^1, P^{1,2}, O^{2,1}$
$B^1, P^{1,2}, O^{2,1}, O^{1,X}$
$B^1, P^{1,2}, O^{2,1}, O^{1,X}, P^{1,1}$
$B^1, O^{2,1}, O^{1,X}, P^{1,1}$
$B^1, O^{2,1}, O^{1,X}, P^{1,1}, B^2$
$B^1, O^{2,1}, O^{1,X}, P^{1,1}, B^2, B^1_2$
$[Success]$

4.2.2 Program Generator

The *program generator* module, as illustrated in Figure 4.1, processes the recorded demonstrations of a task and captures statistical regularities (e.g., relations between sub-goals defining constraints, requirements, etc), as inspired by ethological explanations of PLI in primates presented in Section 2.1.3. Utilizing the statistical regularities (e.g., relations between sub-goals defining constraints, requirements, etc), *program generator* helps the robot to understand how arranging sub-goals leads to the task goal. To do so, this module determines various possible sequences of sub-goals along with some relations between the sub-goals, which is referred to as a solution 'program.' This process involves three steps: *Capturing statistical regularities*, *generating candidates*, and *validating candidates*.

Capturing statistical regularities

Program generator first examines the text file containing all demonstrations of the task and finds sets or relations (containing sub-goals or pairs of sub-goals) defined in Table 4.1. In summary, E will be a set of sub-goals observed in the majority of the demonstrations that reached the task goal (i.e., where successful) starting from the smallest initial task state. Members of E will constitute the n sub-goals in the candidate sequences ($|E| = n$). This is to avoid over-imitation, i.e., imitation of unnecessary actions with regards to the goal demonstrated by a model [92]. If a sub-goal is unnecessary, only a minority of demonstrations likely contain it and therefore, will not be in E.

The system also checks the final state of the recorded demonstrations and finds set D of sub-goals observed in the final state of the majority of successful demonstrations started from the smallest initial task state, ignoring the repetition of sub-goals. Accordingly, $D \subseteq E$ since D contains the sub-goals observed in the majority of demonstrations except for those that were cancelled at some point before the demonstrations ended. Set D will help ensure that the validated sequences result in the desired final task state (i.e., task goal). Note that a validated sequence may have repeated sub-goals such that the final state D is achieved as desired. Sets R, B, N, $C \subseteq \{(e_i, e_j) \in E \times E \mid e_i \neq e_j\}$ define relations between the sub-goals in E, using the definitions in Table 4.1, to help the system find possible ways of arranging sub-goals and efficiently execute the sub-goals in a sequence. Note that relations R and B define constraints for arranging the sub-goals, while relations N and C help the robot validate a candidate sequence of sub-goals and execute the sequence.

To specify 'majority' in the definition of each set in Table 4.1, *program generator* uses 'majority percentages' α_E, α_D, α_R, α_B, α_N, and α_C in $[0, 100]$. In case all the demonstrations are flawless (i.e., there is no recording error, extra or missing sub-goal, etc.), the system may work with $\alpha_E = \alpha_D = \alpha_R = \alpha_B = \alpha_N = \alpha_C = 100$, meaning that relationships between sub-goals are captured from regularities present in all the demonstrations. Nevertheless, setting the parameters to less than 100% will allow a robot to effectively capture information by ignoring erratic instances in the demonstrations. These parameters can therefore be fine-tuned to adjust learning.

As an optional step, relations R and B can undergo a 'filtering' procedure, with parameter t, to reduce the likelihood of inclusion of incorrect regularities. To do so, $(e_i, e_j) \in R$ will be deleted if there exist at least t demonstrations in which e_i is newly-achieved while e_j is already-achieved, or in other words, if sub-goal e_j is achieved before e_i in t instances. Moreover, $(e_i, e_j) \in B$ will also be deleted if there exist at least t demonstrations in which e_i is already-achieved while e_j is newly-achieved. In case *demonstrations observer*

Table 4.1: List of sets (or relations) identified by program generator for extracting statistical regularities from task demonstrations. The definition of each set is used to enable the system to identify the sets. The purpose of understanding each item is also noted. © 2025 IEEE.

Set/ Relation	Definition (strategy of finding)	Purpose
E	$e_i \in E$ if an effect e_i is observed in the majority (α_E%) of successful demonstrations that start from the smallest initial state of the task.	To identify sub-goals necessary for the task and avoid imitation of unnecessary sub-goals.
D	$e_i \in D$ if an effect e_i, regardless of its repetition, is in the final state of the majority (α_E%) of successful demonstrations that start from the smallest initial state of the task.	To identify sub-goals required to be in the final task state to consider the task complete.
R	$(e_i, e_j) \in R$ if $e_i \in E$ is already-achieved before $e_j \in E$ in the majority (α_R%) of demonstrations when e_j is newly-achieved.	To capture sequential regularities when sub-goals e_i may need to be achieved before another sub-goal e_j.
B	$(e_i, e_j) \in B$ if $e_i \in E$ is never already-achieved in the majority (α_B%) of demonstrations when $e_j \in E$ is newly-achieved and if $(e_j, e_i) \notin R$.[1]	To account for sequential regularities when the presence of sub-goal e_i may block some other sub-goal e_j from being achieved.
N	$(e_i, e_j) \in N$ if $e_i \in E$ is never newly-achieved or in the initial state of the majority (α_N%) of demonstrations that have $e_j \in E$ in their initial state and if e_i is not repeated in E.[2]	To describe cases when the achievement of sub-goal e_j may obviate the need for achieving some other sub-goal e_i.
C	$(e_i, e_j) \in C$ if in the majority (α_C%) of demonstrations where sub-goal e_j is newly-achieved, e_i is no longer already-achieved in the next line.	To consider when achievement of sub-goal e_j may cancel out the effect of sub-goal e_i.

Newly-achieved sub-goal: Sub-goal in the right-most place of a line in a demonstration; *Already-achieved* sub-goal(s): Sub-goal(s) in a line in a demonstration except for the right-most one; *Initial state* of a demonstration: Already-achieved sub-goal(s) in the first line of the demonstration.

[1] Not allowing inverse of relations in R to be in B is to avoid capturing duplicate information in two sets. Nevertheless, some sub-goals may physically block other sub-goals and therefore be part of B, but these blocking sub-goals may not be present in the majority of the demonstrations to be part of R.

[2] Relation N does not carry information about sub-goals that are repeated in E. This is because if demonstrations with different numbers of repetitions of a certain sub-goal exist, it may not be easy to identify exactly which one is not needed in case another sub-goal is achieved.

is implemented so that it ensures the order of detected sub-goals is accurate, but some sub-goals may be missing or wrong, filtering can improve the system's accuracy. Through experiments presented in Section 4.4.2, we will explore the impact of adjusting majority percentages and applying the filtering procedure on program generation.

Generating candidate sequences

The *program generator* module then generates candidate sequences of sub-goals to later validate against the captured statistical regularities to check if they satisfy constraints in R and B, taking into account relations D and C. In the basic form, all permutations of members of E ($|E|!$ candidates), each as a sequence of n sub-goals in the form of $S = (e_1, e_2, ..., e_n)$, will be considered for validation.

Depending on the particular application of the architecture (discussed in Section 4.5), there can be an optional step called *grouping* to narrow down the number of candidates by considering some sub-goals as a block. To do so, set R' will be constructed from R after every member $(e_i, e_j) \in R$ is replaced by (E_i, e_j), where $E_i = \{e \in E \ |(e, e_j) \in R\}$. In other words, all pairs of sub-goals with a common second element will be combined and replaced by one pair with a set as its first element. Then, every $(E_i, e_j) \in R'$ is replaced by (E_i, E_j), where $E_j = \{e \in E \ |(E_i, e) \in R'\}$. R' will be the set of ordered paired 'sets' of sub-goals with the same definition as R: $(E_i, E_j) \in R$ if every $e_i \in E_i$ is already-achieved before every $e_j \in E_j$ in the majority of demonstrations when e_j is newly-achieved. With a similar procedure but in the reversed order, set B' of paired 'sets' of sub-goals will be constructed from B. The module then forms set $H = \{E_1, E_2, ...\}$ containing every E_i and E_j seen in $(E_i, E_j) \in R' \cup B'$ (all first and second elements of members in R' and B'). using members of H, which are possibly overlapping subsets of E, E will be divided into some disjoint sets, i.e., partitioned[7]. The set containing the partitions will be called $E' = \{E_1, E_2, ..., E_k\}$. Here, $E = \cup E_i$, and $\forall E_i, E_j \in E', E_i \cap E_j = \emptyset$. Partitioning E according to the process of grouping ensures there are no constraints in R or B between any two members of $E_i \in E'$ (meaning there is no $(e_i, e_j) \in R \cup B$ such that $e_i, e_j \in E_i$). With grouping, every permutation of members of E' ($|E'|!$ candidates) forms a candidate, after each set $E_j \in E'$ is replaced by a random permutation of its members (since the order of sub-goals within E_i does not matter according to R and B). Again, a candidate will be a sequence of n sub-goals $S = (e_1, e_2, ..., e_n)$.

[7]To find the partitions, $2^{|H|}$ sets of all possible intersections of some members of H will be considered first (each member of H is either included in the intersection or not). Then, for every $e \in E$, the smallest set among the $2^{|H|}$ intersections that contains e will be selected. These sets will constitute the partitions.

Validating candidates and saving the program

Every candidate sequence of sub-goals will be checked by Algorithm 1 to determine if it meets requirements in R and B, taking into account C. Note, in this algorithm "e_i is before/after e_j in Γ" means the index of e_i in sequence Γ is smaller/larger than the index of e_j in Γ. Furthermore, $e_{j,k}$ refers to the $k-$th repetition of sub-goal e_j. The sequence is also checked to ensure that completing sub-goals based on the defined order leads to a final state S equal to the expected final state D. If $S \neq D$, because candidate sequence Γ is constructed with all sub-goals in E and $D \subseteq E$, some sub-goals must have been cancelled due to the particular arrangement of sub-goals in the sequence. In this case, extra steps will be taken in Algorithm 1 (lines 23-34) to check if repeating sub-goals that might have been previously cancelled may change the final state: first, every sub-goal that could be cancelled and contributes to the last state becoming D will be considered for being added to the candidate sequence. This extra sub-goal will be added with an index one step larger than the existing one in the candidate. Then, all the permutations when the added sub-goal is in a place that is after the existing one and is in such a way that the new sequence results in a new final state will be validated.

Ultimately, the *program generator* module saves the set of validated sequences $P = \{S_1, S_2, ..., S_m\}$ along with sets D, R, N, and C associated with the task as the solution program in the robot's long-term memory. If *grouping* strategy is used, all $\prod(|E_i|!), E_i \in E'$ possible sequences resulting from each candidate in all permutations of members of E', called a 'family' of sequences, are included in P. In other words, sequences will be generated with all possible arrangements of sub-goals within each group. For easier visualization when using the *grouping* strategy, in this chapter, a family of sequences will be presented as a sequence that may contain 'sets' of sub-goals when some sub-goals are grouped together. For example, a family of sequences shown as $(P^{3,1}, \{O^{7,3}, O^{2,3}\}, B^5, P^{3,2}, B^1)$ referees to 2 sequences: $S_1 = (P^{3,1}, O^{7,3}, O^{2,3}, B^5, P^{3,2}, B^1)$ and $S_2 = (P^{3,1}, O^{2,3}, O^{7,3}, B^5, P^{3,2}, B^1)$, because the order of achieving $O^{7,3}$ and $O^{2,3}$ may be altered. Section 4.4.1 provides examples to illustrate the program generation process.

4.2.3 Task Progress Observer

The *task progress observer* module obtains the 'current state' of the task, defined as the set of sub-goals detected to be completed at the present moment, to be used by *demonstrations observer*, *sequence selector* and *action planner*. If set D corresponding to a learned program P was among the detected sub-goals, Pro^P will be added to the task state. The method by which the sub-goals are detected is unique to every robotic implementation. Robot sensory

Algorithm 1 Validating candidates

Input: L: List of all candidates, sets D, R, B, and C. **Output:** P: List of validated candidates.

Function $validate(L, D, R, B, C)$:
Initialize P as an empty list
for each sequence Γ in L **do**
    Initialize R_okay as True
    Initialize B_okay as True
    for each $r = (e_i, e_j)$ in R **do**
        if e_i is after e_j in Γ **then**
            $R\_okay \leftarrow$ False
        end if
        if e_i is before e_j in Γ and there exists e_k such that $i \leq k \leq j$ and $(e_i, e_k) \in C$ **then**
            $R\_okay \leftarrow$ False
        end if
    end for
    for each $r = (e_i, e_j)$ in B **do**
        if e_i is before e_j in Γ **then**
            $B\_okay \leftarrow$ False
        end if
    end for
    if R_okay and B_okay are both True **then**
        Initialize set S as the final state of sub-goals in Γ
        if $S = D$ **then**
            Add validated candidate Γ to P
        else
            for each $(e_i, e_j) \in C$ such that $e_i \in D - S$ **do**
                Initialize L' as all permutations of members of Γ by adding $e_{i,k+1}$, where k is the largest index of e_i in Γ
                for each sequence Γ' in L' **do**
                    Initialize set S' as the final state of sub-goals in Γ'
                    Initialize set S'' as the final state of sub-goals in Γ' without $e_{i,k+1}$
                    if $S' \neq S''$ and $e_{i,k+1}$ is after $e_{i,k}$ in Γ' **then**
                        $P' = validate(L', D, R, B, C)$
                        Add validated candidate in P' to P
                    end if
                end for
            end for
        end if
    end if
    if *Grouping* is used for generating candidates **then**
        for each sequence p in P **do**
            Replace p with sequences resulting from every permutation of sub-goals within each group
        end for
    end if
end for
Return P

Algorithm 2 Following a sequence in program P

Input: Sequence $S = (e_1, e_2, ..., e_n)$, sets R, N, and C in P,
l_i: Locational information about objects/places related to sub-goals in S, core objective function f.

Function $follow(S, R, N, C, l_i, f)$
Initialize *Step* as 1
Initialize set *State* as the current state of the task (obtain from *task progress observer*)
for each $1 \leq step \leq n$ **do**
  Initialize $S_c = (e_1, e_2, ..., e_{Step})$
  if $e_{step} \notin State$ **then**
    Initialize k as 1
    for each $1 \leq k \leq step$ **do**
      $State \leftarrow$ Current state of the task
      Initialize $S_{c,k} = (e_k, e_{k+2}, ..., e_{Step})$
      Initialize Do as True
      if $e_k \in State$ **then**
        $Do \leftarrow$ False
      end if
      if There exists $(e_k, e_j) \in N$ such that $e_j \in State$ **then**
        $Do \leftarrow$ False
      end if
      if There exists $(e_k, e_j) \in C$ such that $e_j \in S_{c,k}$ **then**
        $Do \leftarrow$ False
      end if
      if Do is True **then**
        if e_k is not a program **then**
          Call $r(e_k, State, l_k, R)$
          Call $f(e_k, State, l_k)$
        else
          Request *sequence selector* to choose the most appropriate sequence S_t in e_k
          Call $follow(S_t, State, l_k, R)$
        end if
      end if
      $k \leftarrow k + 1$
    end for
  end if
  $Step \leftarrow Step + 1$
end for

Function $r(e_k, State, l_k, R)$:
for each $(e_i, e_j) \in R$ **do**
  if $e_k = e_j$ and $e_k \notin State$ **then**
    Call $r(e_i, State, l_i, R)$
    Call $f(e_i, State, l_i)$
    $State \leftarrow$ Current state of the task
  end if
end for

data can be processed by approaches similar to those used by *demonstrations observer* to detect effects associated with the completion of sub-goals.

4.2.4 Sequence Selector and Action Planner

With multiple sequences available in a program for a task, *sequence selector* and *action planner* modules select and execute the most appropriate sequence every time the robot is asked to execute the task. Both these modules utilize a special process called 'following' a given sequence $S = (e_1, e_2, ..., e_n)$, as defined in Algorithm 2 and explained in more detail later in this section, with two different purposes. *Sequence selector* follows all sequences via 'mental simulation' to decide which sequence leads to the lowest total cost associated with the actions. Mental simulation of sequences takes place by taking into account the specific user's preference in what aspect of robot performance matters most (e.g., task completion time, safety, energy consumption, etc.), the current state of the physical task, and the location of objects and places involved. *Action planner* then follows the selected sequence to guide the robot to perform physical actions according to the program.

The process of following a sequence using a program, compared to linearly acting based on sub-goals $e_1, e_2, ...$ in the sequence, is a procedure that ensures only necessary sub-goals are performed, in a desired order and with minimized extra actions. This process involves repeatedly calling a core objective function $f(e, State, l)$ for a sub-goal e, the current state of the task ($State$) (i.e, set of all detected sub-goals when observing the task), and information about the location of relevant objects and places (l).

The *sequence selector* module follows every sequence S_i in the program with each available 'optimization' function $g_k(e, State, l)$ used as the core objective function. $g_k(e, State, l)$ adds a cost associated with actions that produce sub-goal e to an overall cost $J_i(g_k)$ for the entire sequence S_i. To select a sequence, the total expected cost of performing the task according to every sequence is calculated with all available core optimization functions $g_1(), g_2(), ...$ The overall costs obtained by each optimization function for all sequences are then normalized to be in the range $[0, 1]$ (noted by $\hat{J}_i(g_k)$). The particular user asking the robot to perform the task can provide weights $\omega_1, \omega_2, ...$ for adding together the normalized overall costs generated by each optimization function, to define criteria for the selection of a sequence based on which aspect(s) of robot performance matters most. The sequence selector will select sequence S_j with the smallest total cost, i.e., $S_j = \arg\min_{S_i \in P}(\omega_1 \hat{J}_i(g_1) + \omega_2 \hat{J}_i(g_2) + ...)$. An example of this procedure is given in Section 4.4.1.

The optimization functions are formulated by the robot designers with respect to robot

primitives and may be task-specific. These functions may maintain a history of task states and locations while following a sequence of actions. Some examples are introduced in Section 4.3.4. In mental simulation, task state and relevant locations, including the expected positioning of each arm of the robot, are updated after sub-goal e_i is achieved so that for the next sub-goal e_{i+1}, $g(e_{i+1}, State', l')$ produces an estimate using the updated task state ($State'$) and location of objects and places (l').

Action planner uses another core function $run(e, State, l)$ that executes a robot primitive for acting to achieve the sub-goal e, considering the current task state, and the location of objects and places (l).

Details on following a sequence

Algorithm 2 provides details on the n steps for following sequence $S = (e_1, e_2, ..., e_n)$. In every step, the algorithm considers sub-sequence $S_c = (e_1, e_2, ..., e_{Step})$. If the last sub-goal in S_c is not yet achieved (i.e., $e_{Step} \notin State$), every sub-goal e_k in S_c, from e_1 to e_{Step}, will be checked and the core function $f(e_k, State, l_k)$ will be called, except for when at least one of the following three conditions are met: (1) If $e_k \in State$, meaning that sub-goal e_k is already-achieved; (2) If there exists $(e_k, e_j) \in N$ such that $e_j \in State$. In this case, the presence of sub-goal e_j in the current state of the task may obviate the need for achieving e_k; (3) If there exists $(e_k, e_j) \in C$ such that $e_j \in S_{c,k} = (e_{k+1}, e_{k+2}, ..., e_{Step})$. In this case, the presence of sub-goal e_j will later cancel out the effect of e_k, and therefore, there may be no need for e_k to be achieved.

In following a sequence, in case sub-goal e_k is a program (not a robot primitive), *sequence selector* will choose and follow the most suitable sequence S_t within first. Moreover, before every $f(e_k, State, l_k)$ call, $r(e_k, State, l_k, R)$ defined in Algorithm 2 is called. $r(e, State, l_k, R)$ is a recursive function that ensures the chain of sub-goals that need to be achieved before sub-goal e_k are all achieved. This function calls $f(e_j, State, l_j)$ for every unaccomplished sub-goal e_j that needs to be achieved before e_k (i.e., $(e_j, e_k) \in R$).

When following a sequence $S = (e_1, e_2, ..., e_n)$, the strategy of going back and reviewing sub-goals in sub-sequence $S_c = (e_1, e_2, ..., e_{Step})$ step-by-step offers a number of benefits in real-world situations. Due to robot errors, some actions may not produce the desired effects, or the environment may change during robot actions, or when the robot stops performing the task and resumes after a while. Furthermore, performing some actions may result in the cancellation of some previously produced effects by mistake. This way, the robot will execute the required actions to re-achieve any missing sub-goals before moving

forward. Tests presented in Section 4.4.3 demonstrate the behaviour of a robot with the PLI system if interruptions occur.

4.3 Robotic Implementation

This section presents how we implemented the PLI architecture on an iCub robot [137] to evaluate the system and analyze the learning and execution of two abstract physical tasks, described in Section 4.3.1. iCub is a humanoid robot designed for research in embodied cognition, with applications in HRI experiments and, in particular, robot learning [137]. iCub is capable of recording demonstrations and performing our physical abstract tasks placed on a table in front of it using 38 DOFs in its upper body (including arms, hands, and neck). More specifically, this robot can perform primitive actions that involve picking up objects, pressing buttons, and sliding boxes, as each hand of the robot includes nine DOFs with the ability to control thumb, index, and middle fingers independently, while the remaining two fingers share a single DOFs to provide stability. The iCub operates using the YARP middleware [136] that enabled communication between multiple C++ and Python scripts for implementing different modules of the PLI architecture.

An Intel RealSense camera [106] was mounted on the robot's head for 3D localization when learning and performing tasks. An ArUco marker was attached to each known object or place in tasks, so that the RealSense camera could obtain 3D positions of the centre of each marker in the camera frame. To do so, a Python script was developed that could track and find the centre of each marker using the ArUco library in OpenCV [74, 158] and interface with the camera using the pyrealsense2 wrapper [6]. Locations in the camera frame were then converted to locations in iCub's base frame according to the current configuration of the robot's upper body.

The three primitives listed earlier in Section 4.1 were incorporated into the system[8]. This section describes how the perception capabilities of the robot were utilized to create a *demonstrations observer* module that acquires demonstrations through remote observation. A *task progress observer* module was also designed to evaluate the task state through remote observation. We also describe how the robot could act to achieve the sub-goals in a program, via a *action planner* module. All the noted modules were implemented in C++, while the *program generator* module was implemented in Python. Lastly, we define three examples of core optimization functions for the *sequence selector*.

[8]Code is available at: https://github.com/paliasgh/icub-pli

4.3.1 Abstract Physical Tasks

Considering the capabilities of iCub in performing actions, we designed and built physical models of two conceptual, abstract tasks to test our implementation: a *potato boiling* task and an *organization* task, shown in Figure 4.2. The ID assigned to each object is the ID of the attached ArUco marker.

Task 1: [**Potato boiling task**] In this task, a part of a simple kitchen environment is modelled. Movable box 3 is considered to be a pot that can be in three places: place 0, on the counter-top; place 1, in the sink, and place 2; on the stove-top. Any other positioning of the box will be unknown to the robot (i.e., no sub-goal will be detected). The box is attached to a sliding bearing for ease of movement. Objects 2 and 7 are presumed to be potatoes. Pressing button 5 is assumed to add water to the pan at once when the pan is in the sink. Pressing button 1 is assumed to turn on the stove-top. The goal of this task is to place the pot containing water and potatoes on the hot stove-top.

Task 2: [**Organization task**] This task models a drawer in a freezer. Movable box 3 is assumed to be a drawer that can be in two places: place 0, closed inside the fridge; and place 1, open outside the fridge. To fully open the drawer, a lock needs to be released by pressing button 1. When the drawer is open, access to its inside is blocked by an obstacle (e.g., a big item in the fridge). There is a hidden space, stationary box 0, that can be accessed only if the drawer is open. This constraint is modelled by a cover that is attached to the movable box 3 and slides on top of box 0. Objects 2 and 6 are assumed to be meat packages, while object 7 is assumed to be an ice cream. The goal of the task is to put the ice cream in the drawer and the meat packages in the hidden space and close the drawer.

4.3.2 Demonstrations Observer for iCub

As described earlier in Section 4.2.1, *demonstrations observer* module analyzes robot sensory input to identify and record sub-goals in human-provided demonstrations. In the present implementation, a human teacher can perform the task in front of the robot while the robot is observing the task through its vision system to recognize known sub-goals in the demonstration[9]. This observational learning takes place using data from the RealSense

[9]I also tried to implement a kinesthetic learning modality for the iCub robot. However, after a prototype was created, I observed that the gravity compensation module of the robot was not robust enough to allow easy movement of the robot's arm. Moreover, tactile sensors of the robot were not sufficiently accurate to indicate to the robot when its arm is grabbed, so that the robot changes the joints' control mode in certain parts of the arm into compliance mode to allow kinesthetic teaching. I therefore decided to only keep the observational learning modality.

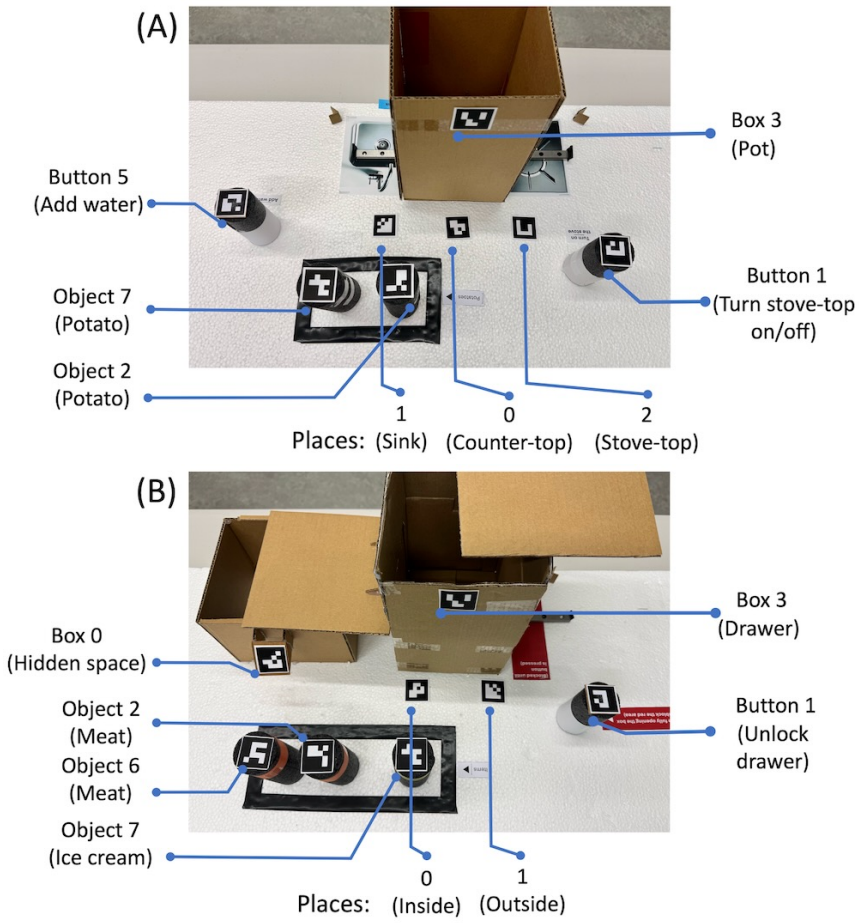


Figure 4.2: Layout of the abstract tasks used for system evaluation: (A) potato boiling task and (B) organization task. Objects and places and their assigned ID are indicated. © 2025 IEEE.

camera. With information about the location of objects and places, *demonstrations observer* unitizes the following rule-based strategies for the detection of sub-goals related to each robot primitive in task demonstrations:

- **Putting object X in box Y:** During observational learning, the robot detects object X is moved into box Y if the object disappears for 3 seconds, as long as box Y was the closest box to where object X was last located, compared to other boxes, and this distance was shorter than a threshold. When evaluating the task state, in case there is only one box in the defined items, object X is considered as moved into box Y if it has disappeared for 3 seconds. If there is more than one box and the object is not visible in an area that defines a box, the target box remains unknown and the task state is reported as explained in Section 4.3.3.
- **Pressing button X**: In observational learning, the robot detects button X is pressed if the z-component of its location has been smaller than the average of this value in the past 2 seconds (indicating that the button has been moving downward) and the condition no longer holds (meaning the button is still). As part of evaluating the task state, the robot considers button X is pressed if the z-component of its location is below a certain value.
- **Sliding box X to place Y**: In observational learning, the robot detects movable box X is moved to place Y if the y-component of its location stays within 2 cm of the y-component of the location of place Y for 1 second. In the task state, the robot considers movable box X to be in place Y if the y-component of its location is within 2 cm of the y-component of the location of place Y.

Note that the list of available items and their type (four types in our case: movable box, stationary box, button, and object), as well as the list of places, with the ID of the ArUco marker attached to each of them, is given to the system beforehand. At the end of the demonstration, the teacher can save the demonstration by pressing a specific button indicating that the task goal was completely achieved or delete the demonstration or another button to indicate the task goal was not completely accomplished.

4.3.3 Task Progress Observer for iCub

Relatively similar strategies to those used by *demonstrations observer*, noted in Section 4.3.2, are employed for detecting the effects related to known sub-goals in the task state via observing the location of objects and places. Given the limitations of the iCub robot, in some

situations, the task state cannot be exactly determined. For example, when at least two boxes are known to the robot, it may not be possible to determine in which box an object that is not in the view is located, because the robot may not see what is inside each box. In such a case, a sub-goal with missing information will be sent to *demonstrations observer*, while all possible sub-goals will be sent to the *sequence selector* and *action planner*.

4.3.4 Sequence Selector for iCub

As noted in Section 4.2.4, *sequence selector* uses core optimization functions to find the optimal sequence(s) of sub-goals in a program. These functions are used when mentally simulating following a solution program and add a cost associated with the achievement of each sub-goal to a total cost. In the present implementation, three core optimization functions are considered: $path(e, State, l)$ estimates the total length of robot hand movements in acting to achieve sub-goals, $risk(e, State, l)$ quantifies the risk of robot failure when performing the actions, and $eff(e, State, l)$ evaluates the time efficiency of the actions when transferring items to a freezer.

- $path(e, State, l)$ returns the sum of Euclidean distances between each end-location of the robot's moving hand when performing a sub-goal. Optimizing total path-length is expected to result in faster task completion when the robot's hand moves at a constant speed.

- $risk(e, State, l)$ returns a quantified risk of failure (1 or 0) when performing some sub-goals. In our implementation, because the robot picks up objects with its left arm, there will be a risk of collision if there exists an object close enough to the left side of an object. Such an instance is detected when there is no sub-goal in the current task state indicating the absence of a close-enough object on the left side of another object.

- $eff(e, State, l)$ is a task-specific optimization function (for task 2 only) and returns a cost associated with a delay in moving items to the freezer that may cause the frozen food to thaw. In case the item is an ice cream (object 7), the cost will be the number of sub-goals that need to be achieved before moving it to the freezer. To do so, the number of times the function call has been made with sub-goals other than $O^{7,0}$ when following a given program is kept track of, until $O^{7,0}$ is called. If the item is meat, the cost will be 0.1 times the number of sub-goals before $O^{2,3}$ or $O^{6,3}$, to indicate the higher urgency in moving ice cream to the freezer, to avoid melting.

For a meat package (objects 2 and 6), the cost will be one-tenth, compared to an ice cream, to indicate the higher urgency in moving ice cream to the freezer.

To let the robot calculate the weighted total cost and select the most suitable sequence for task execution, a two- or three-dimensional vector containing weights for the above-mentioned optimization functions is inputted by the user to the robot (illustrated with an example in Section 4.4.1).

4.3.5 Action Planner for iCub

The iCub robot's primitives to achieve the defined sub-goals related to each primitive by the *action planner* are implemented as follows:

- **Putting object X in box Y:** The robot finds the location of object X, approaches an adjusted location with its left hand and with fingers opened, moves its left hand closer, and finally closes the fingers. Then, it raises the left hand to the top of box Y and opens its fingers. In tests presented in Section 4.4.1, iCub only used a single arm joints and therefore, due to its narrow dexterous space, the robot can only perform this action if Y is in the left-most location. Therefore, if box Y was not reachable, the robot used its right hand to move the box within reach.
- **Pressing button X**: Assuming the button is horizontal, the robot finds the location of button X, moves its appropriate hand to 5 cm above the adjusted location with figures open, moves its hand without change of orientation 8 cm down and then backs up.
- **Sliding box X to place Y**: The robot moves its hand to a starting location, based on whether the box should be slid right or left, with its fingers open. It then starts moving that hand to an ending location, without changing the orientation, until box X reaches place Y, considering the y-component of their locations.

The motion planning was done automatically with the target locations and orientations, using iCub's Cartesian control interface. Note that the programming of the robot's actions involved experimentally finding some calibration values based on our setup, which were used to shift certain locations and distances or serve as starting/ending positions. In this work, due to technical issues, only arm movements of the iCub robot were allowed to perform actions.

4.4 System Validation

In this section, to evaluate the PLI system, we first describe examples of how a robot can generate programs based on multiple task demonstrations of the introduced abstract tasks and improve its performance (Section 4.4.1). We use task completion time and the estimated risk of robot failure as metrics to quantify the performance of the robot when learning tasks. Next, we evaluate the program generation when demonstrations contain errors by defining 'success rate' as a metric to evaluate sequences in the generated programs (Section 4.4.2). We conclude by describing scenarios in which the robot uses the acquired programs for task execution (Section 4.4.3).

4.4.1 Learning and Adaptation

To illustrate how a robot with our PLI system can learn new skills, a researcher took the role of a human teacher and provided multiple task demonstrations to the system. Because these tests were intended to describe the learning process, only correct demonstrations, recorded through observational learning, were used. The majority percentages were set to 100, without the 'filtering' strategy.

Example 1: Improving task completion time in Task 1

This example illustrates how the robot learns the abstract potato boiling task (Task 1) and uses its understanding of the sub-goals to improve task completion time. Figure 4.3 presents the learning outcomes, step by step, when six successful demonstrations were recorded. Here, the 'grouping' strategy was enabled. After receiving demonstration #1, the robot only knew how to follow the same order of sub-goals as seen in that single demonstration. Demonstration #2 did not include moving the pot to the counter-top ($P^{3,0}$), and potatoes were placed in reverse order. Figure 4.3 also includes an example timeline showing how demonstration #2 was broken down into six sub-goals, and seven lines were added to the text file containing the demonstrations. By processing the two demonstrations, the system determined $P^{3,0}$ unnecessary and excluded it from the program. Furthermore, the robot determined that the order of achieving sub-goals $O^{2,3}$ and $O^{7,3}$ does not matter and therefore grouped them together, which resulted in a family of 2 sequences in the program. With demonstration #3, in which the pot was moved to the sink at the beginning, the system discovered more possibilities, and the sequences fell into two families: $P^{3,1}$ comes before or after $O^{2,3}$ and $O^{7,3}$. Following demonstration #4, the robot discovered that

potatoes could be put into the pot at any time. Therefore, five families of sequences, as noted in Figure 4.3, became available.

Relations R, B, N and C, identified using 4 demonstrations, are shown as graphs in Figure 4.3. Relation R determines the sequence of $(P^{3,1}, B^5, P^{3,2}, B^1)$ as part of all sequences in the program. No relation includes any regularity for $O^{2,3}$ and $O^{7,3}$, which implies that these sub-goals can be achieved at any point. To allow the robot to further gain knowledge and form relation N, demonstrations #5 and #6 had non-empty initial states (i.e., some sub-goals were already achieved before demonstrations started). Demonstration #5 only included moving to pan to the stove-top ($P^{3,2}$) and turning it on (B^1) as the potatoes were already inside the pan with water. After this demonstration, the robot understood (captured in N) that the presence of sub-goals $O^{2,3}$, $O^{7,3}$, and B^5 obviates the need for $P^{3,1}$ (moving the pan to the sink). After demonstration #6, only the relation between B^5 and $P^{3,1}$ was kept in N, implying that there is no need to move the pot to the sink when there is water already in it. It should be noted that the extra regularities in N when there were five demonstrations, which implied $O^{2,3}$ and $O^{7,3}$ obviate the need for $P^{3,1}$, would not hinder task execution. This is because relation R will still mandate $P^{3,1}$ to be achieved before B^1 in case $P^{3,1}$ is skipped at its original place when following a sequence. Note, for pairs of sub-goals in relation R after all demonstrations, instances when one sub-goal was already achieved when another sub-goal was achieved are marked on all demonstrations in Figure 4.3.

In every step of this example, the sequence that was estimated to result in the shortest path-length was identified by the *sequence selector* through mental simulation of a real task using core optimization function *path*() introduced earlier in Section 4.3.4, at two initial states: $S_1 = \varnothing$ and $S_2 = \{P^{3,1}, O^{2,3}, B^5\}$ (i.e., one potato was initially in the pot which was in the sink with water in it). The most suitable sequence(s) according to the path-length optimization criterion and the simulated path-length(s) are also noted in Figure 4.3, up to when no further improvement was observed. With the empty initial task state (S_1), the shortest sequence of 5.36 m was obtained after three demonstrations, while with the non-empty initial task state (S_2), the first two demonstrations led to finding an optimal sequence that was estimated to need 2.90 m of robot hand movement. Given the physical arrangement of objects in Task 1 (see Figure 4.2 (A)), potatoes were at their shortest distance to the box when the pot was in the sink. Therefore, the optimal sequence by starting from S_1 was obtained after the robot arranged $O^{2,3}$ and $O^{7,3}$ after $P^{3,1}$. A sequence with this property appeared after three demonstrations. When starting from S_2, because the pot was initially in the sink, only the omission of extra sub-goal $P^{3,0}$ after the second demonstration led to the optimal sequence. In Figure 4.3, corresponding to every mental simulation of path-length, the robot's actual run time executing the given sequence

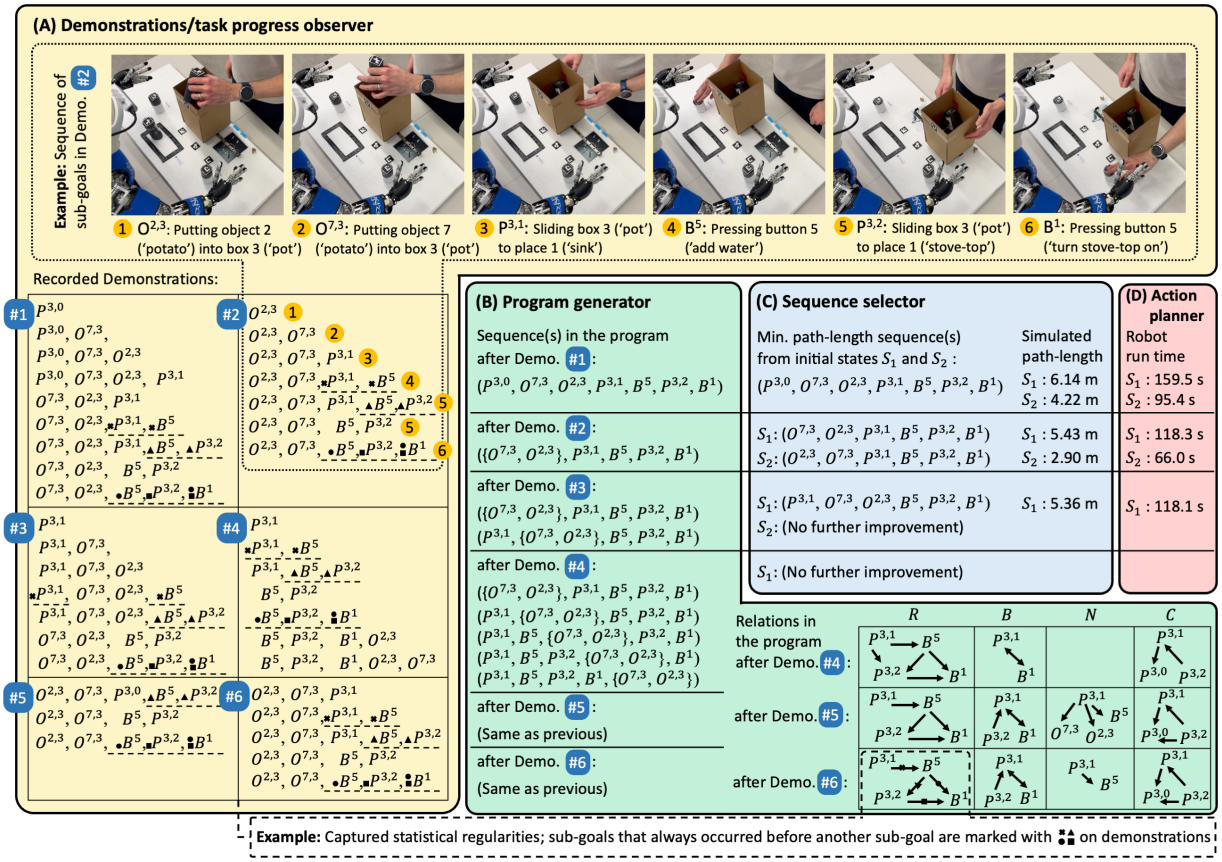


Figure 4.3: The process of robot adaptation to the abstract potato boiling task (Task 1) with six demonstrations. (A) *Demonstration observer* recorded demonstrations #1-#6, step by step. An example timeline shows how demonstration #2 was broken down into sub-goals and was recorded. (B) Programs were generated based on the entire collection of demonstrations up to each step. The sequence(s) in the program are shown. Relations R, B, N, and C in the programs after 4, 5, and 6 demonstrations are also presented. As an example, the statistical regularities seen in all demonstrations that are captured in relation R are marked on all demonstrations. (C) *Sequences selector* chose the sequence(s) in the program with minimum path-length via mental simulation, starting from two different task initial states S_1 and S_2. The selected sequence(s) and the estimated path-length(s) are presented to the point where the smallest possible estimate was achieved. (D) *Action planner* guided the robot to complete the task in multiple instances. Task completion times are provided to the point where no further improvement was achieved. Zoom in for more visible text. © 2025 IEEE.

is also included, showing that task completion time was reduced with shorter estimated path-lengths.

Example 2: Performing Task 1 more efficiently than each demonstration

In this small example scenario, the PLI system could find a sequence to complete the task more efficiently than each individual task demonstration. This is possible because the program is generated by applying the captured regularities within the sub-goals in the demonstrations to all possible arrangements of sub-goals. A test in three parts was performed with the robot learning our abstract potato boiling task (Task 1), similar to Example 1 (Section 4.4.1), considering the empty initial task state (S_1). In parts 1 and 2, the task was demonstrated to the robot following the sequences $(O^{2,3}, P^{3,1}, B^5, P^{3,2}, B^1, O^{7,3})$ and $(O^{7,3}, P^{3,1}, B^5, P^{3,2}, B^1, O^{2,3})$, respectively. Performing the task according to the sequence of sub-goals in each individual demonstration was estimated to require 6.26 m of robot hand movement. In part 3, a program was generated using both demonstrations that included 3 families of sequences: $(P^{3,1}, B^5, P^{3,2}, B^1, \{O^{7,3}, O^{2,3}\})$, $(P^{3,1}, B^5, P^{3,2}, \{O^{7,3}, O^{2,3}\}, B^1)$, and $(\{O^{7,3}, O^{2,3}\}, P^{3,1}, B^5, P^{3,2}, B^1)$. Sequence $(O^{2,3}, O^{7,3}, P^{3,1}, B^5, P^{3,2}, B^1)$ was estimated to need only 5.88 m of robot hand movement.

Example 3: Reducing the risk of failure in Task 2

In this example, the robot learned the abstract organization task (Task 2) and selected a sequence intending to minimize the risk of robot failure. Figure 4.4 presents sequence(s) in the generated program with four demonstrations. In this example, we examined both cases with and without the 'grouping' strategy. Upon receiving demonstration #1, the robot was only able to follow the same order of sub-goals as seen in that demonstration, which started from when the drawer was neither open nor closed (i.e., an empty task state). Demonstration #2 started when the drawer was already closed ($P^{3,0}$ was in the initial state of the task). The teacher first released the lock and opened the drawer (achieve B^1 and $P^{3,1}$, respectively) to place the ice cream in the hidden space ($O^{7,0}$) before placing the meat packages ($O^{2,3}$ and $O^{6,3}$), which required closing the drawer again ($P_2^{3,0}$). An example timeline describing demonstration #2 is included in Figure 4.4.

With the grouping strategy, the first two demonstrations led to three families of sequences, all with a repeated $P^{3,0}$ sub-goal (as this was seen in both demonstrations) and involving inserting the two meat packages in any order. The first two families of sequences consisted of ensuring the drawer is closed to store meat packages when no obstacle blocks

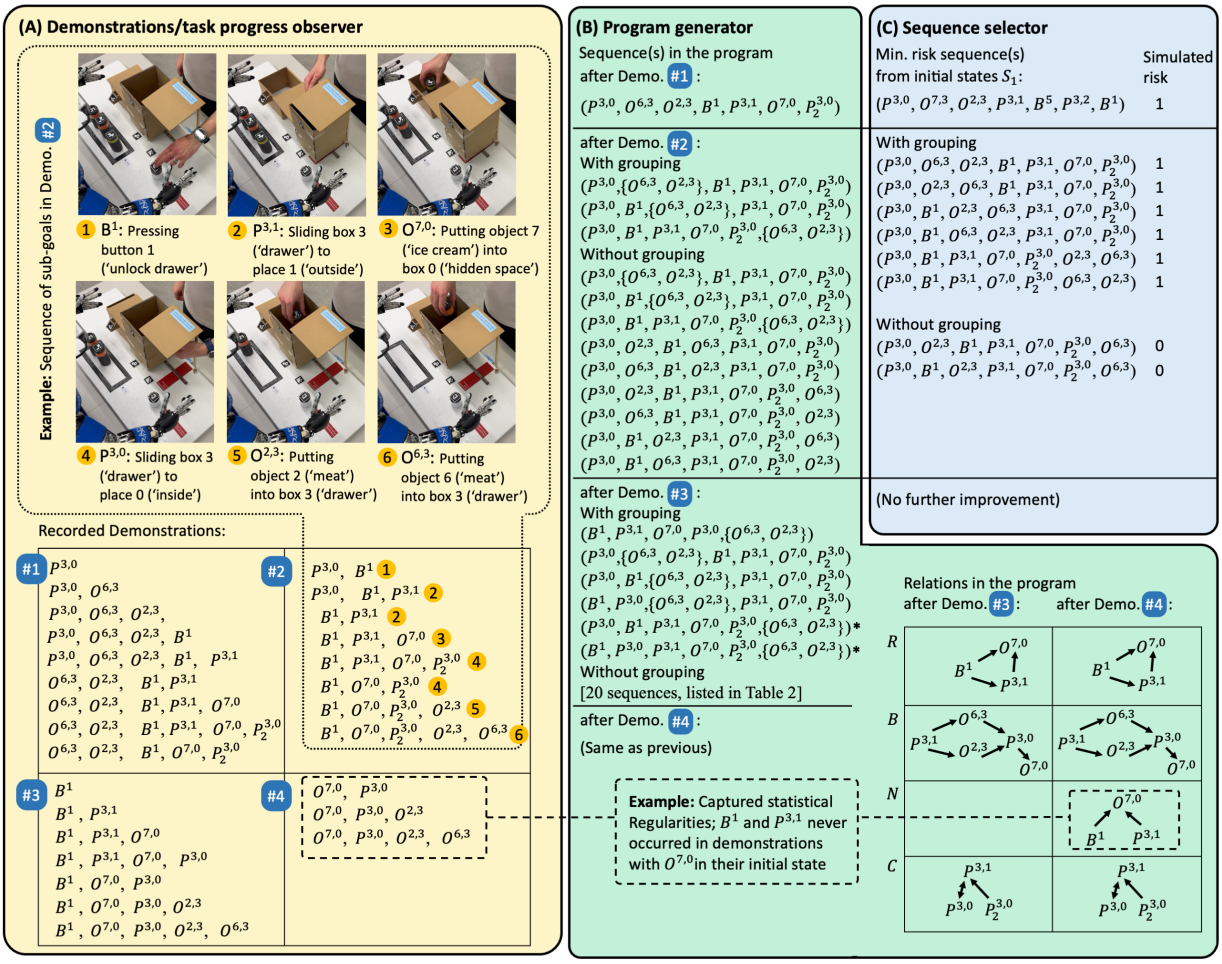


Figure 4.4: The process of robot adaptation to the abstract organization task (Task 2) with four demonstrations. (A) *Demonstration observer* recorded demonstrations #1-#4, step by step. An example timeline illustrates how demonstration #2 was broken down into sub-goals and was recorded. (B) Programs were generated based on the entire collection of demonstrations up to each step. The sequence(s) in the program are shown. Relations R, B, N, and C in the programs after 3 and 4 demonstrations are also presented. As an example, the statistical regularity captured in relation N is pointed out. (C) *Sequences selector* chose the sequence(s) in the program with a minimum estimated risk of robot failure via mental simulation. The selected sequence(s) and the estimated risk(s) are presented to the point where the smallest possible estimate was achieved. Zoom in for more visible text. © 2025 IEEE.

the drawer's entry, and then also unlocking the drawer, either before or after putting the objects, to be able to open it afterwards and put ice cream in the hidden space, and finally closing the drawer again. The third family of sequences leaves storing the meat packages ($O^{6,3}$ and $O^{2,3}$) for the end of the sequence after the drawer is closed again. In this family of sequences, because the meat packages are not going to be stored first, and therefore there is no need to close the drawer, sub-goal $P^{3,0}$ at the beginning of the sequence is unnecessary. This is because both demonstrations contained $P^{3,0}$ as the first sub-goal in addition to $P_2^{3,0}$ to keep the drawer closed at the end, so that $P^{3,0}, P_2^{3,0} \in E$.

To illustrate how this aspect could be improved, demonstration #3 was recorded, starting with an empty task and then achieving other sub-goals similar to the second demonstration. With three demonstrations, the solution program further evolved. Now, set E no longer included $P_2^{3,0}$. Therefore, with the grouping strategy, a family of sequences with only one sub-goal $P^{3,0}$ resulted. In this case, there existed valid arrangements of sub-goals that did not result in the desired task goal, and therefore, the *program generator* added $P_2^{3,0}$ to candidate sequences and produced five other families of sequences. These families were similar to the three families obtained with the first two demonstrations, but with allowing B^1 and $P^{3,0}$ to switch order when possible. Note that without demonstration #2, families of programs with unnecessary $P^{3,0}$ sub-goal, noted with '*' in Figure 4.4, would not have been generated.

We then provided a fourth demonstration, which started from an initial state where ice cream was already in the hidden space and the only next steps taken were closing the drawer ($P^{3,0}$) and inserting the meat packages ($O^{6,3}$ and $O^{2,3}$). The obtained relations, after three and four demonstrations were processed, are also shown in Figure 4.4. In this example, unlike Example 1, most of the information is captured in B, rather than in R. This is due to the presence of obstacles that could block access to the drawer and the hidden space. After demonstration #4, as noted on Figure 4.4, set N contained the information that when ice cream is already placed ($O^{7,0}$ is present), there is no need for unlocking or opening the drawer (B^1 or $P^{3,0}$).

As presented in Figure 4.4, *sequence selector* identified sequence(s) that were estimated as having the lowest risk of robot failure by mental simulation of a real task with core optimization function $risk()$ described in Section 4.3.4. The task's initial state was $S_1 = \{P^{3,0}\}$ and objects arranged from left to right were 2, 7, and 6, as shown in Figure 4.2 (B). Considering the close proximity of the objects, only sequences that include sub-goals related to objects with the order $O^{2,3}$, $O^{7,0}$, and $O^{6,3}$ could lead to a 0 risk. In this example, such sequences became available only without the grouping strategy, after two demonstrations. Note that $O^{2,3}$, $O^{7,0}$, and $O^{6,3}$ were never achieved in this particular order in any of the provided demonstrations.

Example 4: Optimizing Task 2 for different users

This example illustrates how the robot selects a sequence and adapts to different users with varying expectations. Here, the robot used the program of Task 2 learned in Example 3 with four demonstrations. Without the 'grouping' strategy, a total of 20 sequences, listed in Table 4.2, were available. The cost according to all defined core optimization functions ($path()$, $risk()$, and $eff()$) was estimated by mental stimulation of the task from an initial state similar to Example 3. Table 4.2 includes the estimated costs and also the normalized values for each. In this case, because sub-goal $P^{3,0}$ is already achieved and objects can be added only when the drawer is at a particular place, the estimated path lengths required for programs were relatively similar. However, the risk of hitting objects and the cost associated with the delay in moving frozen food items to the freezer varied among the sequences. For instance, six sequences that had sub-goal $O^{7,0}$ in the earliest possible place in the sequence of actions were determined to result in the lowest total $eff()$ cost function.

In this example, three different users, on three different occasions, requested the robot to complete Task 2. Each user imposed their expectations about which aspect of the task matters more by providing weights $[\omega_1, \omega_2, \omega_3]$ for each normalized cost so that a total cost could be calculated. User 1 asked the robot to consider all aspects of the task equally important by setting weights $[1, 1, 1]$. Sequence 16, $(P^{3,0}, O^{2,3}, B^1, P^{3,1}, O^{7,0}, P_2^{3,0}, O^{6,3})$, with a total cost of 0.90, was selected for this user. User 2 requested that the robot pays extra attention to prevent thawing the ice cream, by assigning a higher weight to the $eff()$ cost function (setting weights $[1, 1, 2]$). Either sequences 1, 7, or 13 could be selected for this user. Finally, user 3 asked the robot to pay extra attention to finishing as fast as possible by setting weights $[2, 1, 1]$ and sequence 17 was selected accordingly.

4.4.2 Error Handling

We examined how demonstrations containing errors impacted program generation and how adjusting the 'majority percentages' and applying the 'filtering' strategy, introduced in Section 4.2.2, could improve the learning performance. Multiple sets of 10 or 5 demonstrations were simulated computationally (instead of being recorded by the *demonstrations observer*), and the resulting program was then evaluated.

Simulating a set of demonstrations

A Python script was developed to produce demonstrations in the form of text files, with specific types and repetition of errors in a manner similar to *demonstrations observer*

Table 4.2: Robot sequence selection for the organization task according to three different user expectations. The program generated based on all four demonstrations in Figure 4.4, without grouping strategy, was used. The total estimated cost and normalized values according to core optimization functions $path()$, $risk()$, and $eff()$, explained in Section 4.3.4, are calculated and then added for each user, given the provided weights. The sequence having the minimum total cost (in boldface) was selected for each user. Weights $[\omega_1, \omega_2, \omega_3]$ were set to [1,1,1], [1,1,2], and [2,1,1] for Users 1, 2, and 3, respectively. © 2025 IEEE.

i	Sequence (S_i)	$J_i(path)$	$\hat{J}_i(path)$	$J_i(risk)$	$\hat{J}_i(risk)$	$J_i(eff)$	$\hat{J}_i(eff)$	Total cost, with $[\omega_1, \omega_2, \omega_3]$		
								[1,1,1]	[1,1,2]	[2,1,1]
1	$B^1, P^{3,1}, O^{7,0}, P^{3,0}, O^{2,3}, O^{6,3}$	7.10	1.00	1.00	0.33	2.90	0.00	1.33	**1.33**	2.33
2	$B^1, P^{3,1}, O^{7,0}, P^{3,0}, O^{6,3}, O^{2,3}$	7.10	1.00	2.00	0.67	2.90	0.00	1.67	1.67	2.67
3	$B^1, P^{3,0}, O^{2,3}, O^{6,3}, P^{3,1}, O^{7,0}, P_2^{3,0}$	7.07	0.10	1.00	0.33	4.30	1.00	1.43	2.34	1.33
4	$B^1, P^{3,0}, O^{2,3}, P^{3,1}, O^{7,0}, P_2^{3,0}, O^{6,3}$	7.09	0.55	0.00	0.00	3.60	0.50	1.05	1.50	1.50
5	$B^1, P^{3,0}, O^{6,3}, O^{2,3}, P^{3,1}, O^{7,0}, P_2^{3,0}$	7.07	0.10	2.00	0.67	4.30	1.00	1.76	2.67	1.67
6	$B^1, P^{3,0}, O^{6,3}, P^{3,1}, O^{7,0}, P_2^{3,0}, O^{2,3}$	7.09	0.51	3.00	1.00	3.60	0.50	2.01	2.50	2.50
7	$B^1, P^{3,0}, P^{3,1}, O^{7,0}, P_2^{3,0}, O^{2,3}, O^{6,3}$	7.10	0.93	1.00	0.33	2.90	0.00	1.27	**1.33**	2.33
8	$B^1, P^{3,0}, P^{3,1}, O^{7,0}, P_2^{3,0}, O^{6,3}, O^{2,3}$	7.10	0.93	2.00	0.67	2.90	0.00	1.60	1.67	2.67
9	$P^{3,0}, B^1, O^{2,3}, O^{6,3}, P^{3,1}, O^{7,0}, P_2^{3,0}$	7.07	0.00	1.00	0.33	4.30	1.00	1.33	2.34	1.33
10	$P^{3,0}, B^1, O^{2,3}, P^{3,1}, O^{7,0}, P_2^{3,0}, O^{6,3}$	7.09	0.47	0.00	0.00	3.60	0.50	0.97	1.50	1.50
11	$P^{3,0}, B^1, O^{6,3}, O^{2,3}, P^{3,1}, O^{7,0}, P_2^{3,0}$	7.07	0.00	2.00	0.67	4.30	1.00	1.67	2.67	1.67
12	$P^{3,0}, B^1, O^{6,3}, P^{3,1}, O^{7,0}, P_2^{3,0}, O^{2,3}$	7.09	0.47	3.00	1.00	3.60	0.50	1.97	2.50	2.50
13	$P^{3,0}, B^1, P^{3,1}, O^{7,0}, P_2^{3,0}, O^{2,3}, O^{6,3}$	7.10	0.93	1.00	0.33	2.90	0.00	1.27	**1.33**	2.33
14	$P^{3,0}, B^1, P^{3,1}, O^{7,0}, P_2^{3,0}, O^{6,3}, O^{2,3}$	7.10	0.93	2.00	0.67	2.90	0.00	1.60	1.66	2.67
15	$P^{3,0}, O^{2,3}, B^1, O^{6,3}, P^{3,1}, O^{7,0}, P_2^{3,0}$	7.07	0.00	1.00	0.33	4.20	0.93	1.26	2.19	1.26
16	$P^{3,0}, O^{2,3}, B^1, P^{3,1}, O^{7,0}, P_2^{3,0}, O^{6,3}$	7.09	0.47	0.00	0.00	3.50	0.43	**0.90**	1.36	1.43
17	$P^{3,0}, O^{2,3}, O^{6,3}, B^1, P^{3,1}, O^{7,0}, P_2^{3,0}$	7.07	0.00	1.00	0.33	4.10	0.86	1.19	2.05	**1.19**
18	$P^{3,0}, O^{6,3}, B^1, O^{2,3}, P^{3,1}, O^{7,0}, P_2^{3,0}$	7.07	0.00	2.00	0.67	4.20	0.93	1.60	2.52	1.60
19	$P^{3,0}, O^{6,3}, B^1, P^{3,1}, O^{7,0}, P_2^{3,0}, O^{2,3}$	7.09	0.47	3.00	1.00	3.50	0.43	1.90	2.36	2.43
20	$P^{3,0}, O^{6,3}, O^{2,3}, B^1, P^{3,1}, O^{7,0}, P_2^{3,0}$	7.07	0.00	2.00	0.67	4.10	0.86	1.52	2.38	1.52

and then evaluate the generated sequences. This script worked based on the set of all valid sequences of sub-goals for each task (V). For Task 1, this set contained 30 sequences resulting from the shuffle product of all permutations of $(O^{7,3}, O^{2,3})$ and $(P^{3,1}, B^5, P^{3,2}, B^1)$. For Task 2, 20 sequences included in Table 4.2 constituted set V. To generate a collection of n demonstrations, n random sequences in V were selected and used in one of the following forms:

- C: Without any modifications (correct sequence).
- M_k: k random sub-goals in the sequence were removed. This was to simulate cases when some sub-goals were missing (e.g., were not detected by *demonstrations observer*, were not provided by the teacher, etc).
- WO_k: The order of sub-goals in k random pairs in the sequence was switched. This was to simulate cases when some sub-goals were not detected in their correct place in the sequence of actions.
- WT_k: k random sub-goals in the sequence included incorrect target information to simulate possible inaccuracies in sub-goal detection. For example, sub-goal $P^{3,1}$ could be replaced with $P^{3,0}$, to simulate a situation when box 3 was incorrectly identified as moved to place 0 instead of 1. For example, $O^{2,3}$ could be replaced with $O^{2,0}$, simulating a case when object 2 was detected to be placed inside box 0 instead of 3.
- I_k: The last k sub-goals in the sequence were removed. This was to simulate incomplete demonstrations, either through technical errors or the teacher not demonstrating the task until the final goal was achieved. Note that such demonstrations were labelled as unsuccessful.

Evaluating sequences in programs and success rate

For nine different compositions of correct and faulty demonstrations, we carried out 100 tests each, with majority percentages α_E, α_D, α_R, α_B, α_N, and α_C used by the *program generator* set to $\{60\%, 70\%, 80\%, 90\%, 100\%\}$. The coefficients were kept equal to α so that the number of tests was reasonable. Each test was carried out with a simulated set of demonstrations and the sequences in the generated program were evaluated: a sequence was considered 'correct' if it was in V; if a sequence was only a sub-sequence of a member in V, it was classified as 'incomplete'; Otherwise, it was labelled as 'faulty.' Figure 4.5 presents the total numbers of generated sequences in each of the three categories, averaged

over 100 tests, using different majority percentage values and with or without the filtering strategy. The average ratio between the number of correct sequences and all the sequences is also shown as the 'success rate.' This metric indicates the likelihood of success if one of the available sequences is randomly selected. If no program was generated, the success rate was recorded as 0. Wilcoxon (Mann-Whitney) rank-sum tests for independent samples were used when reporting the improvements resulting from the filtering strategy. Below are the important results in each of the nine cases:

$C \times 5$: With 5 correct randomly selected demonstrations, and with the $\alpha = 100\%$, 19.43 (SD=6.34) correct sequences were available on average over 100 tests for Task 1. For Task 2, on average 14.09 (SD=4.11) correct sequences were found. In this case, no incomplete or faulty sequences were generated in any test. Figure 4.5 (A1,A2,A3,A4) indicate that $\alpha = 80\%$ and smaller reduced the number of correct sequences in both tasks. The filtering strategy was not expected to make any impact when $\alpha = 100\%$ (because no error was present in this case). Filtering significantly improved the system's performance when $\alpha = 60\%$, e.g., resulted in the discovery of 4.92 (SD = 2.65) correct sequences on average for Task 1 compared to 0.52 (SD = 0.50) without filtering ($W = 9975.5, p < .001$).

$C \times 10$: With 10 correct random demonstrations, 27.24 (SD = 3.52) and 18.04 (SD = 2.67) correct sequences were generated on average for Tasks 1 and 2, respectively, with $\alpha = 100\%$. These numbers were close to the total number of correct sequences to perform the two tasks (30 and 20, respectively). As seen in Figure 4.5, the filtering strategy could prevent the accuracy drop when $\alpha \leq 80\%$.

$C \times 8, M_1 \times 1, M_2 \times 1$: With 8 correct demonstrations, 1 with 1 missing sub-goal, and 1 with 2 missing sub-goals, only incomplete sequences were generated in all the cases when $\alpha = 100\%$. This was because the set of required sub-goals, E, by analyzing all the demonstrations was always incomplete. Setting $\alpha = 90\%$ could make the system discover, on average, 16.78 (SD = 9.72) and 14.47 (SD = 7.87) correct sequences for Tasks 1 and 2, respectively. Still, incomplete demonstrations existed, so the success rate was 81.00% (SD = 39.43) and 82.34% (SD = 36.58) for Tasks 1 and 2, respectively. With $\alpha = 80\%$, while fewer number of correct sequences were generated (11.12 (SD = 4.95) and 9.71 (SD = 3.98) for Tasks 1 and 2, respectively), the success rate for both tasks was 100% (SD = 0%). The performance drop observed when $\alpha \leq 80\%$ was prevented when the filtering strategy was used.

$C \times 8, WO_1 \times 1, WO_2 \times 1$: With 8 correct demonstrations, 1 with 1 miss-ordered pair of sub-goals, and 1 with 2 miss-ordered pair of sub-goals several faulty sequences were in the generated programs, when $\alpha = 100\%$. This was because wrong sequential regularities could be captured in relations R and B. Here, similar to the previous case, reducing α

could enhance the performance. For example, with $\alpha = 80\%$, 10.83 (SD = 4.85) and 8.43 (SD = 3.88) sequences were available for Tasks 1 and 2, respectively, which were all correct. In this case, applying filtering increased the number of both faulty and correct sequences when $\alpha = 70\%, 80\%, 90\%$. However, when $\alpha = 60\%$, in Task 1, filtering could significantly improve the success rate from 58% (SD = 49.60) to 98.73% (SD = 7.31), with 19.88 (SD = 6.69) correct sequences identified ($W = 2729.5, p < .001$). A similar effect was also observed for Task 2 ($W = 4510, p < .05$).

$C \times 8, WT_1 \times 1, WT_2 \times 1$: In case 1 demonstration in sets of 10 demonstrations included 1 sub-goal with a wrong target and another demonstration included 2 sub-goals with a wrong target, the system's behaviour was very similar to tests with missing sub-goals, especially when $\alpha = 100\%$. This is because, with large α values, sub-goals with wrong targets are most likely observed only in a minority of demonstrations and therefore are not included in E, similar to when the original sub-goal is missing. However, in this situation, there is a chance that a sub-goal recorded with a wrong target is identical to other correct sub-goals. The system's performance was not high with $\alpha = 90\%$, but setting $\alpha = 80\%$ could improve the performance. In case $\alpha = 80\%$, on average 11.72 (SD = 4.99) and 9.62 (SD = 4.76) correct sequences were generated for Tasks 1 and 2, respectively, while the success rate was 100% (SD = 0%) and 98% (SD = 14.07%) for the two task, respectively. The filtering strategy helped increase the number of correct sequences in case $\alpha \leq 70\%$ and increase the success rate when $\alpha = 60\%$ for both tasks.

$C \times 8, I_1 \times 2, I_2 \times 1$: Including 1 demonstration with 1 and another demonstration with 2 missing ending sub-goal(s) in the sets of demonstrations 10 demonstrations did not substantially impact program generation. This was because demonstrations marked as unsuccessful did not influence sets E and D, and only reduced the information available for the formation of other relations. When $\alpha = 100\%$, 25.81 (SD = 4.94) and 17.6 (SD = 2.62) correct sequences were found on average for Tasks 1 and 2, respectively, and the success rate was 100% (SD = 0%) for both. Similar to the cases with only correct demonstrations, the filtering strategy could prevent the accuracy drop when $\alpha \leq 80\%$.

$C \times 2, (M_1, WO_1, WT_1, I_1) \times 2, C \times 6, (M_1, WO_1, WT_1, I_1) \times 1, C \times 6, (M_2, WO_2, WT_2, I_2) \times 1$: Lastly, we tested three situations where a mix of previously discussed faulty demonstrations was among the provided demonstrations. In the most extreme case, only 2 correct demonstrations along with 8 faulty demonstrations (2 from each category) were given to the system. Without the filtering strategy, the system identified 3.02 (SD = 2.73) correct sequences for Task 1 with a success rate of 77% (SD = 42.30%) when α was set to 70%. When filtering was in place, 15.91 (SD = 6.87) correct sequences were generated using $\alpha = 60\%$, and the success rate was 89.99% (SD = 22.29%). The improvement in the success rate, however, was not found to be statistically significant ($W = 4730, p = .36$).

Similar patterns were also observed in Task 2. Two other cases, with 6 correct demonstrations and 4 demonstrations with 1 or 2 errors of all kinds, were also tested. In both conditions, the system's performance was in a range similar to previous cases.

Overall, in these tests, setting α to 60% along with employing the filtering strategy was a reasonable choice to handle various kinds of errors. Note, in many cases, the filtering strategy increased the number of both correct and faulty sequences. While the success rate was also higher with the filtering strategy, filtering may not be appropriate for cases where performing actions in an incorrect order, even with a low likelihood, is harmful. Still, for any robotic implementation of our system, the likelihood of observing different errors in sets of multiple demonstrations can be measured and compared to the simulated conditions (described here). The best value of α can then be selected according to the condition that is most similar to the real situation.

4.4.3 Task Execution

We describe three example scenarios in this section to demonstrate how robots can perform tasks in the real world using a program. We also illustrate how a robot may successfully achieve a task goal from varying initial task states with the presence of interruptions, i.e., unexpected changes in the task.

In the first scenario, the robot executed sequence $(P^{3,1}, O^{7,3}, O^{2,3}, B^5, P^{3,2}, B^1)$ in the program learned in Example 1 (Section 4.4.1) to complete Task 1 from an empty initial state. This program was estimated in Example 1 to result in the shortest possible path-length, i.e, fastest task completion time, when the initial task state was empty. As illustrated in Figure 4.6 (A), iCub started by moving the pot to the sink ($P^{3,1}$). Once done, it added potatoes to the pot ($O^{7,3}, O^{2,3}$). At this point, a human relocated the pot. As part of 'following' a sequence, the robot completed the chain of sub-goals that needed to be achieved before any new sub-goal. Here, iCub noticed $P^{3,1}$ was missing from the task state and because $(P^{3,1}, B^5) \in R$, it re-achieved the missing sub-goal: it moved the pot back to the sink ($P^{3,1}$) to be able to add water (B^5). iCub then completed the task by moving the pot containing water and potatoes to the stove-top ($P^{3,2}$) and turning the stove-top on (B^1).

In the second scenario, the robot used the same sequence in the same program to perform Task 1 when the initial task state was $\{B^5, O^{7,3}\}$ (see Figure 4.6 (B)). In this case, the robot skipped moving the pot to the sink ($P^{3,1}$) at the beginning as it noticed water was already in the pot (B^5 was in the task state), because $(B^5, P^{3,1}) \in N$. Sub-goal $O^{7,3}$ was also skipped as it was already in the task state. Here, the robot only added the remaining

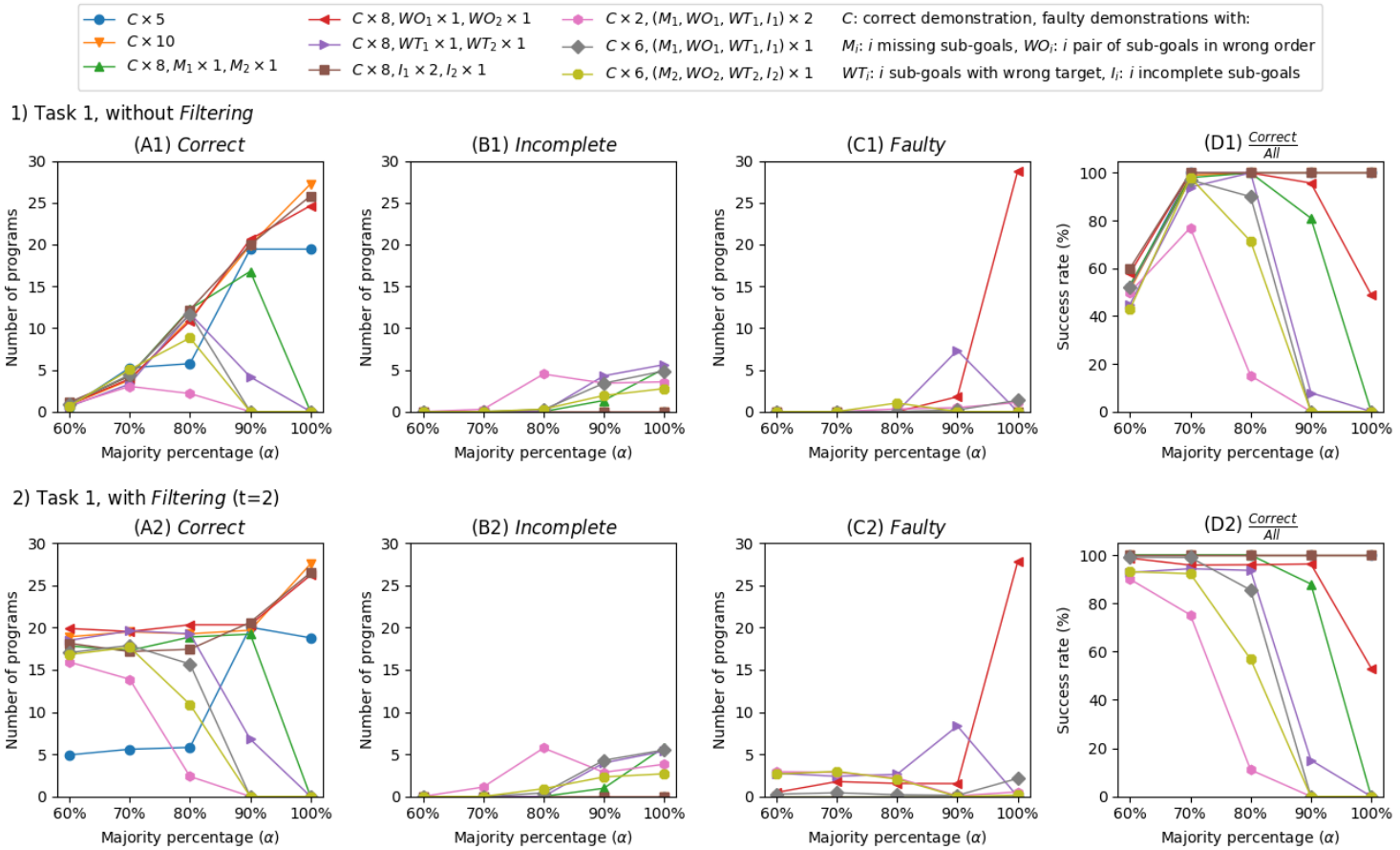


Figure 4.5: Performance evaluation of the PLI architecture in Task 1. In nine conditions, different sets containing random correct and faulty demonstrations were provided to the system in simulation. After each test, repeated 100 times, the generated sequences were classified as *correct*, *incomplete*, or *faulty*. (A,B,C) The average number of generated sequences, using different majority percentages (α) and with or without the *filtering* strategy. (D) Success rates: the percentage of correct sequences over all the available sequences (*Correct* + *Incomplete* + *Faulty*). © 2025 IEEE.

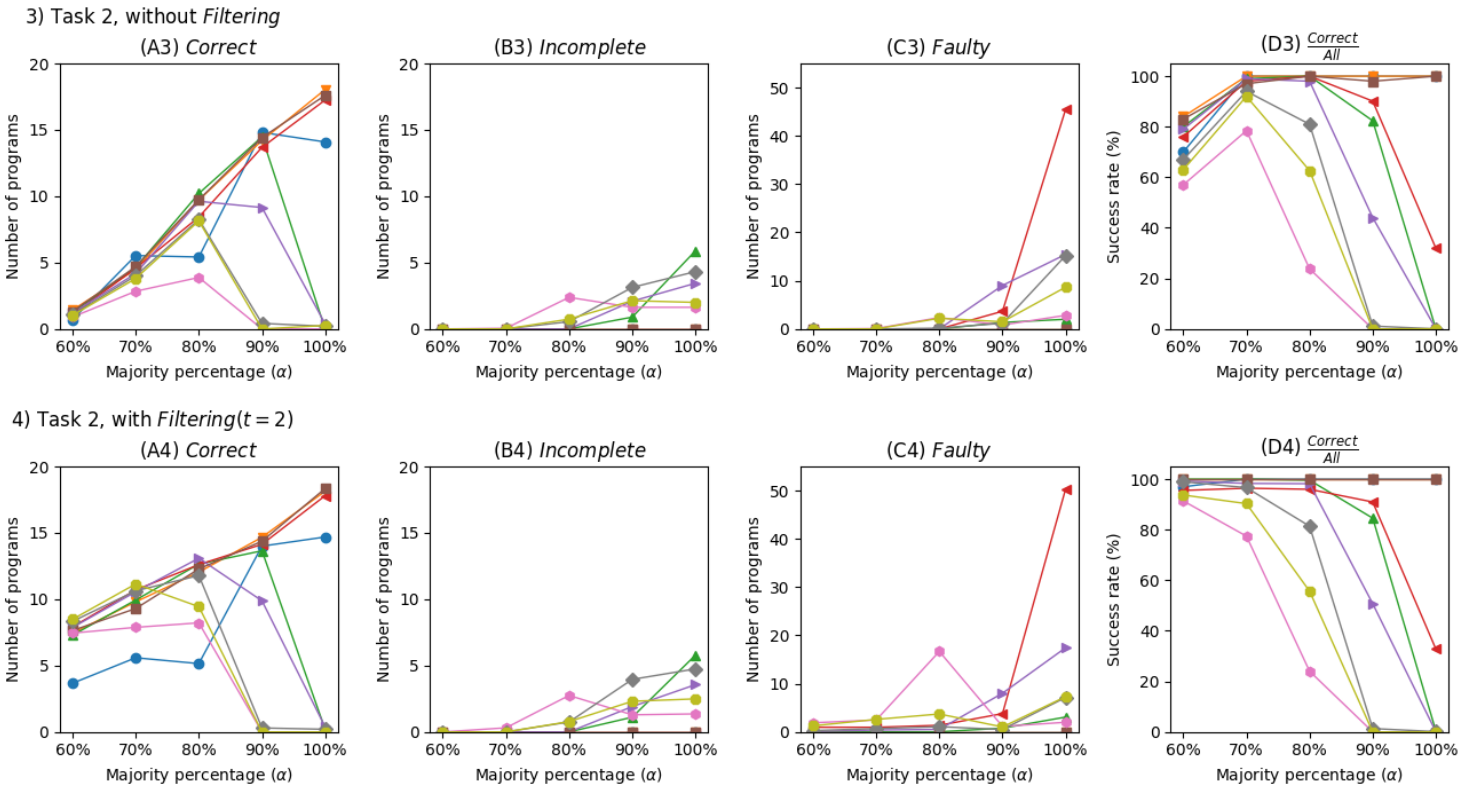


Figure 4.5: [Cont.] Performance evaluation of the PLI architecture in Task 2. In nine conditions, different sets containing random correct and faulty demonstrations were provided to the system in simulation. After each test, repeated 100 times, the generated sequences were classified as *correct*, *incomplete*, or *faulty*. (A,B,C) The average number of generated sequences, using different majority percentages (α) and with or without the *filtering* strategy. (D) Success rates: the percentage of correct sequences over all the available sequences (*Correct* + *Incomplete* + *Faulty*). Note that the vertical scales in (C3) and (C4) differ from (C1) and (C2). © 2025 IEEE.

potato to the pot ($O^{2,3}$), moved it to the stove-top ($P^{3,2}$) and turned the stove-top on (B^1) to complete the task.

In the third scenario, Figure 4.6 (C), the robot performed Task 2 using sequence ($P^{3,0}, B^1, O^{2,3}, P^{3,1}, O^{7,0}, P_2^{3,0}, O^{6,3}$) obtained in Example 3 in Section 4.4.1. Because the drawer was closed initially ($P^{3,0}$ was already achieved), iCub started by releasing the lock (B^1) and then putting the first meat package ($O^{2,3}$). Next, the robot opened the drawer ($P^{3,1}$) and placed the ice cream in the hidden space that was now accessible ($O^{7,0}$). Once done, iCub closed the drawer ($P_2^{3,0}$) and completed the task by putting the second meat package in the drawer ($O^{6,3}$).

4.5 Discussion

We introduced a novel imitation learning system for robots that enables them to understand the sequential structure of goal-directed actions. By learning from human demonstrations, the robot comes up with multiple possible sequences of sub-goals in the form of a program to perform tasks flexibly and to better adapt to contextual variations and different user preferences. To accomplish this, the robot determines which order of sub-goals in demonstrated sequences can be altered or left unspecified and which sub-goals are necessary for future sub-goals. We described how different modules in the proposed PLI framework work together, explained how such a system was implemented on an iCub robot, and evaluated the effectiveness of this approach in various cases.

As a part of system evaluation in Section 4.4.1, we presented examples of the learning and adaptation process of a robot to two physical abstract tasks. We explained how the system discovered more possibilities in achieving task goals and found more effective strategies for task performance, according to previously defined optimization functions and through mental simulation of the real tasks, when demonstrations that covered different situations were provided. For selecting a sequence, optimization functions to minimize the length of robot hand movement when performing the actions (to eventually reduce task completion time, Examples 1 and 2), to lower the risk of the robot failure when manipulating objects (Example 3), and to better arrange actions in a particular organization task to prevent ice cream from thawing when transferring frozen food to a freezer (Example 4) were employed. We also described a situation in which different users, each with unique expectations from the robot, customized robot sequence selection according to their own preferences. To use with our PLI system, robot designers can build many other optimization functions to give users a variety of options to customize robot behaviour for different scenarios and tasks. For example, a function may add to the total cost if there is a delay

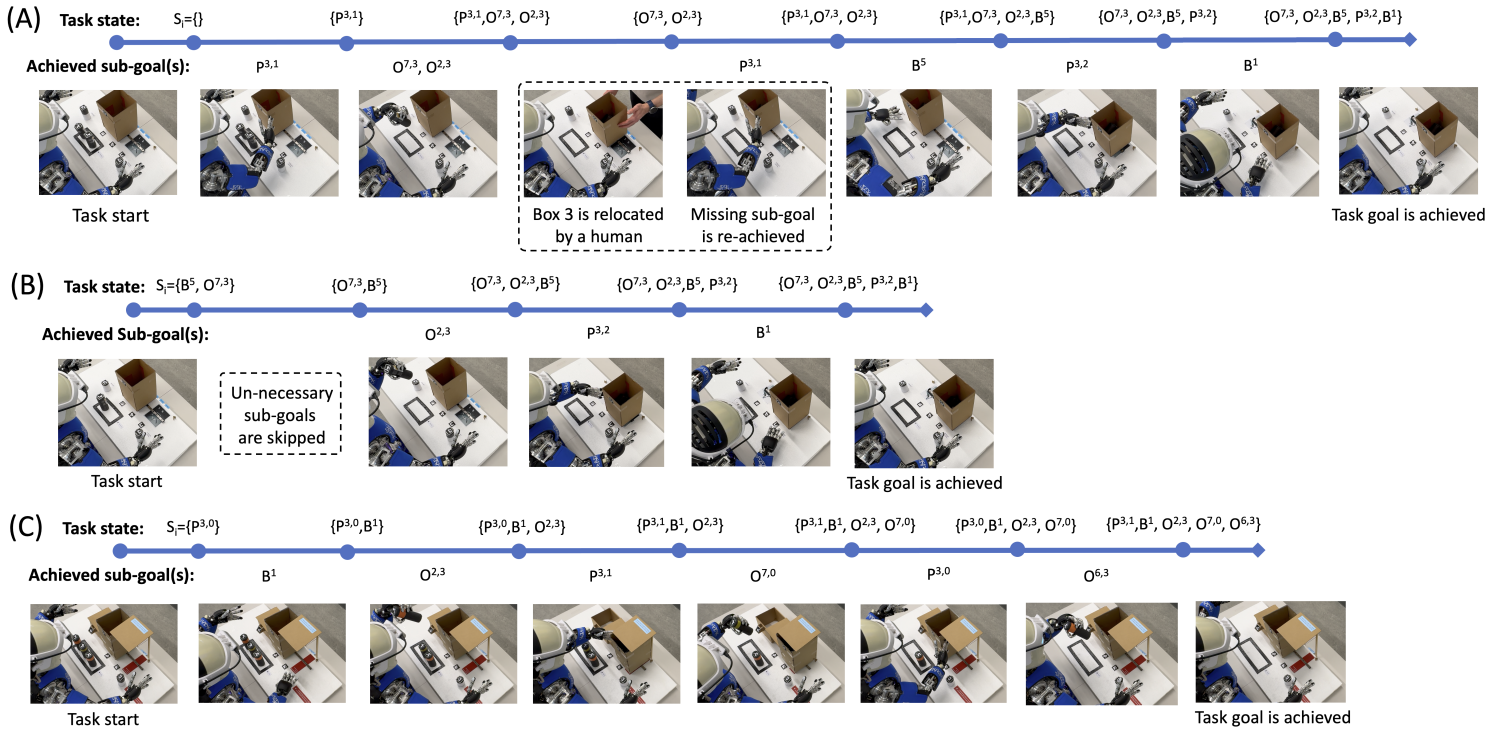


Figure 4.6: Timelines of three example task execution by the iCub robot: (A) performing the abstract potato boiling task with an interruption, (B) performing the abstract potato boiling task from when some task steps have been previously completed, and (C) performing the abstract organization task. Snapshots of the robot achieving sub-goals are shown. The task state in every step, achieved sub-goals, and important events are also noted. © 2025 IEEE.

between turning on the water and placing the vessel below it, or if an oven stays hot for some time before food is placed inside it, to reduce home energy consumption. Another function can be designed to increase the safety of the robot handling fragile objects, e.g., by adding to the total cost if eggs are dropped into the pot before water is added.

A robot equipped with the present PLI system can be envisioned to record task demonstrations in two ways: when it is being explicitly taught by a human teacher (as was the case in this chapter) or when it is moving around and observes humans performing tasks, e.g., around a house. In both cases, the learning progress of the robot may be impacted by the recorded demonstrations. Diversity in the demonstrations will help *program generator* observe different situations and therefore, obtain more information about sub-goals in a task to better discover possibilities in structuring the sub-goals. In Section 4.4.2, we observed that when demonstrations contained five 'random' ways for task execution, fewer than 20 sequences were found on average for Task 1. Meanwhile, when a researcher carefully provided effective, diverse demonstrations of the same task (Example 1 in Section 4.4.1), only four demonstrations were sufficient to obtain 30 sequences. Due to the importance of diversity in demonstrations, the full potential of the PLI system may not be realized in some situations, e.g., when only a few people teach a robot and they all tend to follow a specific general approach. Moreover, when a robot is observing humans performing tasks, it may take time for the robot to observe different approaches to performing one specific task. In real-life applications, it is not yet known whether or not a robot can generate programs with more sequences when receiving demonstrations from non-expert robot users compared to the theoretical situation of using randomly picked task demonstrations (although this aspect will be experimentally tested in Chapter 5). Nevertheless, our approach can benefit robots by offering an effective way of accommodating variability in demonstrations, which is a challenge in LfD [155].

As outlined in Section 4.2.2, *program generator* can be configured to group some sub-goals together in generated sequences. The 'grouping' strategy can reduce the search space for valid sequences and make programs easier to describe by forming 'families' of sequences. As humans, we may most of the time perform similar, repetitive actions altogether rather than holding some of the sub-goals in mind to come back and complete later. For example, when making a salad, we usually add all the tomatoes at once 'to clear our mind.' This has been shown to reduce cognitive and memory load in humans [198, 159]. For a robot, in contrast, considering the possibility of achieving some sub-goals in any place in sequences of actions only requires extra memory and calculations, which is not difficult for today's modern robots. The proposed PLI system could allow various places for sub-goals to complete a task, even if some were always demonstrated together. Letting the sub-goals separate allows for more flexibility, which may lead to improved efficiency in new situations.

Such a case was observed in Example 3 in Section 4.4.1: while sub-goals $O^{6,3}$ and $O^{2,3}$ were achieved one after another in all the demonstrations, they were not adjacent in all the generated sequences without the grouping strategy. In that particular case, only the sequences generated without the grouping strategy could yield zero cost based on the optimization function that preferred picking objects in a form that reduces the risk of robot failure.

In light of the above discussion, it would be better for a robot not to use grouping for sequence selection to enhance its performance, but to use grouping, when possible, for communicating programs to users. Grouping sub-goals according to the procedure noted in Section 4.2.2 will minimize the total number of groups. Nevertheless, more sophisticated strategies may be used to help the system decide whether or not to keep sub-goals in their existing groups. For example, pauses during task demonstrating after completing certain sub-goals have been shown to indicate the completion of end goals [189]. Such end goals may be considered to guide the PLI system to better group sub-goals.

Throughout the main part of system evaluation in Section 4.4.2, which focused on assessing the performance of the system by analyzing the generated programs, it was observed that by adjusting system parameters even in a limited manner and also employing the strategy of 'filtering,' the system could learn multiple correct sub-goal sequences even if demonstrations contained errors. In this work, seven simulated conditions of combining different errors in sets of 10 demonstrations were considered when searching for the best parameters. To obtain system parameters more accurately, a condition should be simulated that represents the average combination of errors that are expected to be in sets of demonstrations of different tasks in a real situation, i.e., when non-experts' demonstrations are recorded by the robot. Robot designers may then test different sets of values for majority percentages in simulation, to find the most suitable system parameters.

Despite the strategies in place for handling errors, the PLI system can greatly benefit if errors in sub-goal detection by the *demonstrations observer* are reduced. As explained in Section 4.3.2, *demonstrations observer* can be designed to use multiple modalities of recording demonstrations at the same time. Combining different modalities may help the robot better identify sub-goals in instances where a single modality fails to accurately detect sub-goals. In such cases, a system can be implemented such that if there is a conflict between learning modalities, the robot, e.g., verbally asks the teacher to repeat the last sub-goal. As an example, if by observational learning the robot detects box 1 is moved to place 0 ($P^{1,0}$) and by kinesthetic learning, the robot detects box 1 is moved to place 1 ($P^{1,1}$), it may ask the human teacher for clarification. For example, the robot can ask "please repeat the last step involving box 1 in the demonstration." Then, only if the teacher does not repeat that part of the demonstration (it may be impossible because, e.g., the

environment has changed inevitably) or the inconsistency persists, the sub-goal is recorded with missing information. Nevertheless, if there is a large difference in the accuracy of the two or more learning modalities, having them work together may not be desirable, as it could result in many instances of the robot asking questions when learning or may lead to many unknown sub-goals. In such cases, employing one accurate learning modality would yield better results.

With the PLI learning architecture, robots use their own primitives, and this will allow the transfer of programs from one robot to another. A new embodiment can use different methods for detecting sub-goals and acting to achieve them, which can be pre-programmed by its designers. The correspondence problem [140] is not relevant here: as long as the robot can achieve the needed sub-goals, even in different ways, a program learned by another robot can be still useful. Furthermore, by allowing learned programs to be utilized as robot primitives in the PLI architecture, robot teaching can take place in a modular form, which may improve user experience. For example, in the abstract potato boiling task discussed in this chapter, a teacher may first teach a smaller task of 'adding potatoes to the pot' and then, after making sure the small task is properly learned, include it in the demonstration of the bigger task.

The PLI system could consider learned programs as new robot primitives (i.e., sub-goals that the robot can detect or achieve) to allow learning new tasks in a hierarchical or recursive manner. For example, the entire previously-learned process of "boiling potatoes" may be recorded as a single sub-goal (noted by BP) in the demonstrations of "making potato salad." Whenever the robot needs to achieve sub-goal BP in the program of the bigger task, it proceeds down the hierarchy and runs the boiling potatoes program.

Since humans often project their own mental models onto robots that appear human-like [120], we expect that a biologically inspired learning framework that allows robots to learn in a way comparable to humans and primates can potentially help increase the acceptance of such robots. By making aspects of the learning process visible to human teachers, non-expert users may better understand a robot's learning progress, correct issues with robot learning, and teach robots more effectively. As an example, a human can verify the information captured in relation R, if the robot can explain or present such information to the teacher, to prevent robot mistakes. To do so, the robot may ask "I understood that when boiling potatoes, the pot needs to be inside the sink before I can add water, is this correct?"

4.6 Limitations

The presented PLI system utilizes a robot's pre-programmed capabilities for detecting effects associated with a set of known sub-goals and acting to achieve them. Implementing the PLI system on a new robot requires programming robot primitives in a way suitable for the new embodiment. In our work, an iCub robot was programmed to only detect the occurrence of three types of sub-goals in abstract tasks. Moreover, in a real-life context, decomposing tasks into sub-goals may be challenging [129]. In fact, tasks may consist of several sub-goals reached at the same time or in a continuous series of events. Human teachers may not always define sub-goals clearly. Furthermore, sub-goals may not always result in visible physical effects on the surroundings, e.g., 'waiting' for a certain amount of time to boil eggs or stirring a liquid for some time. Several studies in robot LfD have focused on 'learning' to extract sub-goals in human demonstrations, e.g., [127, 80, 213, 76]. Future work can incorporate more effective, automated, and generalizable techniques for sub-goal detection in the *demonstration observer* module.

Our proposed method is relatively simple in terms of how data is processed and decisions are made by the robot. While this minimalist approach can be particularly effective for learning tasks that consist of a small number of sub-goals, knowledge acquisition and performing tasks when many sub-goals are present may be computationally costly because of some large search spaces and high time complexity of the proposed algorithms, e.g., during greedy search in Algorithm 1 for program generation and during mental simulation of actions based on multiple sequences for task execution. There are also limitations when capturing statistical regularities used in programs. For instance, at least a few demonstrations that start from a reasonably small initial task state are required so that sets E and D (noted in Table 4.1) can be constructed, set N may not include information about repeated sub-goals, etc.

Another limitation of our approach is the simplistic strategy for sequence selection. At present, to find the optimal sequence for the robot's actions, *sequence selector* adds the costs associated with different aspects together and compares the total costs. Technical aspects of this work can be improved in future research. As an example, fuzzy logic may replace the simple technique for sequence selection [217] so that a user only defines certain 'degrees of importance,' rather than numerical weights, for each criterion, e.g., indicate that task completion time is of 'medium importance' and robot safety is 'absolutely important.'

Concerning the generalizability of the system, while the current implementation can learn tasks of different kinds, it does not allow for customization or substitution in the sub-goals, to generalize programs to cover other tasks that may have extra sub-goals compared

to what was originally learned. Furthermore, the current system does not allow changes or replacements in the program in real-time when performing a task. A more effective approach for *sequence selector* and *action planner* may allow dynamic sequence selection based on the task progress and current situation. The present work also discusses tuning of system parameters for learning new tasks in a limited setting. These generalizability limitations can be addressed with subsequent developments over time, which can be a challenge.

4.7 Future Directions

The present learning framework creates opportunities for realizing social learning of robots, i.e., learning through feedback and interaction with humans [33, 209]. One strategy for achieving this would be implementing a 'learning from conversation' modality as a part of *demonstration observer* to identify and record sub-goals in a task from verbal instructions. With such capabilities, a robot may even ask for clarifications from the human teacher in real-time during the conversation. In this area, LLMs may demonstrate great potential. Future work can employ LLMs to extract and record sub-goals from human-provided descriptions of tasks, possibly with the help of knowledge from other sources for widely known tasks, in a manner similar to *demonstration observer*. LLMs may also contribute to other aspects of our PLI architecture. For example, relations identified and used by the *program generator* may be evaluated using an LLM according to the definition and purpose of each item, so that robot the robot asks for clarification or correction if needed. As another example, LLMs may facilitate user input for *sequence selector* by mapping a verbal request from a user to a vector containing weights for available cost functions.

The PLI approach opens up new possibilities for explainable learning systems. Here, generating programs is led by a limited number of sets and relations, each with a logical definition as noted in Table 4.1. Selecting a sequence is also driven by costs obtained via optimization functions that are defined for our physical world. Acting based on a sequence is also governed by relations that define why a particular sub-goal is being achieved, as described in Algorithm 2. Future work can focus on developing a component that can express the system's behaviour in any step of learning and performing tasks, to support explainability. LLM may be useful in this context as well, by enabling the robot to explain its internal state, intents, goals, etc., through a natural conversation. Here is an example scenario. In case a user asks a robot why a particular order of sub-goals was not considered, the robot can explain and point to the optimization function based on which the sequence was selected. If the questioned sequence was not part of the valid sequences, the robot can

explain the reason, i.e., which relations or sets were in conflict with the sequence. If the user believes the questioned sequence is indeed valid, the robot can explain the contradictory relations/sets and also indicate according to which demonstrations such relations/sets are known, so that the user can decide to delete or correct the demonstrations, if they are faulty.

4.8 Summary and Conclusion

In Phase 2 of my research, I designed a minimal, biologically inspired PLI system for humanoid robots inspired by learning in primates. With this architecture, a robot can learn from people with different styles in doing a particular task. After acquiring some level of structural knowledge about a given task, the robot can take into account a new user's preferences and select a suitable sequence of actions among the multiple possible solutions for performing that task in a new setting. Having access to programs with multiple actions grouped into sets allows the robot to perform better by knowing several possibilities of ordering the actions to achieve a goal, resulting from receiving different demonstrations, which may lead to better performance. The proposed system can work with any modality for demonstrating to a robot, help a robot cope with interruptions and errors, which are very likely to occur in a domestic scenario [59], and improve efficiency in a manner that a user may expect. This work contributes to developing robots that can learn from and meet the needs of a diverse group of users.

I also implemented the proposed learning architecture on an iCub humanoid robot and evaluated the effectiveness of the system in multiple scenarios, when either a researcher took the role of a teacher for the robot and provided demonstrations, or demonstrations were simulated computationally. This work addressed O2 most notably with the design and implementation of a robotic learning framework. Addressing O2, however, requires evaluation of the learning system in a broader context. In the next chapter, I report on an experiment that was conducted to further evaluate the system by involving a group of non-expert users teaching physical tasks to the robot.

Chapter 5

Experiment 2: Exploring Humans' Awareness of a Robot's Limitations and Capabilities as an Indicator of Suitable Teachers for the Robot

After designing a robot learning framework, implementing it on an iCub robot, and testing it in a controlled setting, we now transition to utilizing the system with non-expert humans as teachers for a robot. In this chapter, I report Experiment 2, which was conducted to evaluate human-robot teaching interaction with the proposed PLI learning system, and more importantly, to identify factors in human teachers that could contribute to better robot learning outcomes in human-robot teaching, with an ultimate goal of answering the question of 'who to imitate.'

I propose a factor in human teachers for robots, namely, human teachers' awareness of and attention to the robot's capabilities and constraints, and explore how this factor affects robot learning outcomes. To define 'teachers' awareness of and attention to a robot's limitations and capabilities' (in the short form: *teachers' awareness*), I took into account the inherent limitations faced by robots that learn by visually observing the environment as well as the constraints imposed by the learning mechanism that a particular robot uses. More specifically, I considered that (1) such a robot could only learn from demonstrations visible to its vision system, and (2) in case using the proposed PLI for task learning, such a robot could only learn from certain primitive actions in human demonstrations (as noted in Section 4.2). Accordingly, a teacher with high awareness when teaching a robot is expected

to (1) ensure objects and presented actions in the demonstrations are visible to the robot, and (2) demonstrate primitive actions in a way that the robot can learn effectively (e.g., avoid performing unknown actions, combining task steps, etc.).

Through an experiment, I systematically explore the role of human teachers' awareness of a robot's limitations and capabilities in improving robot learning outcomes. To manipulate teachers' awareness and potentially increase this factor in certain participants, some groups of human teachers for robots were provided with extra resources to help them understand robot limitations and learning strategies. This experiment studies whether teaching the robot with an emphasis on attention leads to improved robot learning outcomes, primarily to address O3. In this experiment, 72 participants, who had no expertise in robotics, taught three physical tasks to an iCub humanoid robot by providing multiple demonstrations. Depending on the experimental condition, the provided resources were: (1) a screen next to the robot that displayed what the robot could see (to give a visual perspective of the robot to the teacher), and/or (2) virtual activities that put participants in the position of the robot when labelling task actions in some hypothetical demonstrations (to let them experience the strategies a robot with such a learning mechanism uses to learn, i.e., identifying primitive actions in demonstrations). Throughout this chapter, I examine the impact of experimental conditions on various qualitative and quantitative measures of robots' learning performance (e.g., quality of recorded demonstrations, program generation scores, etc.) to address O3. Moreover, I study trends in the characteristics of non-experts' demonstrations to contribute to addressing O1 and assess the effectiveness of the PLI system when used by non-expert users to contribute to addressing O2. I also examine the data to find patterns in non-experts' teaching that a robot can use to identify the most suitable teachers to improve its learning outcomes, contributing to O4.

This chapter outlines the investigated research questions in this phase (Section 5.1), details the methodology when designing and conducting the experiment as well as the measures used to study the research questions (Section 5.2), presents the results (Section 5.3), discusses how the results contribute to answering the research questions (Section 5.4), explains the limitations of the experiment and the employed methodology (Section 5.5), and proposes future research directions to address these limitations (Section 5.6). Research findings are summarized in Section 5.7 together with some concluding remarks.

5.1 Research Questions

The experiment presented in this chapter addresses the following research questions related to the main research objective of the thesis (O1, O2, O3, and O4):

RQ1 In observational learning of a robot from non-expert humans, does giving humans extra resources for understanding robot limitations and learning strategies (i.e., by showing them what the robot can see and/or putting them in the position of the robot when labelling task actions)

- **RQ1.1** increase their **attention to ensuring objects and presented actions in the demonstrations are visible to the robot**?
- **RQ1.2** increase their **attention to demonstrating primitive actions so that the robot can learn effectively**?
- **RQ1.3** improve the **quality of the recorded demonstrations**?
- **RQ1.4** improve the **robot learning outcomes**?
- **RQ1.5** impact their **perceptions of the robot's social attributes**?
- **RQ1.6** impact their **perceptions of teaching workload**?

RQ2 Is observational learning with a Program-Level Imitation (PLI) framework effective when non-expert participants teach a robot?

RQ3 Over multiple rounds of non-experts teaching a robot and watching the robot perform what it has learned through observational learning, how do the characteristics of the provided demonstrations change?

RQ4 Does data from non-experts teaching a robot indicate any pattern to find better teachers for observation learning of the robot?

Our study is exploratory in nature due to its specific focus on how humans teach robots. We expected that showing participants what the robot could see would improve their attention to the visibility of the demonstrations, and putting participants in the position of the robot when labelling task actions would improve their attention to demonstrating primitive actions so that the robot can learn effectively. It was also expected that an increase in both aspects of participants' attention would enhance the quality of the recorded demonstrations and improve robot learning outcomes.

RQ1 is directly related to O3 as it focuses on understanding how providing human teachers for a robot with extra resources for understanding robot limitations and learning strategies when they teach a robot may impact different aspects of robot learning. RQ2 is related to O2 by evaluating our PLI learning framework with non-expert participants. Answering RQ3 contributes to O1 as it focuses on understanding how some aspects of

visual task demonstrations provided by non-expert humans change over time when they teach multiple tasks to a humanoid robot. RQ4 asks for patterns in non-experts' teaching that a robot can use to identify the most suitable teachers for its observational learning, and is therefore related to O4.

5.2 Methodology

To study the research questions, we conducted a 2×2 between-participants experiment (experimental factors are described in Section 5.2.1) involving non-expert participants teaching three physical tasks (explained in Section 5.2.2) to an iCub humanoid robot [137]. The robot was visually observing task demonstrations using the camera on top of its head. The experimental procedure will be outlined in detail in Section 5.2.4. In summary, demonstrating each task with an abstract model located on a table in front of the robot (see Figure 5.1) consisted of two phases for all the participants:

1. *Placing* 'objects' involved in the task in a designated area
2. *Teaching* the task to the robot by carrying out task actions

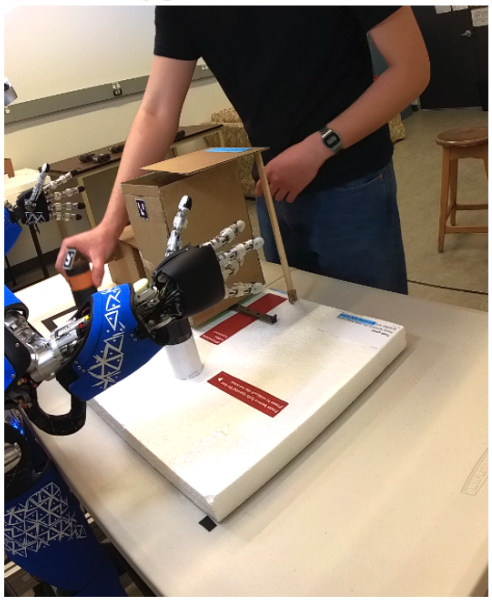
After receiving five demonstrations, the robot performed the task based on its own learning. The robot used the PLI learning mechanism introduced in Chapter 4. Details on how the learning system was adjusted for the experiment are presented in Sections 5.2.3. Measures and metrics used to quantitatively and qualitatively study the research questions are also outlined in Section 5.2.5.

5.2.1 Resources Provided to Participants

The variables manipulated in this experiment, summarized in Table 5.1, were the presence or absence of two extra resources in the teaching session to help human interaction partners better understand the limitations of the learner robot and its learning strategies. These resources were:

- **Having a screen next to the robot that displayed the Robot's Camera View (RCV) in real-time throughout the session:** Participants in the *RCV* condition could see the robot's vision system output throughout the experiment on

(A) *Placing* phase



(B) *Teaching* phase

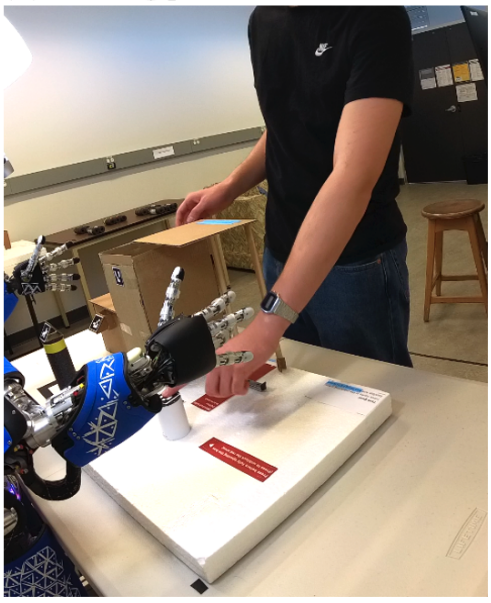


Figure 5.1: A participant in the experiment demonstrating a task to the iCub robot in two phases: (A) *placing* task objects in the designated space, which acted as introducing objects involved in the tasks to the robot; and (B) *teaching* the robot by physically performing the task actions.

a 12" screen, as illustrated in Figure 5.2. For participants in *No-RCV* condition, the screen was facing away from them. The robot's camera view output was mirrored so that participants, who were observing the task from the opposite side, could better relate the robot's view of the tasks with their own view. Green boxes appeared around markers that were attached to different components and places in the tasks when they were detected by the robot.

The presence of this screen provided participants in certain groups with the opportunity of VPT, as discussed in Section 2.3 and used in many other previous works (e.g., [208, 220]). In this work, VPT is used in the specific context of human-robot teaching. Offering the RCV was intended to increase participants' awareness of the visual limitations of the robot (e.g., the robot may not see some task components if they are obstructed by participants' hands).

- **Completing a virtual Goal-Highlighting Activity (GHA) before the first time demonstrating each task to the robot:** A virtual activity was completed

by participants in *GHA* condition before demonstrating each of the three experimental tasks to the iCub robot. The *GHA*s put participants in the position of a hypothetical robot with a learning mechanism similar to the robot in the experiment when the robot needed to recognize and label observed task steps. A tablet was handed to participants to complete each GHA on multiple pages. There were three variants of GHA presented to participants in the session, with scenarios about a future advanced robot learning how to make coffee, start a washing machine, and declutter a kitchen space; and in an order that was counterbalanced with a Latin square design. On the starting page of each GHA, participants saw a written description of a scenario indicating the following (shown in Figure 5.3 (A)): "*A future advanced robot observes a person [making coffee] in the setting pictured below. This robot can only learn from watching someone perform these five basic actions: (1) Pouring, (2) Pressing a button, (3) Mixing with a spoon, (4) Moving an object, and (5) Opening/closing a drawer. Please help the robot recognize the steps in the task on the next pages.*" A picture of the initial state of the task in a real setting, from the hypothetical robot's perspective, was also included. Then, in the five following pages, participants observed snapshots of the task when the person was performing each step and answered "*What is the action being presented to the robot at this time?*", by selecting one option among the five basic actions that were listed earlier, as shown in Figure 5.3 (B). Section C.1 in Appendix C presents snapshots from all pages in all variations of GHA used in our experiment.

By completing the GHAs three times during a session, participants in certain groups experienced solving a problem similar to the one faced by the robot in the experiment, i.e., mapping the observed actions to one of the known primitive actions. Offering these activities was intended to increase participants' awareness of the limitations of this robot in only being able to effectively learn only from watching primitive actions highlighted in demonstrations (e.g., not from watching two primitive actions being combined, or sub-goals being achieved using actions different from the programmed primitive actions).

5.2.2 Abstract Physical Tasks

We built physical models of three simple household tasks, in an abstract form, using very lightweight material (i.e., foam and cardboard) to be taught to the robot. The robot would not have been able to safely fetch and move real household objects such as potatoes and pots. Tasks 1 and 2 were previously used in the work presented in Chapter 4. Written

Table 5.1: The experimental factors and their levels for each of the four groups. These factors were the presence or absence of two extra resources in the teaching session to help human teachers better understand the limitations and capabilities of the learner robot and its learning strategies.

		Robot's Camera View	
		No Robot's Camera View (*No-RCV*)	Having a screen showing the Robot's Camera View (*RCV*)
Goal-Highlighting Activity	No Goal-Highlighting Activity (*No-GHA*)	Group 1	Group 2
	Completing Goal-Highlighting Activities (*GHA*)	Group 3	Group 4

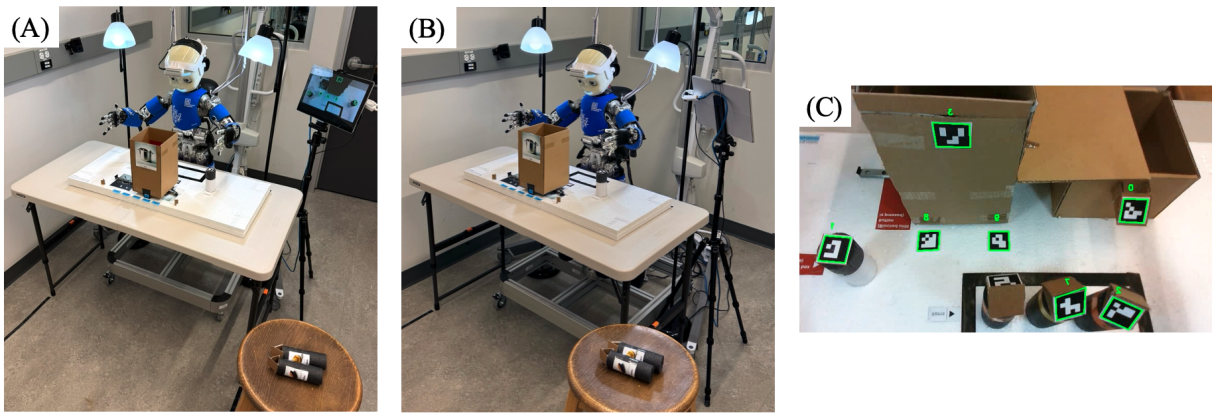


Figure 5.2: (A,B) The experimental setup for task learning of the iCub robot in *RCV* and *No-RCV* conditions, respectively. The screen on the right side of the robot displayed the robot's mirrored camera view to participants in *RCV* condition, while in *No-RCV* condition, this display faced away from participants. An abstract model of a task (Task 1) is set up on a table in front of the robot, with task objects initially on a stool (bottom right corner of photos). (C) The iCub's camera view that was offered to participants in RCV condition, with detected markers highlighted with green boxes.

descriptions and labels were attached to components of each task (e.g., buttons, sliding and stationery boxes, and objects) to denote their assumed roles in the experiment. Task

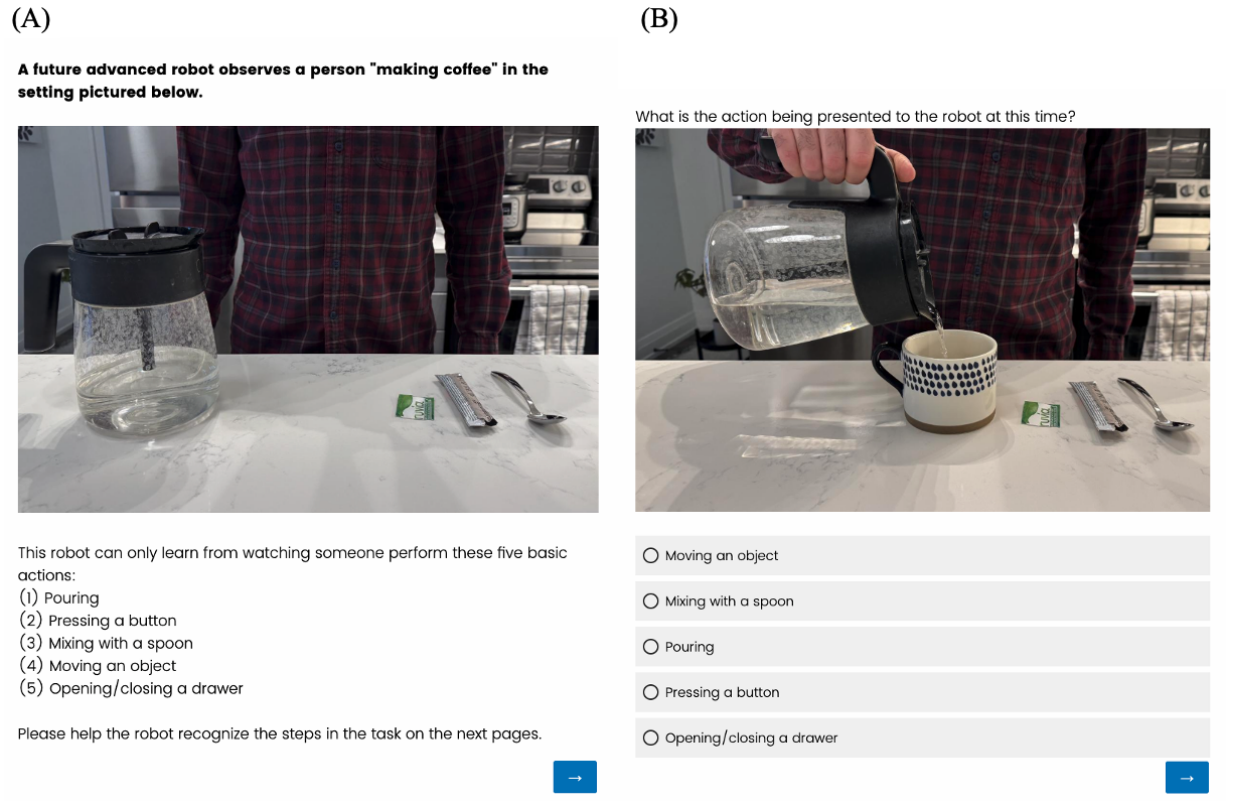


Figure 5.3: Snapshots from a Goal-Highlighting Activity (GHA) around the scenario of a robot learning how to make coffee. (A) The initial page described the scenario and listed the primitives of the 'future advanced robot' when learning tasks. (B) On the following five pages, participants were asked to help the robot by recognizing and selecting the primitive related to the demonstrated task step. Only one option could be selected. Zoom in for more visible text.

components and places involved in the tasks had Aruco markers attached to them to be detected and tracked by the robot's vision system. Here is a summary of each task:

- **Task 1** modelled a simple kitchen environment (Figure 5.4 (A)) for boiling two 'potatoes.' A box ('pot') could be moved into three different places: in a 'sink,' on a 'stove,' and on a 'countertop.' In between these three places were labelled as 'unknown,' since the robot was not able to detect any sub-goals. The stove could be turned on by pressing a button, and 'water' could be added by pressing another button when the pot was in the sink. Two task objects were assumed to be two different potatoes. The task goal was to place the pot filled with water containing the two potatoes on the hot stove.

- **Task 2** modelled a 'drawer' in a constrained space (Figure 5.4 (B)). A tall orange box (the box on the left in the figure) had a cover for an adjacent short yellow box, attached to its side, and could be moved into places marked as 'fully open' or 'fully closed,' corresponding to the state of the yellow box. The orange box could be closed by another cover when placed at 'fully open' (i.e., if moved further to the left from the setting shown in the figure). The constraint indicated by the red label (visible at the top left of the task in the figure) required pressing a button before moving the orange box to the 'fully open' place. Another label on the top cover indicated that objects needed to be placed in boxes when the orange box was fully closed or fully open. Two orange and one yellow objects were also part of the task. The goal of the task was to organize the three task objects in the matching coloured boxes and move the orange box to the 'fully open' place.

- **Task 3** modelled a waste-sorting scenario and consisted of three boxes ('bins') labelled with 'food waste,' 'plastic,' and 'paper' for placing task objects 'banana peel,' 'empty water bottle,' and 'newspaper' in them, respectively (Figure 5.5). A fourth task object, 'battery,' was meant to remain on the table because it did not belong to any of the three bins. A button had to be pressed after moving objects into the bins to close all bins.

Objects in all tasks were in the form of cylinders about 12 cm tall. As shown in Figure 5.6 (A), there was a triangular prism attached to the top of each object that was used for demonstrating tasks to the robot, with the marker positioned on only one face. By doing so, the robot would only detect the marker if the face displaying it was facing the robot. For example, in the robot's vision system output presented in Figure 5.2 (C), among the three task objects placed in the designated area (bottom right of the output),

(A)



(B)



Figure 5.4: (A) Tasks 1 and (B) Task 2 used in the experiment, from participants' point of view. Task 1 modelled a simple kitchen environment, including a pot, sink, and stove, with the goal of starting to boil two potatoes. Task 2 modelled a drawer in a constraint space, including two boxes and a button to remove a constraint, with the goal of placing objects in the matching colour box/bin. The task goal was indicated at the bottom left of each task. Zoom in for more visible text.



Figure 5.5: Task 3 used in the experiment, from participants' point of view. This task consisted of models of four waste items and three bins, with the goal of placing the three relevant items into the appropriate bins and closing all the bins at the end by pressing a button.

the robot detected the two objects on the right. The left-most object was not detected since the marker was not fully visible to the robot. If a participant paid attention to what the robot could see, they had to rotate the object until the robot could 'see' the marker. Note, to avoid providing a hint to participants about the markers on objects, we used another set of objects shown in Figure 5.6 (B) with markers on top (which were visible from any direction) for the task performance of the robot.

For each task, a set of correct sequences was defined by the researchers. These sets for Tasks 1 and 2 were similar to what was used in Chapter 4, except that in Task 1, we allowed turning on the stove at any time, and therefore, 120 sequences of 6 sub-goals were considered correct. For Task 2, 18 sequences with 7 sub-goals and 2 sequences with 6 sub-goals were in this set. For Task 3, we considered all $3! = 6$ permutations of moving 3 waste items into their right bin as correct sequences.

5.2.3 Robot Learning and Performance Mechanism

For learning a task, the robot watched the five demonstrations provided by each participant and obtained 3D locations of task components, using an Intel RealSense camera [106] that

(A)  (B) 

Figure 5.6: Example of abstract task objects used for (A) providing demonstrations to the robot and (B) robot practising the tasks. The markers were used by the robot to identify each object. Objects used for providing demonstrations could be detected by the robot only if the face with the marker was facing the robot.

was mounted on its head. These locations were then converted to values in the iCub robot's base coordinate system [3]. As noted before, the robot in the experiment was using the Program-Level Imitation (PLI) architecture introduced in Chapter 4. Using the following strategies, the robot processed the locations of task components and recorded the achievement of any of the listed known sub-goals related to the task components. Note, compared to those detailed in Section 4.3.2, the sub-goal detection strategies used by the iCub robot in the current experiment were slightly modified to increase the accuracy, especially for having non-experts teaching the robot.

- **Putting object X in box Y:** The robot checked if any of the conditions A or B was met[10]: A) If the x-component of the location of object X was bigger than the same value for any box while the object's z-component was smaller than a threshold (to indicate that it may be in a box), object X was considered to be inside the closest box. B) If object X became invisible for 3 seconds, object X was considered to be inside box Y, in case box Y was the closest box to where object X was last located, and this distance was shorter than a threshold[11].

[10]To balance the experiment, for about half the participants, condition A was checked before condition B and for the rest, condition B was checked before condition A.

[11]For running the experiment with non-experts who did not know the underlying learning mechanism and to reduce sub-goal detection errors, when learning Task 2, if the orange box was detected to be either

- **Pressing button X**: The robot checked if the z-component of button X was smaller than the average of this value in the past 2 seconds (indicating that the button has been moving downward for some time) and the condition no longer holds (meaning the button is still) to consider button X is pressed.
- **Sliding box X to place Y**: Movable box X was detected to be moved to place Y if its y-component was within 2 cm of the y-component of place Y for 1 second.

All boxes, buttons, and places in each task had already been registered in the learning system to be part of the task. However, to make teaching closer to a realistic situation, objects in each task were introduced to the learning system by the participant placing them in front of the robot. Accordingly, only objects that were visible to the robot after the *placing* phase were considered by the robot as part of the task in the *teaching* phase.

Once all five demonstrations were recorded by the robot, a program possibly containing multiple sequences of sub-goals to complete the task was generated (as detailed in Section 4.2.2). Based on our past evaluations of the learning system (e.g., high success rate observed between 70%-80% in Figures 4.5 and 4.5), we set the majority percentages equal to 80% without *filtering* or *grouping* strategies. This means that relationships between sub-goals were identified based on 4 out of 5 demonstrations, and the valid sequences consisted of sub-goals observed in 4 out of 5 demonstrations.

For the robot performing a task itself, task objects were placed by the experimenter in such an order that at least one 'safe' sequence existed in which the robot had to pick objects from left to right (the experimenter checked the sequences in the generated program before setting up the task). In our implementation, because the robot reached for objects with its left arm, the risk of collision with another object was minimized when following such sequences (i.e., $\hat{J}_i(risk) = 0$ as defined in Section 4.3.4). The robot then evaluated all the safe sequences using the optimization function estimating the path-length of the robot's hand (i.e., $path()$), to select a sequence with the shortest task completion time, and performed the task using the optimal sequence.

5.2.4 Experimental Procedure

Step 0 - Pre-session module: Before the session, participants completed an online module to give consent, confirm that they were not robotics or human-robot interaction

'fully open' or 'fully closed' after an object disappeared for 3 seconds, the object was assumed to be inside the box that was not covered.

researchers, and answer a few demographic questions and questions about their previous experience with robots (snapshots are provided in Section C.2 in Appendix C). To prevent making participants self-conscious of their behaviour, there was no indication given in the study information sheet that the study was to understand the impact of their awareness and attention.

Step 1 - Introduction: When participants entered the experimental room, they were asked to stand in front of the robot in a marked area (see Figure 5.2 (A)). The first task was already set up in front of the robot. Participants were told that they needed to teach three simple household tasks to the robot by performing them in front of it. Next, participants were told the following three facts about the robot:

- "*This robot can only learn from watching someone perform these three basic actions: (1) Moving an object from the table to a box, (2) Sliding a box to a different place, (3) Pressing a button.*" The experimenter pointed to an example from each type of task component (i.e., button, box, etc.) when mentioning them.
- "*The robot observes the task and locates the markers on the objects using the camera on its head.*" In *RCV* condition, the explanation continued with: "*When the robot detects a marker, it will become green on the screen.*" The experimenter also pointed to a green box on the screen to show an example.
- "*You will need to teach each task five times. While you can repeat the same instruction in multiple rounds, the robot benefits from seeing variations in how you put task actions together.*"

Information about the *placing* and *teaching* phases in every round was then given to participants, and they were asked to teach according to any method they felt was correct.

Step 2 - Completing a goal-highlighting activity [GHA condition only]: Before teaching the first task, participants in Groups 3 and 4 were asked to complete a GHA on a tablet, at their own pace. Once they were done, the experimenter noted that "*Similar to the robot in the activity you just completed, the robot in our experiment can only learn from watching someone perform a specific set of basic actions. These actions are: ...*" and reminded participants of the three basic actions the iCub robot could learn from.

The experimenter then left the experimental room and observed the study through one-way mirrors to issue commands to the robot from an adjacent room, e.g., to start/stop the learning and task performance of the robot.

Step 3 - Participants giving demonstrations: Each participant provided five demonstrations of each task to the robot. In every demonstration round, task objects were initially placed by the experimenter horizontally and randomly on a stool next to the task, as shown in Figure 5.2 (A,B). Additional photos from the experimental setup and material are provided in Section C.3 in Appendix C. The robot first gave a summary of the experimental procedure (which was given to participants in detail by the experimenter in Step 1) and then asked participants to "*place task objects in front of me.*" After participants confirmed that the objects were placed in front of the robot (end of the *placing* phase) by saying 'ready,' the robot said "*You can start teaching me now,*" and then, participants could start performing task actions (start of the *teaching* phase). After finishing the demonstration, participants said 'done' to end the recording of the demonstration by the robot. The experimenter then came back to the experimental room to reset the task and also put objects back on the stool for the next round. To avoid technical difficulties, i.e., speech recognition errors, and to ensure the same interaction experience for each participant, the experimenter listened to 'ready' and 'done' signals and started and stopped learning manually. The robot detected and saved the identified sub-goals, along with time (individually and also in the form of a demonstration to be used in the PLI system, as detailed in Section 4.2.1), task components that were not visible to the robot after the teaching phase in every frame (as will be explained and used later in Section 5.2.5), task objects that were not visible after the *placing* phase, and the video output of its vision system.

Step 4 - Robot trying the task: After five demonstrations were provided, the experimenter reset the task, placed the set of objects for the robot (as noted in Section 5.2.3), and asked participants to sit and watch the robot 'giving the task a try.'

After Steps 2-4 (or Steps 2 and 3 for participants in Groups 1 and 2) for the first task, the experimenter switched the task, and the same steps were repeated for the two remaining tasks.

Step 5 - Ending questionnaire: Participants answered an ending questionnaire on a tablet (snapshots are included in Section C.2 in Appendix C). The first part asked all participants questions Q1-Q3 listed in Table 5.2 to reflect on their attention to different aspects of teaching the robot. Participants in Groups 2 and 4 also answered questions Q4-7 (see Table 5.2) to self-report how often they checked the screen displaying RCV and how useful they found it. The above questions were asked on semantic differential scales using sliders, recording values ranging from 0 to 100. These values were not visible to participants (only the labels on each side were visible). Participants in Groups 3 and 4 answered Q8 to comment on the usefulness of the completed GHAs. All participants were then asked to rate how the robot learned the tasks (Q9), comment on their experience

when teaching the robot (Q10), as well questions to evaluate the social attributes of the robot (RoSAS questionnaire [146] using 7-point Likert scales) and the workload of the task of teaching the robot (NASA-TLX questionnaire [85] using semantic differential scales with sliders).

5.2.5 Measures and Metrics

The following subjective and objective measures and metrics were employed in this experiment to study research questions:

RQ1.1, RQ1.2: Participants' quantitative self-reflections when answering Q1-7 in the ending questionnaire, recorded on a scale ranging from 0 to 100 (with 100 being the highest consideration/frequency/usefulness as noted in Table 5.2), were analyzed to assess different aspects of their attention. I classified participants' comments in response to open-ended parts of questions Q6-7 as positive, negative, or indeterminate.

RQ1.3: To assess the quality of the recorded demonstrations to address RQ3.3, we defined and analyzed metrics that quantify various types of issues in the demonstrations, as well as the overall variation among them. In particular, we considered:

- **Performance issues in demonstrations:** I watched the video of each demonstration that was recorded by the robot's camera to identify and classify issues with 'how the task was demonstrated by participants to the robot' (classified and discussed in Section 5.3.3). I also logged the sequence of sub-goals the robot was expected to record from the performed task demonstration (referred to as *annotated sequence of sub-goals*) and compared it with the closest correct way of completing each task noted in Section 5.2.2, to identify the impact of each type of performance issue (also discussed in Section 5.3.3). The impact of performance issues on demonstrations could be:
  - **Extra wrong sub-goal:** When the robot was expected to detect an extra sub-goal affecting the correctness of the demonstration.
  - **Extra sub-goal:** When the robot was expected to detect an extra sub-goal not affecting the correctness of the demonstration.
  - **Missing sub-goal:** When the robot was expected not to detect a sub-goal that was necessary for the correctness of the demonstration.
  - **Wrong target:** When the robot was expected to detect an incorrect target box when an object was moved to a box in the demonstration.

Table 5.2: Items in the ending questionnaire answered by participants. Only participants in groups with access to the RCV answered Q4-7. Only participants who completed GHAs answered Q8.

	Question	Response type
Q1	When placing task objects in front of the robot, did you pay attention to placing them in such a way that objects were visible to the robot?	Slider (It did not come to my mind/ I considered it a lot)
Q2	When teaching the robot, did you pay attention to teaching tasks in such a way that objects were visible to the robot?	
Q3	When teaching the robot, did you pay attention to performing specific actions our robot could learn from, and no extra actions? (These actions were: putting an object inside a box, sliding a box to a different location, and pressing a button.)	
Q4	How often did you check the robot's camera view when placing task objects in front of the robot?	Slider (Not at all/Very often)
Q5	How often did you check the robot's camera view when showing tasks to the robot?	
Q6	How useful or not did you find the robot's camera view when placing task objects in front of the robot?	Slider (Not useful at all/ Very useful) + open-ended
Q7	How useful or not did you find the robot's camera view when showing tasks to the robot?	
Q8	When showing tasks to our robot, how useful or not did you find the activities (on the tablet before each teaching task) of helping a robot recognize the steps in the task performed by a human?	Open-ended
Q9	How well or badly do you think the robot learned the tasks?	Slider (Very badly/Very well) + open-ended
Q10	Overall, how was teaching the robot? Do you have any thoughts/suggestions on this process? Did you find anything particularly difficult?	Open-ended
	RoSAS questionnaire [146]	7-point Likert
	NASA-TLX questionnaire [85]	Sliders

- **Wrong order:** When the robot was expected to detect a sub-goal in an incorrect position in the sequence of sub-goals in the demonstration.

- **Sub-goal detection issues in demonstrations:** A script compared the *annotated sequences of sub-goals* with the sequence of sub-goals that was detected by the PLI system (referred to as *robot-detected sequences of sub-goals*)[12], and counted the occurrence of the following errors in each demonstration:

  - **Extra sub-goal:** When a robot-detected sub-goal was not in the annotated sequence of sub-goals.
  - **Missing sub-goal:** When an annotated sub-goal was not in the robot-detected sequence of sub-goals.
  - **Wrong target:** When the robot detected a wrong box, compared with the annotated sub-goal, when an object was moved to a box.
  - **Wrong order:** When a sub-goal in the robot-detected sequences of sub-goals was in an incorrect place, compared to the annotated sequence of sub-goals.

  As annotated sequences of sub-goals varied in length, we divided the total number of sub-goal detection issues by the total number of sub-goals in the demonstration. This fraction, multiplied by 100, was then called *normalized sub-goal detection error rate* for each task demonstration. This percentage does not necessarily indicate the percentage of sub-goals that were detected with an error, since there could be more than one detection issue per sub-goal. I also checked the videos of demonstrations when each sub-goal detection issue occurred and identified the causes for the issues (discussed in Section 5.3.3).

- **Variability of demonstrations:** I first reviewed each set of five *annotated sequences of sub-goals* each participant provided for each task 'to make them comparable.' In this process, modified sets of sub-goals were formed that only included major task steps, after extra sub-goals were removed, missing sub-goals were added, and some sub-goal issues were eliminated[13], so that all five sequences consisted of the same set of sub-goals. Considering each modified set, a script calculated the minimum number of pair-wise changes that were needed to convert each sequence to every

[12]The annotated sequences of sub-goals could contain performance issues, and the robot was expected to detect them the same way.

[13]For example, in demonstrations of Task 1, if the annotated demonstrations included moving the battery to a box, this sub-goal was removed since it did not contribute to the task goal.

other sequence (i.e., Kendall tau distance) in the set. The average of the 20 values was considered as the *demonstration variability score* for each set of demonstrations each participant provided for each task.

RQ1.4: With the PLI learning system, the robot generated a program potentially containing multiple sequences of sub-goals to achieve the goal of each task, based on the five demonstrations each participant provided. I considered the following measures and metrics to assess the robot learning outcomes:

- **Number of sequences in each program**
- **Program generation score**: For each task, a script compared the sequences in the robot-generated program with correct sequences (defined in Section 5.2.2) to identify any issues and to assign a score to each sequence, based on the following classes for the sequences:
  - **Valid minimal sequence:** In case a robot-generated sequence was among the correct sequences, a score of 1 was assigned.
  - **Incomplete sequence (sequence with missing sub-goals):** In case a robot-generated sequence S_r was a proper subset of a correct sequence S_v, a score of $|S_r|/|S_v|$ was assigned.
  - **Less-efficient sequence (sequence with extra sub-goals):** In case a robot-generated sequence S_r was the same as a correct sequence S_v but had extra sub-goals, a score of $|S_v|/|S_r|$ was assigned to it. In Task 3, a sequence with an extra sub-goal was considered to be wrong with a score of 0[14].
  - **Wrong sequence:** If a robot-generated sequence did not fall under any of the previous cases, it was considered to be wrong and a score of 0 was assigned to it.

  To quantify the success of program generation, the sum of scores assigned to all sequences in the program generated for each task using the demonstrations each participant provided was calculated as a *program generation score*, which is reported and analyzed in Section 5.3.4.

[14]This was because in Task 3, the only extra sub-goal could be moving the battery to one of the bins, which resulted in a wrong sequence, rather than a less efficient sequence.

- **Subjective robot learning performance evaluations**: We also assessed participants' quantitative subjective evaluation of robot learning performance as indicated by their answers to Q9, recorded on a scale ranging from 0 to 100 (with 100 being the most positive evaluation, as noted in Table 5.2). Note that this evaluation is limited as each participant only saw one performance of the robot for each task, which may also have included robot errors in object manipulation, object detection, etc.

RQ1.5, RQ1.6: Answers to each sub-scale in the RoSAS questionnaire, on scales from 1 to 7, were averaged to determine perceived competence, warmth, and discomfort of the robot. Participants' answers to six sub-scales of mental demand, physical demand, temporal demand, performance, effort, and frustration in the NASA-TLX questionnaire, on scales from 0 to 100 (with 100 being the highest workload/performance), were directly included in the analysis. This approach was also used in the experiment reported in Chapter 3.

RQ2: To evaluate the effectiveness of the PLI framework when non-experts taught some tasks to the iCub robot, we assessed participants' subjective viewpoints on robot learning:

- **Subjective robot learning performance evaluations**: We reviewed participants' responses to the open-ended part of Q9 (see Table 5.2), evaluating whether robot learning resulted in positive or negative outcomes from the teachers' standpoint.
- **Subjective evaluations of the teaching process**: We conducted a Thematic Analysis (TA) on the comments provided by participants reflecting on their experience of teaching the robot when answering Q10, because participants provided detailed answers to this question. We used the 6-step method developed by Braun and Clarke (2006) [32], which involved identifying codes and then merging them into themes.

To study RQ2, we also qualitatively evaluated the effectiveness of the PLI by further analyzing the sequences in the generated programs, which were classified when addressing RQ1.4. As will be detailed in Section 5.3.7, we also performed simulated tests comparing the learning outcomes when receiving demonstrations from individual participants to when learning from five randomly observed demonstrations.

RQ3: We considered the following characteristics of the demonstrations as they could be calculated by the robot itself and potentially be used in finding optimal teachers:

- **Time spent on providing demonstrations:** Similar to the experiment presented in Chapter 3, we analyzed the total time participants spent on providing demonstrations of each task.

- **Average time spent on demonstrating each sub-goal:** In our experiment, participants may have achieved the goal of a particular task with varying numbers of sub-goals. We therefore analyzed the time spent on providing demonstrations divided by the number of sub-goals achieved for each task demonstration, which estimated the average time spent on each individual sub-goal.

- **Fraction of frames where task components or objects disappeared**: In the *learning* phase, the PLI system evaluated the state of tasks every 200 milliseconds (i.e., 5 frames of the robot's vision system output per second) and recorded if any of the components defined as part of the task were not visible in each frame. We defined a metric called *normalized task components disappearance rate* for each demonstration by first dividing the number of frames in the demonstration when each object was not visible to the robot by the total number of frames in the demonstration (normalized for time) and, then, dividing this number by the total number of task components that were part of the task, i.e., excluding objects that were not visible to the robot after the *placing* phase (normalized for number of components within the view). With a similar approach, we also defined *normalized task objects disappearance rate* considering only 'objects' in the task as opposed to all the components in the task[15].

5.2.6 Participants

72 participants, 18 in each group, who were University of Waterloo students aged between 18 and 40, took part in our experiments. Participants confirmed that they were not a robotics or HRI researcher to be eligible to participate in our study. Figure 5.7 illustrates the age distribution of participants in each group. Moreover, Figure 5.8 presents participants' gender distribution as well as the number of participants in each group who had programmed a robot before, had interacted with a human-like robot before, had interacted with an iCub robot, or had taught something to a robot in another study. Note that participants who stated that they had programmed a robot before mostly stated that their experience dates back to high school projects or competitions. Participants who had interacted with a humanoid robot, including iCub, or even had taught something to a

[15]This metric was calculated only using data recorded by the robot itself, assuming the robot knows the number of important components (or objects) in the task before the start of teaching.

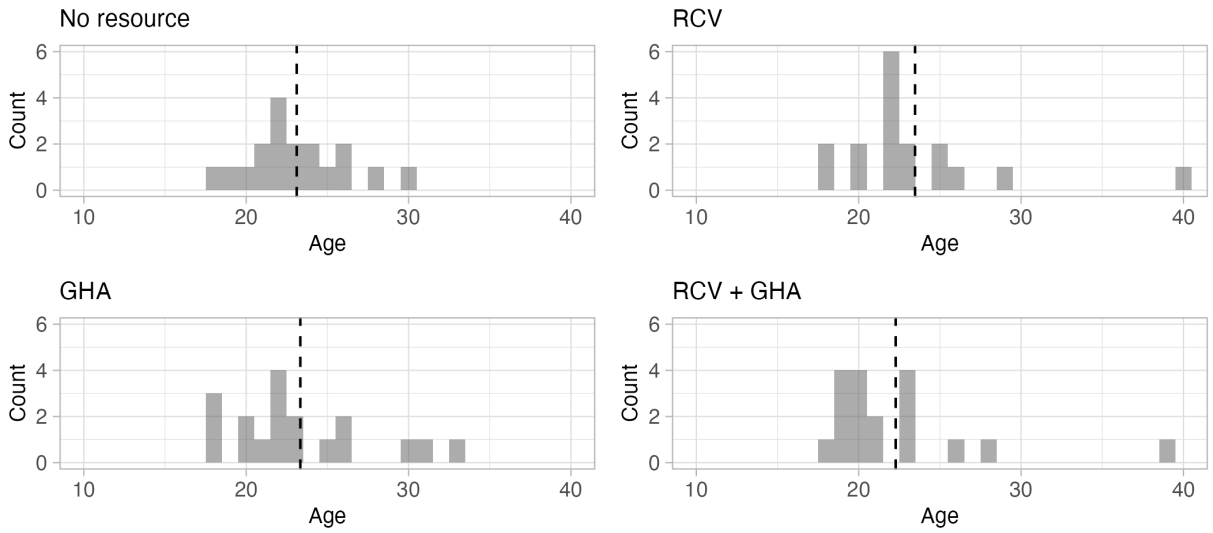


Figure 5.7: Histograms of participants' age across the four experimental groups. Dashed lines indicate the averages.

robot, likely had only taken part in other HRI studies on campus. The study received full ethics clearance from the University of Waterloo Human Research Ethics Board (included in Appendix A).

5.2.7 Statistical Analysis

In our analysis, we mainly used LMMs [24] to check for significant effects of the two experimental factors, as well as interaction effects between them, while also considering other possible confounding factors. For the analysis with regards to RQ1-3, participants' gender and age, and in some cases (for RQ1.3, RQ1.4, RQ3), also participants' self-reported attention to the visibility of objects in two demonstration phases (answers to Q1-2) and to demonstrating primitive actions (answers to Q3), the order of teaching the tasks (noted as *task order*) and providing demonstrations of each task (noted as *demonstration round*), and the nature of the task were controlled for and were kept in final models only if they could improve model fit, minimizing AIC [29]. A random intercept was fit based on the participant. We visually checked if the assumptions of LMMs were met, namely that residuals were independent and normally distributed, using histograms and Q-Q plots for the residuals and scatter plots of residuals against the fitted values. If not, we log-transformed our data (the only case was the analysis of program generation scores when

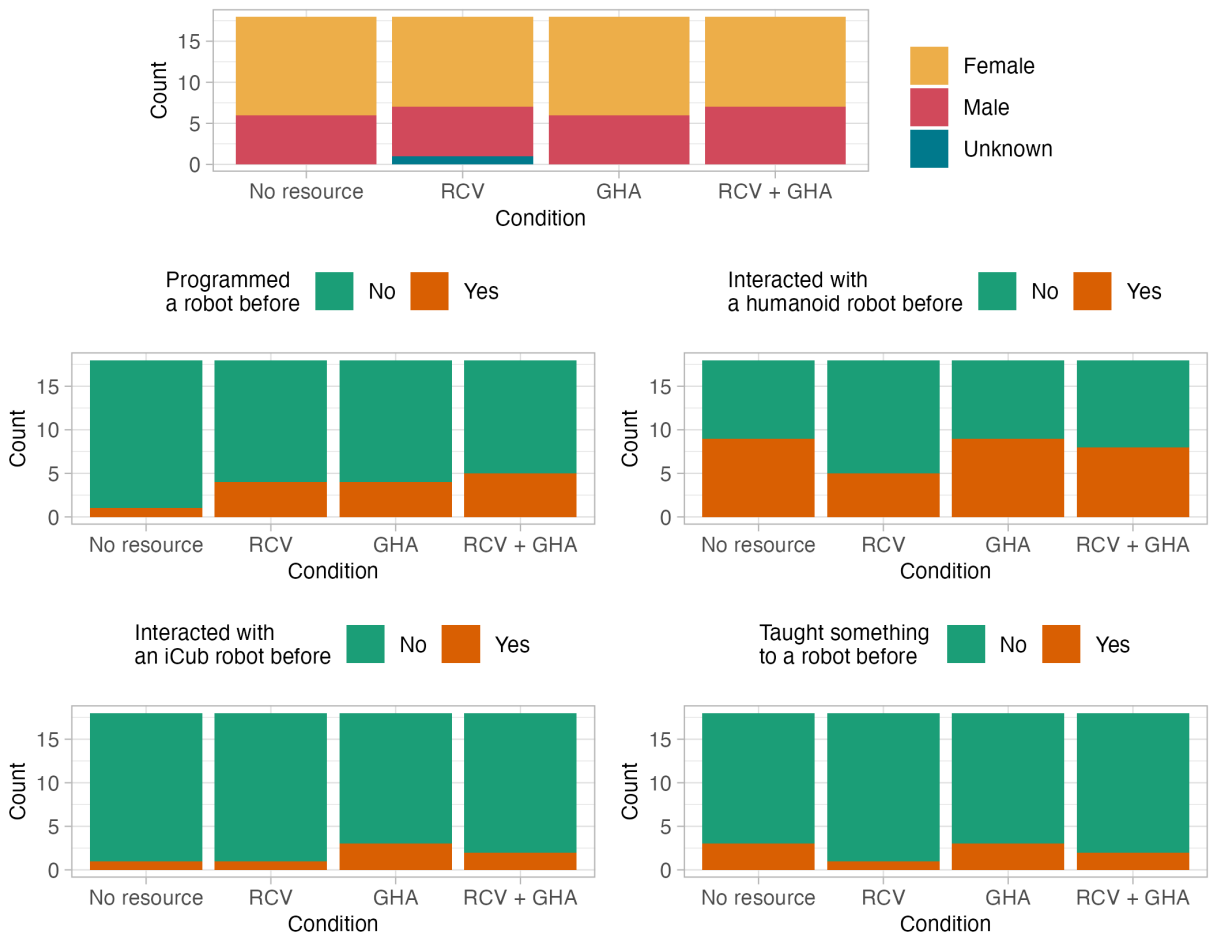


Figure 5.8: Participants' gender distribution and level of experience with robots across the four experimental groups.

studying RQ1.4). Wilcoxon (Mann-Whitney) rank-sum tests for independent samples were also used for comparing differences between specific levels of our experimental factor or between different tasks since our data was not normal according to the performed Shapiro tests. The p-values from the pairwise tests were adjusted for multiple hypothesis tests using the Holm-Bonferroni method. Note that data from 4 out of 216 task demonstrations were not included in the analysis due to technical issues.

5.3 Results

5.3.1 [RQ1.1] Attention to the Visibility of Demonstrations

Participants' average self-reported attention to objects being visible to the robot in the two phases of *placing* and *teaching* (their responses to Q1 and Q2) is presented in Figure 5.9 across the four experimental groups. According to the LMM presented in Table 5.3, also visible in Figure 5.9, participants' self-reported attention was significantly higher when they had access to the RCV compared to when they did not have that resource ($se = 3.74, t = 2.670, p < .01$). Participants who completed the GHAs reported significantly lower overall attention to the visibility of objects compared to those who did not complete the activities ($se = 3.74, t = -2.335, p < .05$). Moreover, participants generally reported higher attention to the visibility of objects in the *placing* phase compared to in the *teaching* phase ($se = 2.92, t = 3.808, p < .001$). In the *placing* phase, the attention difference between participants with and without access to the RCV was only approaching significance according to a Wilcoxon test ($W = 488, p < .1$). In the *teaching* phase, however, this difference was statistically significant ($W = 393.5, p < .01$). The attention difference between participants who did and did not complete the GHAs was significant both in the *placing* phase ($W = 807, p < .05$) and the *teaching* phase ($W = 835.5, p < .05$).

For participants who had access to the RCV, Figure 5.10 (A) presents the histogram of their self-reported frequency of checking the tablet displaying the RCV and perceived usefulness of the RCV in both *placing* and *teaching* phases (responses to Q4-7). While the majority of participants reported that they checked the RCV very often and found it very useful in the *placing* phase, participants' responses for the *teaching* phase varied considerably. According to Wilcoxon tests, the average participants' self-reported frequency of checking the RCV was significantly lower when *teaching* the robot, compared to when *placing* task objects ($W = -953, p < .001$). The same difference between *placing* phase and *teaching* phase was also observed regarding the perceived usefulness of the RCV ($W = -918.5, p < .001$).

Table 5.3: Linear Mixed-effects Model predicting participants' self-reported attention to objects being visible to the robot in task demonstrations

Covariate	**Estimate**	**SE**	**t**	**p**
Robot's Camera View *RCV*	9.981	3.739	2.670	**
Goal-Highlighting Activity *GHA*	-8.731	3.739	-2.335	*
Phase *Teaching*	-11.099	2.915	-3.808	***

No-RCV and *No-GHA* conditions and *Placing* phase are baselines. * = $p < .05$; ** = $p < .01$; *** = $p < .001$.

As illustrated in Figure 5.10 (B), no participant commented negatively on the usefulness of the robot's camera view in the *placing* phase, with 30 positive comments recorded. The following is an example of a positive comment: *"I noticed that the camera view had green boxes when an object was recognized, and I was trying to ensure that all codes were being green before operating. The camera view was helpful to provide me the robot's perspective for more accurate object placing etc."* Considering the *teaching* phase, comments from 11 participants were classified as negative. These participants often cited the difficulty of checking the RCV in the *teaching* phase due to their focus on completing the task, e.g., participants mentioned that *"I didn't check the robot's view because I was focused on doing the task,"* or *"I was more focused on completing the task correctly as to not make mistakes and confuse the robot."* Nevertheless, 19 participants indicated that having access to the RCV was useful in the *teaching* phase.

5.3.2 [RQ1.2] Attention to Demonstrating Primitive Actions

Participants' average self-reported attention to performing specific actions our robot could learn from (their responses to Q3) is presented in Figure 5.11 (A). We found no significant effect of any experimental factor or confounding variable on this measure when fitting LMMs. The effect of the RCV on increasing participants' attention was found to be only approaching significance ($se = 5.48, t = 1.693, p < .1$).

As shown in Figure 5.11 (B), participants commented equally positively and negatively on the usefulness of the completed goal-highlighting activities. 12 participants provided comments that were classified as negative, such as *"I was not sure how that activity was connected to the robot. I just thought they are independent. Didn't put that much focus on connecting the activity to the robot,"* and *"I didn't really find a lot of similarities between the activities on the tablet and the task itself. I think more related activities can be more*

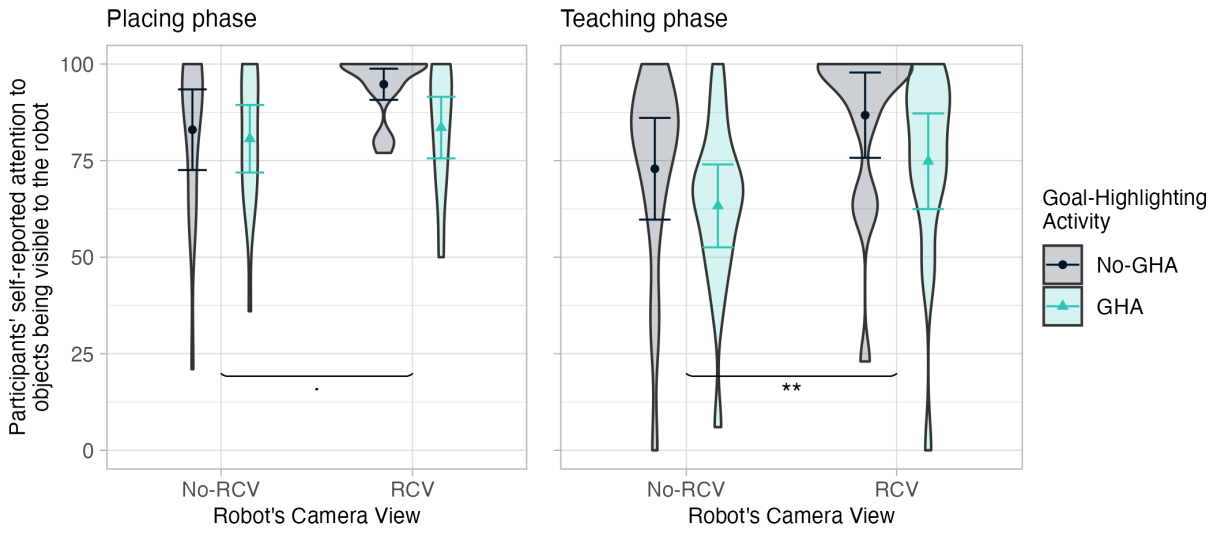


Figure 5.9: Average participants' self-reported attention to objects being visible to the robot in two phases of *placing* and *teaching* in task demonstrations. Error bars indicate 95% confidence intervals. The distribution of responses in each condition is also shown. . = $p < .1$; ** = $p < .01$.

useful." These comments highlighted the difficulty of finding a relationship between the virtual GHAs and the tasks of teaching the iCub robot. Examples from the 13 positive comments include: *"They helped break down the similar thought process of the learning robot," "I found it helpful as it gave me an idea of how the robot's observation might work and how to divide one task into smaller steps to make it more comprehensive for the robot,"* and *"I found it useful, made me think to ensure I perform a certain way."*

5.3.3 [RQ1.3] Quality of Recorded Demonstrations

As noted in Section 5.2.5, we utilized three metrics to assess the quality of the recorded demonstrations: performance issues when participants provided the demonstrations, sub-goal detection issues when the robot recorded the demonstrations, and the variability among the demonstrations.

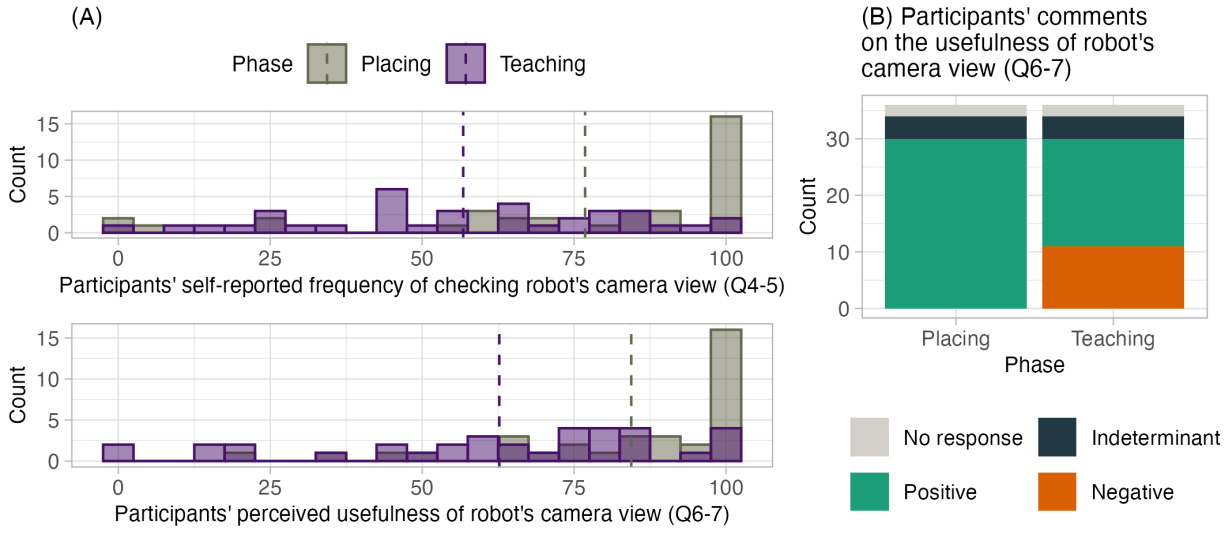


Figure 5.10: (A) Histogram of participants' self-reported frequency of checking the robot's camera view and its perceived usefulness in two phases of *teaching* and *placing* in task demonstrations. Dashed lines indicate averages. (B) Participants' comments on the usefulness of the RCV, classified as positive, negative, and indeterminant.

Performance issues in demonstrations

By watching the recorded videos of the RCV showing the provided demonstrations, we identified 12 types of performance issues, summarized in Table 5.4 (analysis of the occurrence of each issue is presented later in this section). These performance issues when providing demonstrations were determined to be from three main sources: (1) being unaware of the limitations of the implemented PLI system (Types #1-4), (2) teaching errors (Types #5-10), or (3) inefficiency in performed task actions (Types #11-12).

1) Being unaware of the limitations of the implemented PLI system: In some cases, performance issues when providing demonstrations occurred due to not considering the limitations of the PLI system in detecting sub-goals. These cases were when participants performed actions leading to correct sub-goals in the physical world, but they performed them in a way that the robot was not able to correctly detect the achievement of sub-goals. Grabbing and moving two or more objects at the same time (with both hands) was spotted in some demonstrations of all tasks (Type #1, noted in Table 5.4). When the sub-goals of putting objects in boxes were achieved in such a way in the demonstrations, the robot could not have been able to detect the order of each sub-goal as intended by the

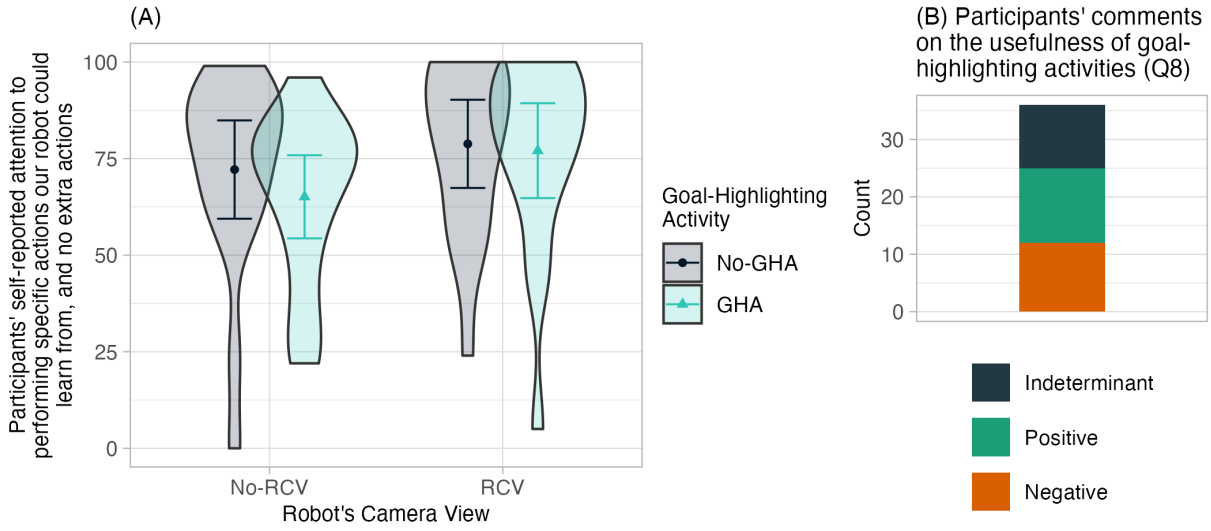


Figure 5.11: (A) Average participants' self-reported attention to performing specific actions the robot could learn from in task demonstrations across four groups. Error bars indicate 95% confidence intervals. The distribution of responses in each condition is also shown. (B) Participants' comments on the usefulness of the GHAs, classified as positive, negative, and indeterminant.

human teacher. Note, due to the nature of tasks used in the present work, Type #1 issues occurred with objects that did not necessarily need to be moved to boxes in a particular order, and therefore, such issues had no impact on the correctness of demonstrations. In some demonstrations of Task 3, participants moved a waste item to an arbitrary box initially and then transferred the item to the correct box (Type #2), or moved the battery to a bin and then took it out and placed it on the table again (Type #3). While these actions produced correct sub-goals at the end, the system could only detect the first box where an object was moved to, and therefore, the ways of achieving sub-goals could have led to the detection of sub-goals with wrong targets or extra wrong sub-goals, respectively, for Types #2 and #3. When demonstrating Tasks 1 and 2, some participants first picked an object and then moved a box while holding the object with their hand (Type #4). In such cases, it was possible that the system recorded a sub-goal indicating the object was moved to a box before the box was fully moved to its new place, which could have resulted in a sub-goal with a wrong target or wrong order between sub-goals.

2) Teaching errors: Another common source of issues in provided demonstrations was teaching errors, when the correctness or informativeness of the performed tasks was compromised due to participants' errors. In some cases that occurred in all tasks, participants forgot to achieve a required sub-goal in their demonstrations or to fully press buttons so that the robot could detect a sub-goal (Type #5). Such issues led to missing sub-goals. When demonstrating Task 2, participants sometimes picked task objects at the wrong time, dropped them from the table, etc. (Type #6), which resulted in the system detecting unintended object-related sub-goals (wrong order) due to objects becoming invisible to the robot. When demonstrating Tasks 1 and 2, some participants started demonstrating before saying ready to the robot (Type #7), so that the demonstration was recorded from a non-empty state of the task. While this issue did not directly impact the correctness of demonstrations, it led to a loss of information when processing the collection of five demonstrations by the PLI system. In some cases when demonstrating Task 2, participants moved a box without pressing a required button but then corrected their actions by moving the box back, pressing the button, and moving it again (Type #8). As a result of these actions, the first sub-goal of the box being moved to the new place was recorded as an extra wrong sub-goal. Type #9 issues were when, in demonstrating Task 1, a few participants pressed the button to add water to the pot at the wrong time, i.e., before moving the pot to the sink. In some other cases when demonstrating Task 1, participants took the objects (i.e., potatoes) out of the pot, but this had no impact on demonstrations since the PLI system was not programmed to detect sub-goals when objects are removed from inside the boxes.

3) Inefficiency in performed task actions: Lastly, we identified instances when

the participants did not demonstrate the most efficient ways of performing the tasks (i.e., sequences with the fewest number of sub-goals) to the robot. In particular, participants sometimes moved the box unnecessarily between open and closed places (Type #11) when demonstrating Task 2, which led to extra sub-goals. When demonstrating Task 1, some participants placed the pot on the countertop one or more times (Type #12), which did not contribute to the task goal (e.g., potatoes could be added once the pot was in the sink or on the stove).

Figure 5.12 presents the occurrence of each type of performance issue in demonstrations of each task provided by each participant. As can be seen, certain types of issues (e.g., Types #1 and #10) were common across a large number of participants when demonstrating Task 1. Demonstrations of Tasks 2 and 3 contained fewer performance issues than Task 1. Figure 5.13 (A) illustrates the total number of performance issues from each source in demonstrations provided by each participant in each condition. Demonstrations by only 3 participants among those who completed the GHAs contained more than 5 issues related to being unaware of the system's limitations, while this number for No-GHA participants was 6. While demonstrations by none of GHA participants contained more than 5 performance issues due to action inefficiency, 4 No-GHA participants demonstrated with such issues more than 5 times. Issues due to teaching errors appeared to occur consistently across all groups.

Figure 5.13 (B) shows the occurrence of all performance issues averaged over the 15 demonstrations of 3 tasks provided by participants in each group. It can be seen that the impact of completing the GHAs was influenced by the access to the RCV. The LMM that was fit on this data, presented in Table 5.5, indicated that the effect of GHA was only significant in No-RCV condition ($se = 0.08, t = -3.055, p < .01$): when participants did not have access to the robot's camera view, completing GHAs led to demonstrations with significantly lower performance issues. According to Figure 5.13 (A), when participants had access to the robot's camera view, the average occurrence of performance issues with and without completing GHA was at the same level. The main effect of the RCV was significant according to the LMM: demonstrations provided by participants who had access to the robot's camera view included a significantly lower number of performance issues ($se = 0.08, t = -2.239, p < .05$). It was also observed that performance issues when demonstrating Task 1 occurred more often compared to when demonstrating Task 2 ($se = 0.04, t = 11.470, p < .001$) or Task 3 ($se = 0.04, t = 14.238, p < .001$).

Table 5.4: Different types of identified performance issues in task demonstrations. 12 types of issues are classified under 3 sources. The task(s) each issue was observed in, along with the impact of the issue on the provided demonstrations, are also noted.

#	Source of issue	Identified issue	In task(s): 1	2	3	Impact on provided demonstrations
1	Being unaware of system's limitation	Grabbing and moving two objects at the same time	✓	✓	✓	Wrong order[1]
2		Moving an object to a wrong box and then moving it to the right one			✓	Wrong target
3		Moving an object to a wrong box and taking it out			✓	Extra wrong sub-goal
4		Moving a box while an object is being held	✓	✓		Wrong target/order
5	Teaching errors	Forgetting a required sub-goal or to fully press a button	✓	✓	✓	Missing sub-goal
6		Demonstrating an object-related sub-goal at a wrong time		✓		Wrong order
7		Starting the task from a non-empty task state	✓	✓		Information loss
8		Moving a box without pressing a required button		✓		Extra wrong sub-goal
9		Pressing a button before moving a box into a required place	✓			Wrong order
10		Taking an object out of a correct box	✓			-
11	Actions ineffieciency	Unnecessarily moving a box between places		✓		Extra sub-goal
12		Moving a box to an unnecessary place	✓			Extra sub-goal

[1] This type of issue had no actual impact on demonstrations in our experiment because objects involved in this issue did not need to be moved to boxes in any particular order.

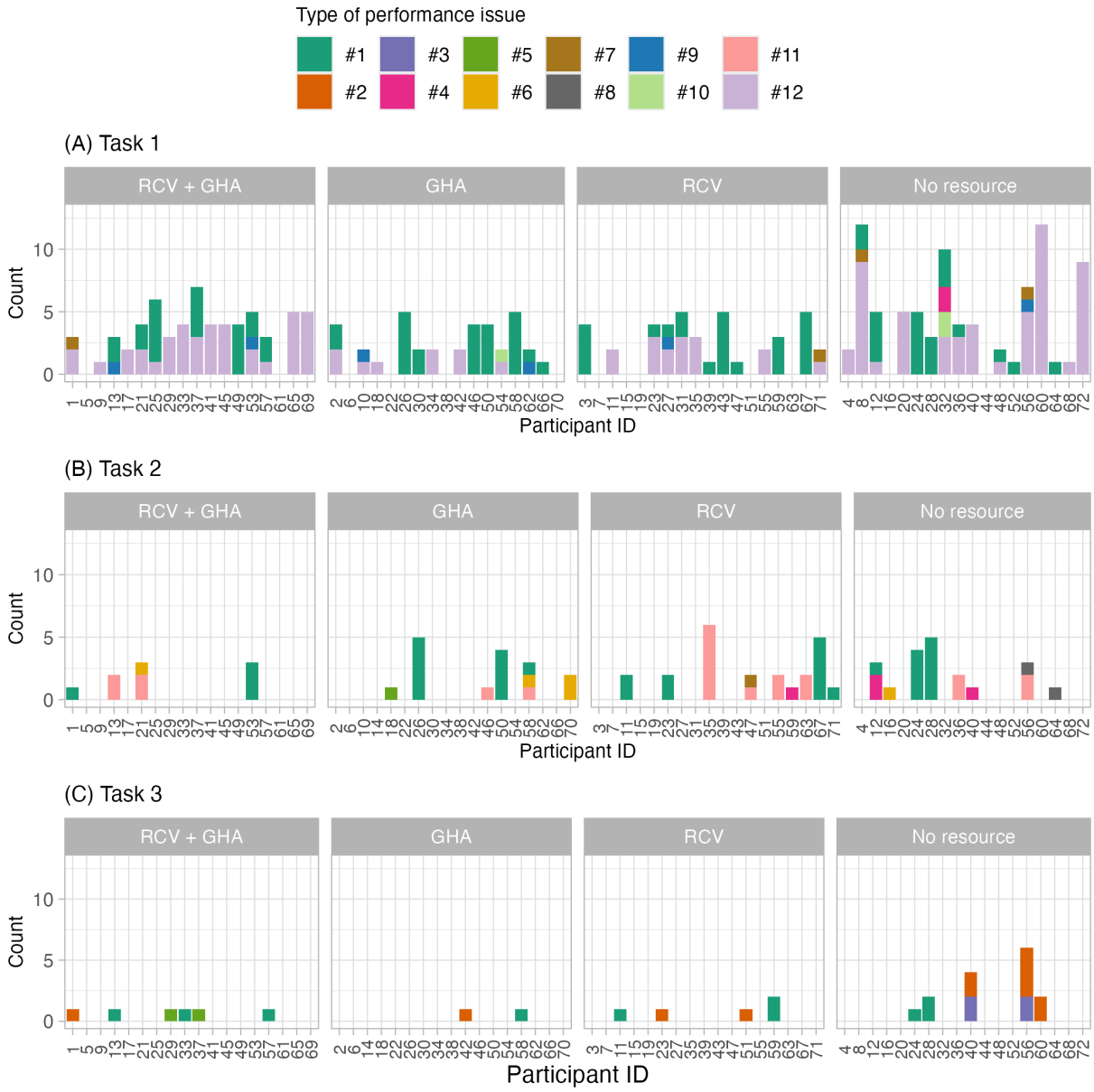


Figure 5.12: The occurrence of each type of performance issue in demonstrations of each task provided by each participant in the four experimental groups. Different types of performance issues, i.e., issues with what was performed in the demonstrations, are labelled in Table 5.4.

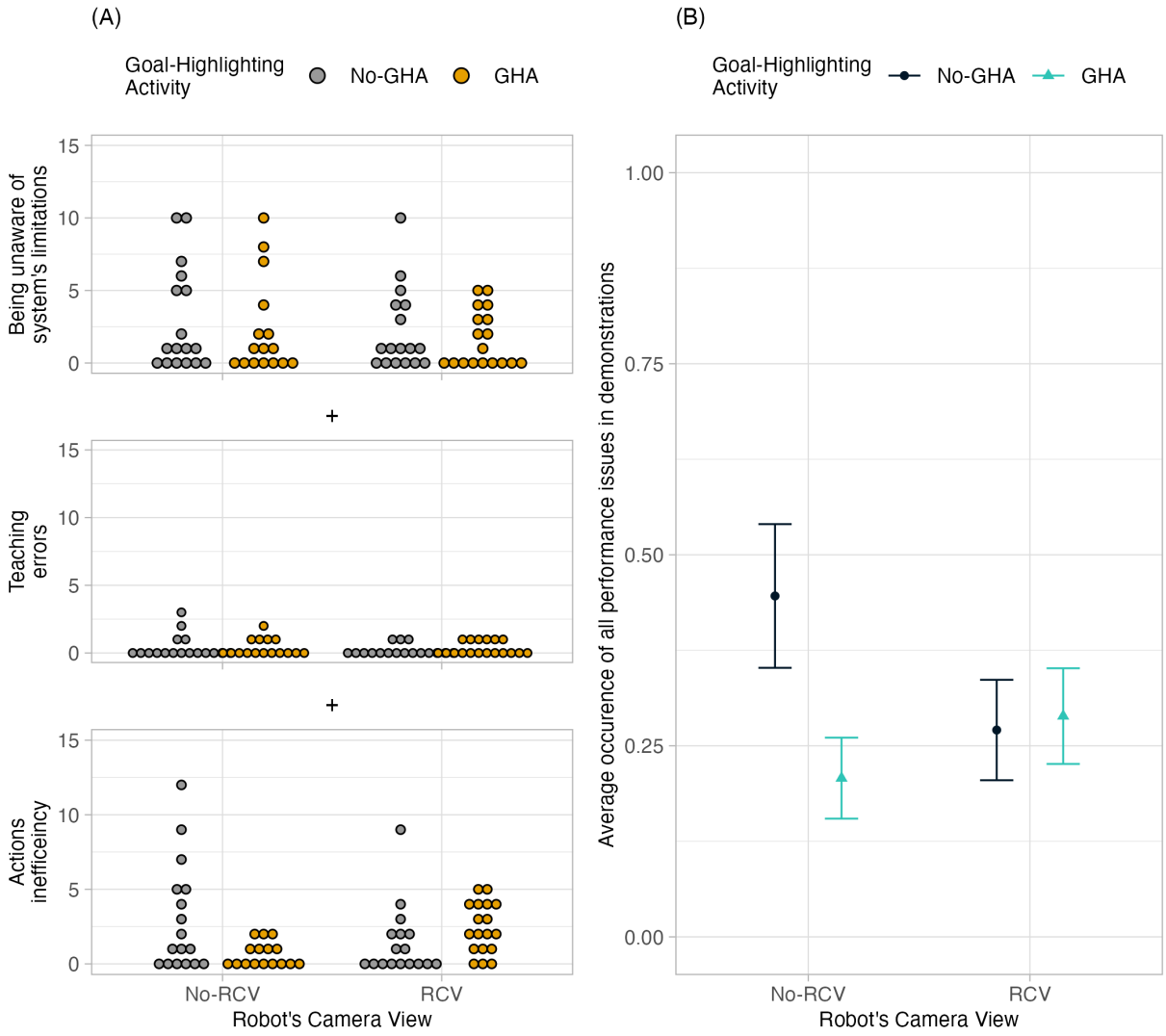


Figure 5.13: (A) Total number of performance issues due to each of the three identified sources in demonstrations provided by each participant across four groups. (B) The average occurrence of all performance issues in task demonstrations. Error bars indicate 95% confidence intervals.

Table 5.5: Linear Mixed-effects Model predicting the occurrence of all performance issues in demonstrations

Covariate	Estimate	SE	t	p
Goal-Highlighting Activity *GHA*	-0.240	0.0786	-3.055	**
Goal-Highlighting Activity *No-GHA*: Robot's Camera View *RCV*	-0.176	0.0786	-2.239	*
Goal-Highlighting Activity *GHA*: Robot's Camera View *RCV*	0.082	0.0786	1.042	
Task *2*	-0.430	0.0375	-11.470	***
Task *3*	-0.538	0.0378	-14.238	***
Task order	-0.031	0.019	-1.652	

No-RCV and *No-GHA* conditions and Task *1* are baselines. * = $p < .05$; ** = $p < .01$; *** = $p < .001$.

Sub-goal detection issues in demonstrations

A script compared the annotated sequence of sub-goals for each demonstration with the robot-detected sequence to identify sub-goal detection errors and calculated a sub-goal detection error rate. Figure 5.14 (A) presents the average sub-goal detection error rates for participants across the four groups and Table 5.6 provides the LMM fit on this metric. According to the LMM, the sub-goal error rate in demonstrations provided by participants who had access to the RCV was lower compared to participants who did not have this resource ($se = 2.01, t = -4.054, p < .001$). Higher participants' attention to the visibility of objects in the *placing* phase led to a significant reduction in the normalized sub-goal detection error rate ($se = 0.06, t = -3.441, p < .001$). The model also indicated a significant effect of task on the normalized sub-goal detection error rate: sub-goal detection errors occurred more often in Task 2 and Task 3 compared to Task 1 ($se = 1.16, t = 2.529, p < .05$ and $se = 1.17, t = 20.592, p < .001$, respectively). We observed significant effects of the order of the task taught to the robot and the teaching round within each task: sub-goal detection error rates were higher when demonstrating tasks later in the session ($se = 0.58, t = 2.225, p < .05$) and in later rounds when demonstrating a particular task to the robot ($se = 0.33, t = 2.980, p < .01$).

Table 5.7 summarizes the identified sources for the sub-goal detection issues. Figure 5.15 illustrates the occurrence of the issues with different sources in the demonstration of each task provided by every participant. One main source of sub-goal detection issues was invisible initial placement of objects (i.e., markers were not visible to the robot after the

placing phase), when demonstrating any of the tasks (Source #1). Such issues led to missing sub-goals in the robot-detected sequences.

Sub-goal detection issues also occurred in the *teaching* phase due to a number of reasons. In some cases, which happened in all of the tasks (Source #2), objects became invisible once they were picked by the participant. In these cases, robot-detected sequences were affected in multiple ways. The impact on the detected sub-goals was sometimes a wrong target (e.g., when an object became invisible closest to a box that was not its actual target), a wrong order between sub-goals (e.g., when an object became invisible before being picked to be moved to a box), and an extra sub-goal (e.g., when the battery in Task 3 became invisible while remaining on the table and the PLI system detected it as being moved into a box). In some cases when demonstrating Tasks 2 and 3 (Source #3), task actions were performed at a fast pace and very close to one another and the PLI system either detected sub-goals with a wrong target (e.g., when the box was not yet detected to be in its new place after an object was moved into it) or in a wrong order. Note that Sources #2 and #3 could not be easily separated and, therefore, in the rest of the analysis, are considered together. In some cases of demonstrating Tasks 1 and 2, inaccuracies in detecting the location of the boxes led to detecting extra sub-goals by the robot (Source #4), e.g., when a box was not actually moved to a new place but the system identified a sub-goal. Lastly, we observed a case leading to a missing sub-goal when a button was not identified accurately in Task 1 (Source #5).

The total number of sub-goal detection issues in the two phases of *teaching* and *placing* in demonstrations provided by each participant is presented in Figure 5.14 (B). Considering the sub-goal detection issues in the *placing* phase, the figure shows a reduction when participants had access to the RCV. While sub-goal detection issues occurred more than twice in demonstrations from only 6 RCV participants, this number for No-RCV participants was 22. Sub-goal detection issues in the *teaching* phase, however, did not appear to occur differently based on the level of the RCV factor.

Variability among demonstrations

The average demonstration variability scores for participants in each of the four groups are presented in Figure 5.16. The LMM fit on this data (Table 5.8) did not show any main effect of the experimental factors, but indicated an interaction between them ($se = 0.27, t = 2.101, p < .05$). As can be seen in Figure 5.16, there was a decreasing trend in the variability among demonstrations with completing the GHAs in the absence of the RCV, but an increasing trend when participants had access to the robot's camera view. The model also indicated significant differences in demonstration variability scores between

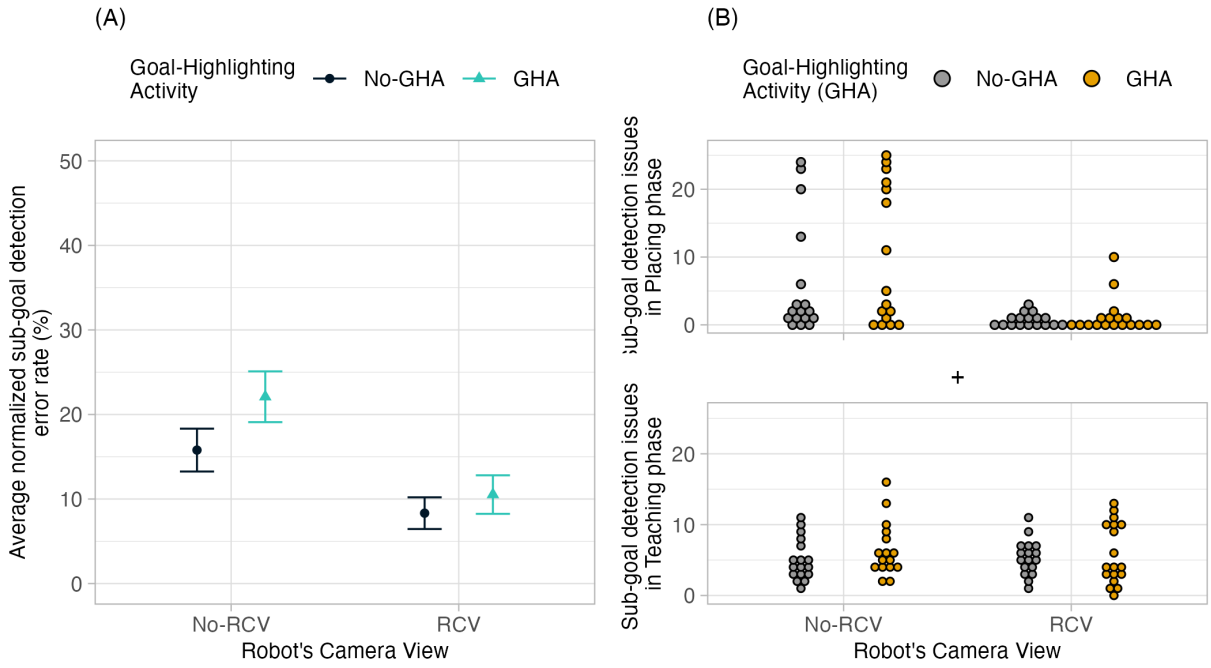


Figure 5.14: (A) The average normalized sub-goal detection error rate (%) in task demonstrations. Error bars indicate 95% confidence intervals. (B) Total number of sub-goal detection issues occurred in the *placing* and *teaching* phase of demonstrations provided by each participant across four groups.

Table 5.6: Linear Mixed-effects Model predicting normalized sub-goal detection error rate in demonstrations

Covariate	**Estimate**	**SE**	**t**	**p**
Robot's Camera View *RCV*	-8.146	2.009	-4.054	***
Attention placing phase	-0.206	0.060	-3.441	***
Task *2*	2.924	1.157	2.529	*
Task *3*	23.987	1.165	20.592	***
Task order	1.290	0.581	2.225	*
Round	0.996	0.334	2.980	**

No-RCV condition and task *1* are baselines. * = $p < .05$; ** = $p < .01$; *** = $p < .001$.

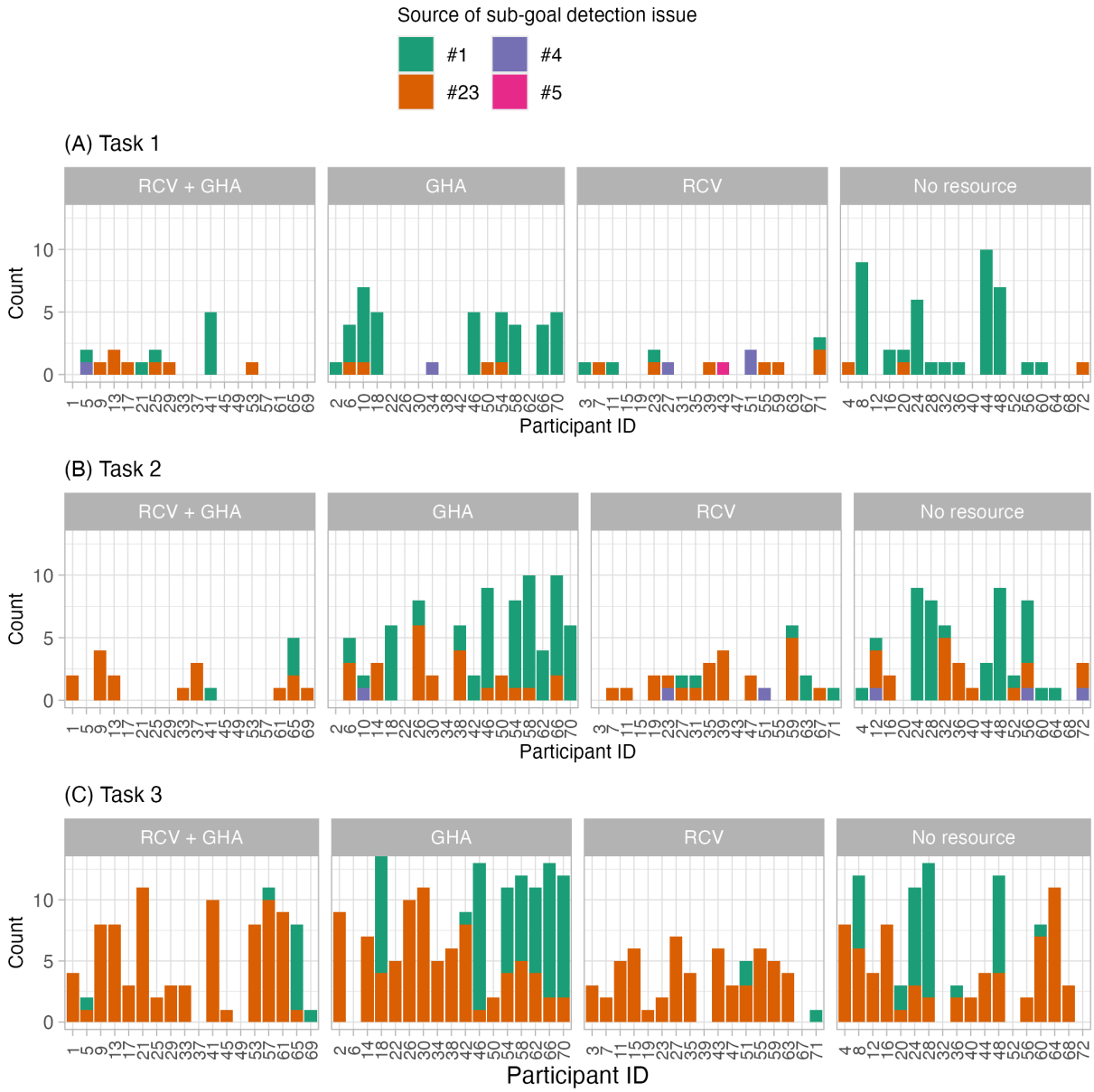


Figure 5.15: The occurrence of each source of sub-goal detection issue in demonstrations of each task provided by each participant in the four experimental groups. Different sources of sub-goal detection issues are labelled in Table 5.7.

Table 5.7: Different sources of observed sub-goal detection issues in task demonstrations. 5 sources of issues are classified as having occurred during the *placing* or *teaching* phase. The task(s) each issue was observed in, along with the impact of the issue on the recorded demonstrations, are also noted.

#	Phase	Observed issue	In task(s):			Impact on recorded demonstrations
			1	**2**	**3**	
1	*Placing*	Invisible object placement	✓	✓	✓	Missing sub-goal
2	*Teaching*	Objects becoming invisible[1]	✓	✓	✓	Wrong target/order/extra sub-goal
3		Fast actions		✓	✓	Wrong target/order
4		Inaccurate box location detection	✓	✓		Extra sub-goal
5		Inaccurate button location detection	✓			Missing sub-goal

[1] No sub-goal detection issue was counted when the battery in Task 3 was placed invisibly to the robot. This was because we could not make sure this was not a result of the participant considering the object irrelevant.

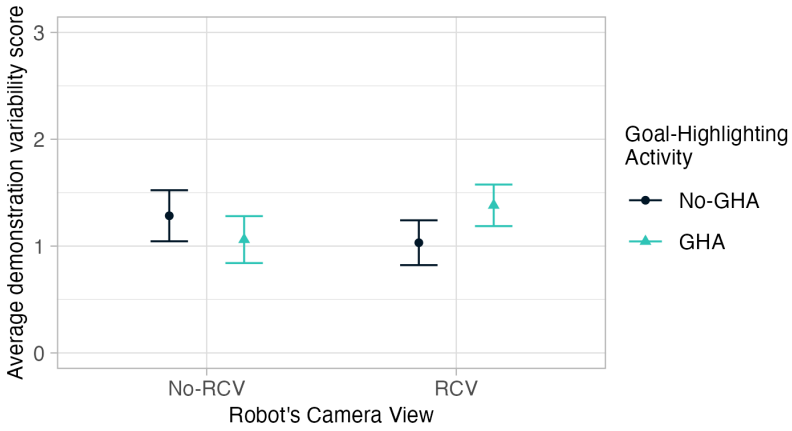


Figure 5.16: Average demonstration variability scores for participants across four groups. Error bars indicate 95% confidence intervals.

tasks. Overall, demonstrations of Task 1 had fewer variations compared to demonstrations of Task 2 ($se = 0.01, t = -2.852, p < .01$), but more variations compared to demonstrations of Task 3 ($se = 0.01, t = 3.165, p < .01$).

5.3.4 [RQ1.4] Robot Learning Outcomes

We investigated a number of measures and metrics (noted in Section 5.2.5) to evaluate the robot learning outcomes, mostly determined by the generated programs. These measures

Table 5.8: Linear Mixed-effects Model predicting demonstration variability scores in demonstrations

Covariate	**Estimate**	**SE**	**t**	**p**
Robot's Camera View *RCV*	-0.245	0.194	-1.264	
Goal-Highlighting Activity *GHA*	-0.238	0.194	-1.230	
Task 2	0.271	0.095	2.852	*
Task 3	-0.303	0.096	-3.165	**
Robot's Camera View:Goal-Highlighting Activity	0.574	0.273	2.101	*

No-RCV and *No-GHA* conditions and task *1* are baselines. * = $p < .05$; ** = $p < .01$.

Table 5.9: Linear Mixed-effects Model predicting number of sequences in the generated programs

Covariate	Estimate	SE	t	p
Robot's Camera View *RCV*	-0.059	0.724	-0.081	
Goal-Highlighting Activity *GHA*	-0.830	0.717	-1.158	
Task 2	0.240	0.547	0.439	
Task 3	-1.804	0.551	-3.273	**
Attention Highlighting	0.006	0.011	0.590	
Robot's Camera View:Goal-Highlighting Activity	2.209	1.015	2.177	*

No-RCV and *No-GHA* conditions and task *1* are baselines. * = $p < .05$; ** = $p < .01$.
'Attention Highlighting' refers to the participants' self-reported attention to performing specific actions that the robot could learn from (responses to Q3).

are: the number of sequences in programs, program generation scores, and participants' subjective learning evaluations.

Number of sequences in each program

Figure 5.17 presents the average number of sequences in the generated programs with demonstrations from participants in each group. In this measure, trends similar to those observed in the demonstration variability scores were detected. The LMM fit on this data, presented in Table 5.9, did not indicate any main effect of experimental factors, but highlighted an interaction between them ($se = 1.02, t = 2.111, p < .05$). When participants demonstrated tasks to the robot without access to the RCV, completing the GHAs led to a reduction in the total number of sequences, but when participants demonstrated tasks to the robot with access to the robot's camera view, an opposite trend was seen. The model also indicated that for Task 3, significantly fewer sequences existed in the generated programs, compared to Task 1 ($se = 0.55, t = -3.273, p < .01$).

Program generation score

As noted in Section 5.2.5, the sum of scores assigned to all sequences in the program generated for each task was calculated for every participant. Because the range of possible scores was different across tasks (e.g., the maximum score for Task 1 could be 120, but for Task 3, it could be only 6), we performed z-score normalization on the program generation

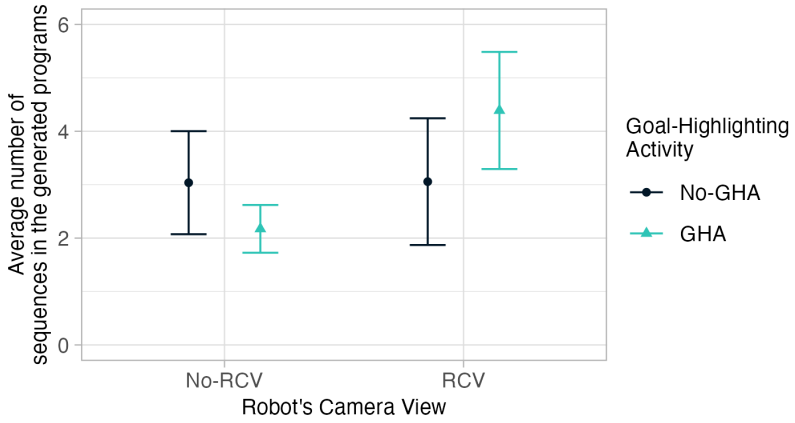


Figure 5.17: Average number of sequences in the generated programs based on demonstrations of participants in four groups. Error bars indicate 95% confidence intervals.

scores of each task and calculated *normalized program generation scores*. The average normalized program generation score is shown in Figure 5.18 for each task and in each group.

The LMM presented in Table 5.10 indicated the significant effect of having access to the robot's camera view on higher program generation scores ($se = 0.14, t = 2.854, p < .01$). The model also predicted an order effect: the normalized program generation scores were significantly lower when participants taught tasks later in the experiment ($se = 0.07, t = -2.015, p < .05$). We also ran Wilcoxon tests to study differences between program generation scores for each task and found that the program generation scores for Task 2 were significantly higher with participants who completed GHAs compared to those who did not when participants had access to the robot's camera view ($W = 252, p < .01$).

The classification of sequences in the task programs generated based on demonstrations by each participant is presented in Figure 5.19. In all the cases, a varying number of sequences was present in each generated program. When learning Task 1, sequences in programs were valid in most cases, while in learning Task 3, we often observed incomplete sequences. The performance of program generation when learning from non-experts will be discussed more in-depth later in Section 5.3.7.

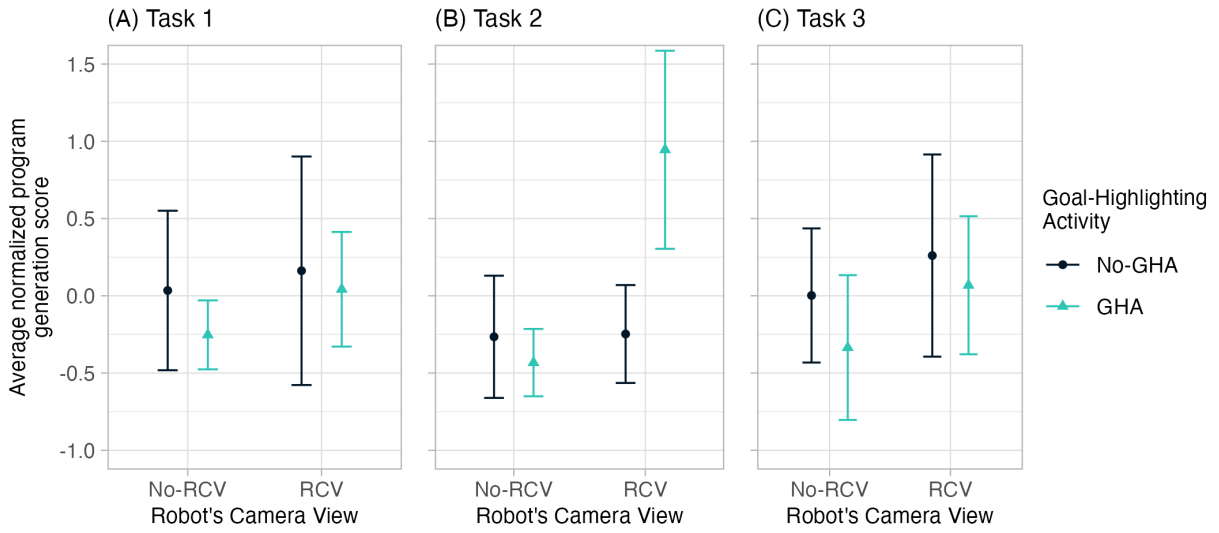


Figure 5.18: Average normalized program generation score when learning each task across four groups. Error bars indicate 95% confidence intervals.

Table 5.10: Linear Mixed-effects Model predicting normalized program generation score based on demonstrations

Covariate	**Estimate**	**SE**	**t**	**p**
Robot's Camera View *RCV*	0.40	0.14	2.854	**
Task order	-0.15	0.07	-2.015	*

No-RCV condition is the baseline. * = $p < .05$; ** = $p < .01$.

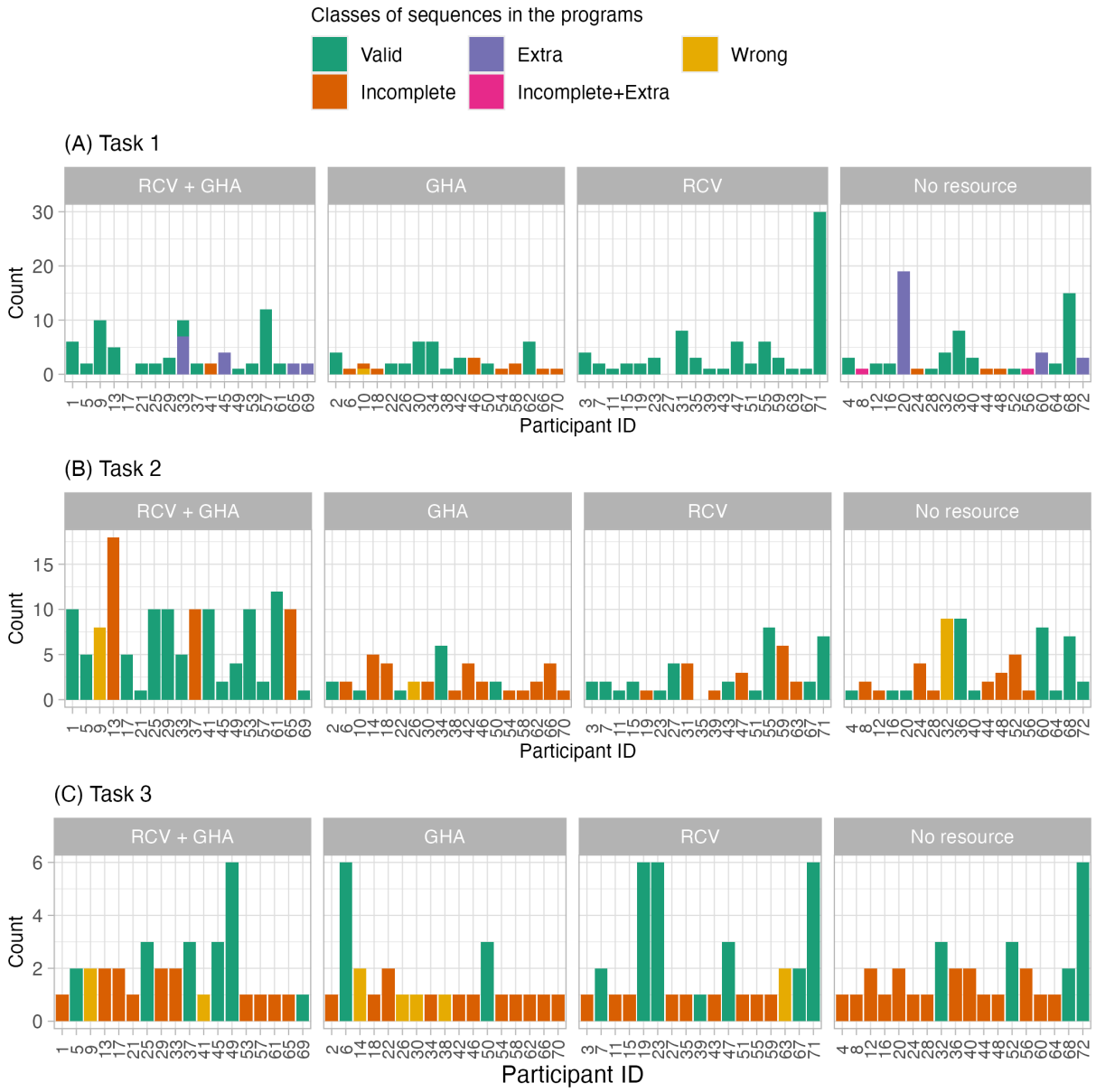


Figure 5.19: The number of sequences from different classes in the task programs generated using demonstrations from each participant in the four experimental groups. Sequences in each program are classified as valid, incomplete, extra, or wrong, as explained in Section 5.2.5. Note that the vertical axes for the three tasks have different scales due to varying possible correct ways of performing the tasks.

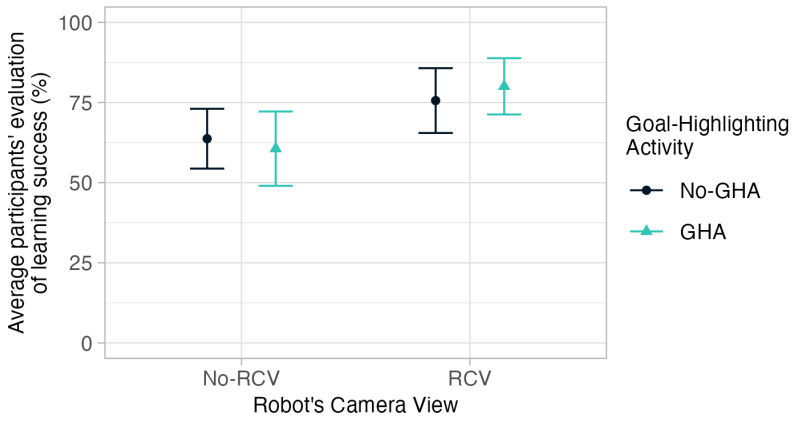


Figure 5.20: Average participants' subjective evaluation of learning success in each session across four groups. Error bars indicate 95% confidence intervals.

Subjective robot learning performance evaluations

Figure 5.20 presents average participants' responses to the question of "how well or badly do you think the robot learned the tasks?". Based on the LMM fit on this data, responses were affected only by the presence of the robot's camera view: participants in groups with access to the RCV evaluated the learning success of the robot higher compared to other participants ($estimate = 15.667, se = 4.63, t = 3.382, p < .01$). Participants' responses to the open-ended part of this question are discussed later in Section 5.3.7.

5.3.5 [RQ1.5] Perceptions of Robot's Social Attributes

Average participants' perceptions of the social attributes of the robot, on three scales of competence, warmth, and discomfort in the RoSAS questionnaire, are illustrated in Figure 5.21. Overall, the ratings of warmth and discomfort appeared to be at the same level across all the groups (no difference was found according to LMMs). However, we found that ratings of robots' competence were significantly higher when participants who had access to the robot's camera view taught the robot ($estimate = 0.534, se = 0.22, t = 2.422, p < .01$). RCV was the only significant factor in this model.

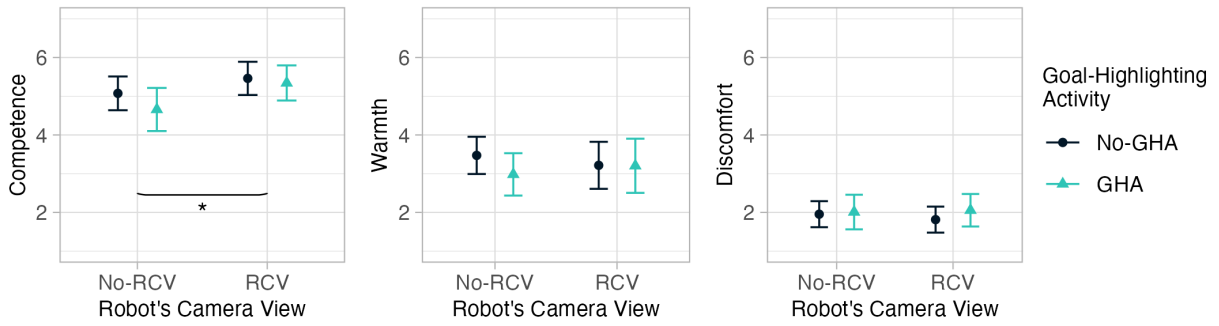


Figure 5.21: Average participants' perception of robot's social attributes (RoSAS questionnaire [146] scales) across four groups. Error bars indicate 95% confidence intervals. * = $p < .05$.

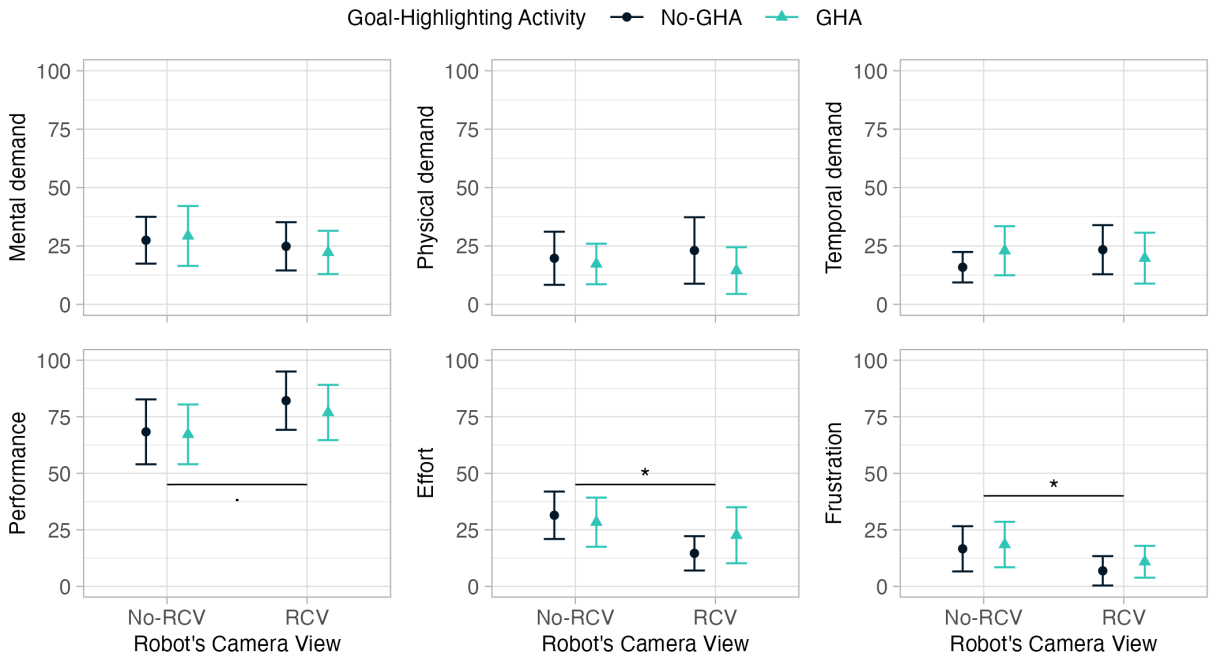


Figure 5.22: Average participants' perception of teaching workload (NASA-TLX questionnaire [85] scales) across four groups. Error bars indicate 95% confidence intervals. . = $p < .1$; * = $p < .05$.

5.3.6 [RQ1.6] Perceptions of Teaching Workload

Results for the six scales of the NASA-TLX questionnaire are presented in Figure 5.22 for each condition. No significant impact of experimental conditions was observed in the scales of mental demand, physical demand, and temporal demand according to LMMs. While the average perceived performance of teaching tasks in Figure 5.22 appeared to be higher for participants who had access to the robot's camera view, compared to No-RCV participants, the effect of RCV according to the LMM was only approaching significance ($estimate = 11.722, se = 6.10, t = 1.924, p < .1$). The LMMs also indicated that participants who had access to the RCV perceived significantly less effort ($estimate = -11.306, se = 4.878, t = 8.673, p < .05$) and less frustration ($estimate = -8.667, se = 3.957, t = -2.190, p < .05$) compared to those without access to this resource. We observed no significant effect of the considered confounding factors in any of the models.

5.3.7 [RQ2] Effectiveness of PLI when Non-experts Teach a Robot

Participants' comments on the performance of the robot when learning the tasks and on the process of teaching the robot, as well as sequences in the generated programs in the experiment and also in a simulated scenario, were analyzed to study RQ2.

Subjective robot learning performance evaluations

When answering the open-ended part of the learning evaluation question (Q9 in Table 5.2), in most cases, participants explained examples of instances when the robot made errors or performed successfully, exhibited unexpected behaviours, etc. For instance, a participant noted *"Performed the first task okay (didn't sort one of the objects), second task perfectly, third task poorly (did not organize 2/3 objects)."* Overall, form 69 recorded responses, 18 participants only noted positive aspects of robot learning, 46 mentioned both positive and negative aspects, and 5 only noted negative observations.

Three participants indicated that they noticed the robot performed a variation of the task that was most optimal, possibly different from what they demonstrated to the robot. Examples include: *"[...] Also, I notice what the robot performed might be different from what I taught,"* and *"It learned quite quickly and usually performed the most optimized variation of the task."*

On the other hand, five participants noted aspects of robot learning and performance that did not appear to be optimal. This included the slow movements of the robot and

the robot inaccurately picking up objects, in comments such as *"The actions were mostly correct, but it did not correct for mistakes when it misdropped an item. It also moved very slowly and not in an optimal or human way,"* *"The articulation of the robot was very impressive. While I added variations to the tasks, some including unnecessary actions, the robot performed the most efficient pathway and accurately performed each task,"* and *"I think the robot did very well with the tasks, there were no mistakes. The only issue is the speed and precision the robot had. It was slow to identify an object, and the precision of its hand to pick up the object wasn't very accurate."*

An interesting observation was that one participant was able to correctly notice that the learning of the robot is based on identifying task structures, i.e., set of rules, as they indicated *"The robot seems to pick up some 'rules' (organizing the orange item in the orange bin, pushing the button to allow the bin to move past a point), but it doesn't seem to pick up ALL the rules. [...]."*

Subjective evaluations of the teaching process

We analyzed 72 participants' responses to the question asking them to evaluate the teaching experience and whether they found any particular aspect of teaching difficult (Q10 in Table 5.2). Overall, by employing the TA method, the following three themes were identified by merging relevant codes. Under the first theme, five sub-themes were accepted with at least 3 codes associated with each of them. Another researcher coded 25% of the data and, using Cohen's Kappa [32], we performed an inter-rater reliability analysis to check the agreement between the two coders. The results indicated an almost perfect agreement ($Kappa = 0.955, SE = 0.032, CI_{95\%} = [0.892, 1.000]$).

1) Positive teaching experience: 55 participants indicated that they had a positive experience when teaching the robot in the experiment. The following sub-themes investigate the responses in more depth:

- **Ease of teaching the robot**: 33 participants' comments indicated that teaching the robot by the method used in the experiment was fairly easy. For instance, participants commented that *"I do not think there was anything difficult in the teaching process (P22),"* and *"I think it was good, it did not feel too difficult apart from making sure everything was in view of the robot (P37)."* Among these 33 comments, 2 also cited that the tasks used in the experiment were easy tasks.
- **Enjoyment of teaching**: 32 participants noted that teaching the robot was a fun and enjoyable experience. Example comments include *"Teaching the robot was*

interesting (P28)," *"Fun, I love the experience (P35),"* and *"Actually the experience was above expectation, I really enjoyed it a lot (P55)."*

- **Enjoyment of watching robot's performance**: Comments from 9 participants included a statement indicating that they enjoyed watching the robot performing the learned tasks itself (e.g., *"Very interesting to see the robot give the task a try after the teaching as well (P36)."*).
- **Novelty of teaching experience**: 3 participants noted that the experience of teaching the robot was a new experience for them, unlike anything they had experienced before. These comments were *"It was amazing!!! Teaching a robot was truly a new experience to me and I found it interesting (P43),"* *"It was so new (P45),"* and *"I've never taught the robot before, so it was a very different experience for me (P58)."*
- **Improved experience with time**: 4 participants noted that their experience improved over time in the teaching session. For example, a response that also indicated *enjoyment of teaching* was: *"I liked it! It was quite terrifying seeing it at first but the voice was actually quite calming and the atmosphere was nice. I enjoyed this experience (P56)."*

2) Difficulties with aspects of teaching: 7 participants noted that they found some aspects of teaching the robot difficult. Among them, 4 participants cited that teaching the robot and repeating demonstrations five times was time-consuming (e.g, *"It took quite a lot of time to teach the robot simpler tasks (P38)."*), 2 participants noted the complexity of Task 2 (e.g., *''It was not difficult at all, except for the first one, I forgot to close the lid of the box for the first two rounds since that task has the most steps to remember (P21)."*), and 1 participant mentioned that they found it difficult to include variations in their demonstration (*"I found it difficult to do things in different ways to help the robot learn better (P52)."*).

3) Robot's outstanding learning ability: 7 participants commented that the robot learned the tasks well in the experiment. In particular, 4 comments noted the learning performance despite the limited number of provided examples (e.g., *"I'm impressed the robot was able to learn from only 5 demonstrations (P11)."*) and 2 comments cited the variations among the demonstrations (*"It learned very well considering how many variations i provided the robot with (P4),"* and *"I think the robot is quite smart as I had mix up the orders of how I was adding/moving things, and it still got it right most of the time (P13)."*).

The second part of the teaching evaluation question (Q10 in Table 5.2) asked if participants wanted to share any thoughts or suggestions on the process of teaching the robot.

33 participants responded to this part of the question (responses are presented in Table 5.11). Some participants proposed a new strategy for themselves for teaching the robot in the future, some proposed technical improvements related to the robot and also to the human-robot teaching interaction, and some expressed curiosity about different aspects of the teaching experience and robot learning. Moreover, 7 comments included reflections on the process of teaching a robot, with 4 noting that they thought about including task variations in their demonstrations (e.g., *"While teaching it I tried to mix up the way I did things so the robot would have more options to learn from which I think is more beneficial (P36)."*), 2 noting the usefulness of robots learning household tasks (e.g., *"I think that this robot can be very useful in households to do repetitive tasks (P24)."*), and one noting the usefulness of teaching and testing approach (*"Just going back and forth between teaching and testing would help understanding what was wrong during teaching (P30)."*).

Classification of the sequences in the generated programs

As could be seen in Figure 5.19, for learning Task 1, the PLI system generated programs with at least two sequences when processing sets of demonstrations from 51 (71.8%) out of 71 participants (note that Task 1 teaching data from one participant was excluded). Robot learning resulted in programs with only a single sequence from 18 (25.3%) out of 71 sets of demonstrations. The system failed to generate any programs in 2 (2.8%) out of 71 cases. Valid sequences existed in the task programs with demonstrations from 49 (69%) participants. In 23 out of these 49 cases (46.9%), the PLI system generated programs with at least one 'unseen' sequence, i.e., sequences of sub-goals that were not recorded by the robot as part of the five task demonstrations (as a result of specific placements of sub-goals in the demonstrated sequences and also performance or sub-goal detection issues). The number of unseen sequences in those 23 programs of Task 1 was on average 4.78 ($SD = 5.59$).

For Task 2, learning a program with at least one sequence was unsuccessful in only one (1.4%) out of 72 cases. Programs for this task with at least two sequences were generated from demonstrations provided by 50 (69.4%) out of 72 participants. In 39 cases (54.2%), the generated programs contained at least one valid sequence. Among these 39 programs, 17 cases (43.6%) contained at least one unseen sequence, with an average of 5.29 ($SD = 2.26$) unseen sequences within the 17 programs.

Considering Task 3, only 19 (27.1%) out of 70 generated programs had valid sequences (note that Task 2 data from two participants was not included in the analysis). The system was able to identify all 6 possible ways of performing this task from 6 participants. In most cases, 44 (62.8%) out of 70, sequences in the programs were incomplete. The PLI system

Table 5.11: Participants' thoughts or suggestions on the robot teaching process.

Category		**Participants' comment**
Reflecting on teaching and proposing a strategy		*"I should have understood that I need to move the items to exact locations and be a bit slower with my movements (P10)."* *"If I were able to learn more about how the robot understands the movement, I could have done better (P38)."* *"Clearly for the second one, I generalized a bit too much and should have just done the task the same way each time (P47)."* *"Maybe I was rushing involuntarily, and it may have affected how the robot was following the visual instruction (P66)."*
Proposing a robot-related improvement	Learning modality	*"The robot could understand voices [...] (P8)."*
	Actions speed	*"The robot still move not very fast/quickly (P15)."* *"Also the timing, it felt very slow, so not suitable for tasks where you need to act very fast (P57)."* *"I would only suggest [improving] the speed and maneuverability of the robot (P69)."*
	Motions accuracy	*"The robot should lower the hands, cuz I found the objectives are easy to fall outside the box in task 1 (P16)."* *"When the robot is picking something up, it tends to tilt its body in that direction while moving its arm, when it can just simply move its arm to do it. Also, the robot always raises the object too high before dropping them (P49)."* *"It only uses one hand at a time can't work concurrently. That's what I felt a bit odd (P57)."* *"Improving the motion of the robot so it wouldn't have to bend all the way forward to drop an object into the box (P69)."* *"Maybe the orientation of the garbage piece while placing it in the box should be considered by the robot to prevent this from happening and increase accuracy (P72)."*
	Voice clarity	*"The instructions [...] were a bit difficult to hear (P24)."* *"Sometimes I found understanding the robot was difficult because of all the noise it created (P25)."* *"The voice was a bit muffled (P31)."* *"During the first experiment, it was a bit challenging understanding the robot's accent (P44)."*

Table 5.11: [Cont.] Participants' thoughts or suggestions on the robot teaching process.

Category	**Participants' comment**
Proposing an experimental improvement	*"It was a bit weird to work with the foam cylinders instead of real objects (P2)."* *"Buttons to press could be stationary so it doesn't move (P28)."* *"Rearrangement of items of task and seeing how that impacts (P40)."* *"For the third task of garbage collection, have a separate area to put batteries; otherwise, very hard to see if the robot actually intends to leave the battery in place or not (P41)."* *"It would be more realistic if there is real items instead of codes, and maybe he could recognize through some photo and pattern recognition model (P55)."* *"I think that maybe it might be better for the items to be bigger/maybe the QR code since it seems like it cannot catch on the QR codes for the items provided in front of it (P58)."*
Expressing curiosity	*"I wonder if variations on a procedure might confuse it, and if it might be better to let it master one way of doing things before teaching another (P14)."* *"I hope more changing actions can be taught to the robot and I'm wondering if it can learn from that [...] (P15)."* *"Not sure if I could have been faster or slower [...] (P20)."* *"I wondered how the ordering of the objects used would affect the end action. I also wondered if, for less linear tasks, the order by which things were done would influence the action of the robot in the end (P25)."* *"I was not sure about the speed at which I should do the task. Most of the task I did slowly so that robot can see my actions but few times I tried to do it fast also (P45)."* *"I only worried about if some unexpected incidents happened in the teaching process, will they influence the actions of the robot? (P54)."* *"I try to teach it in curriculum learning, from easy task to more complex task. I don't know whether I could only teach the basic tasks or I could do that curriculum learning tasks (P60)."* *"What if you put the items and tasks in the same order, [...] (P63)."* *"I wonder how the robot learned to what extent it has to push the pot/boxes. Are the steps that the robot learns quantized? Also, I did not try pushing a button two times and I'd like to see what the robot would learn in this case (P68)."*

was able to generate a program based on demonstrations from any of the participants. With the 19 valid programs, 11 (57.9%) included at least one unseen sequence. On average, 1.62 ($SD = 0.67$) unseen sequences were in those 11 programs.

Learning randomly versus learning from participants

To further study the role and effectiveness of non-experts teaching the robot using the PLI system, we compared the program generation performance of the actual case when demonstrations were planned by non-expert human participants in the experiment (*experimental* case) with a simulated case when five demonstrations were 'randomly' picked (*random* case). The random case simulates a situation where a robot randomly observes how potentially different humans perform the tasks. In the random case, a script was executed 72 times and selected 5 random sequences from all 360 *robot-detected sequences of sub-goals* based on demonstrations provided by participants for each task, and computationally generated 5 task demonstrations. This was to include performance and sub-goal detection issues similar to those observed in the experiment. Then, the PLI system generated a program with each of the generated sets of demonstrations.

According to a Wilcoxon test ($W = 1428, p < .001$), the average program generation score when learning Task 1 randomly from multiple people ($M = 6.11, SD = 4.99$) was significantly higher compared to learning it from individual demonstrators in the experiment ($M = 3.48, SD = 4.55$). When learning Tasks 2 and 3, however, no significant difference was observed in the program generation scores between the experimental case and the random case ($W = 2530.5, p = .807$ and $W = 2513, p = .979$, respectively).

5.3.8 [RQ3] Characteristics of Demonstrations Over Time

To answer RQ3, we characterized the provided demonstrations by the total demonstration time, the time spent on demonstrating each sub-goal, and a fraction that captured the rate of disappearance of task components or objects.

Time spent on providing demonstrations

Figure 5.23 (A) presents the average time spent on providing each of the 15 demonstrations in a session, in a chronological order (i.e., from the beginning to the end of sessions, regardless of the nature of the tasks). The LMM provided in Table 5.12 predicted a significant impact of completing the GHAs on reducing the total demonstration time ($se =$

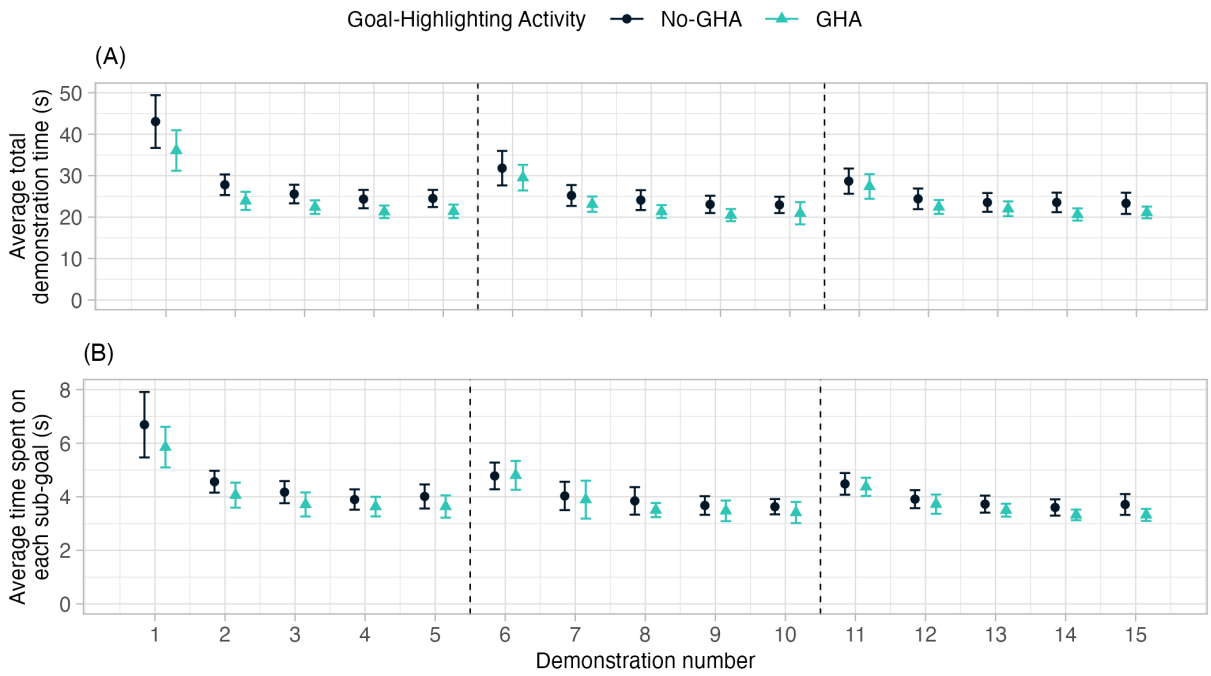


Figure 5.23: Average time spent on (A) providing demonstrations and on (B) demonstrating each sub-goal over fifteen demonstrations (three tasks are separated by dashed lines) provided in a session for participants who did or did not complete goal-highlighting activities. Error bars indicate 95% confidence intervals.

$0.982, t = -2.843, p < .01$), which can also be seen in Figure 5.23 (A). The model also indicated that participants spent a longer time demonstrating Task 2 ($se = 0.475, t = 2.896, p < .01$), while they spent a shorter time demonstrating Task 3 ($se = 0.474, t = -14.750, p < .001$), both compared with Task 1. More importantly, we observed that participants provided task demonstration faster after more rounds of demonstrating each task ($se = 0.137, t = -16.787, p < .001$) and also after demonstrating more tasks in a session ($se = 0.237, t = -7.001, p < .001$).

Average time spent on demonstrating each sub-goal

As shown in Figure 5.23 (B), the average time spent on achieving each sub-goal in demonstrations also followed trends similar to the total demonstration time. The LMM fit on data indicated the impact of completing GHAs on demonstrating each sub-goal faster ($se = 0.174, t = -2.467, p < .05$). In Task 3, participants demonstrated each sub-goal

Table 5.12: Linear Mixed-effects Model predicting total demonstration time

Covariate	**Estimate**	**SE**	**t**	**p**
Goal-Highlighting Activity *GHA*	-2.791	0.982	-2.843	**
Task *2*	1.374	0.475	2.896	**
Task *3*	-6.995	0.474	-14.750	***
Demonstration round	-2.301	0.137	-16.787	***
Task order	-1.661	0.237	-7.001	***

No-GHA condition and Task *1* are baselines. ** = $p < .01$; *** = $p < .001$.

Table 5.13: Linear Mixed-effects Model predicting total time spent on each sub-goal in demonstrations

Covariate	**Estimate**	**SE**	**t**	**p**
Goal-Highlighting Activity *GHA*	-0.428	0.174	-2.467	*
Task *2*	-0.014	0.092	-0.149	
Task *3*	0.603	0.092	6.572	***
Demonstration round	-0.351	0.027	-13.250	***
Task order	-0.324	0.046	-7.055	***
Attention placing phase	-0.015	0.005	-2.923	**
Gender *Male*	0.465	0.180	2.581	*

No-GHA condition and Task *1* are baselines. * = $p < .05$; ** = $p < .01$; *** = $p < .001$.

slower compared to Task 1 ($se = 0.092, t = 6.572, p < .001$). We also observed significant reduction trends in demonstration times of individual sub-goals, both across five demonstration rounds for each task ($se = 0.027, t = -2.250, p < .001$) and across the three tasks in a session ($se = 0.046, t = -7.055, p < .01$). Two confounding variables were also found to significantly impact this measure. Higher participants' self-reported attention to the visibility of objects in the *placing* phase led to faster sub-goal demonstration times ($se = 0.005, t = -2.923, p < .01$). Male participants spent more time demonstrating each sub-goal to the robot compared to female participants ($se = 0.180, t = 2.581, p < .05$).

Fraction of frames where task components or objects disappeared

The average fraction of frames where task components and objects were invisible to the robot is presented in Figure 5.24 for all demonstrations in a session. According to the LMM presented in Table 5.14, completing the GHAs led to a higher task components disappearance rate ($se = 0.003, t = 2.621, p < .05$). The normalized task components disappearance rate increased significantly over five demonstration rounds of each task ($se = 0.000, t = 6.190, p < .001$).

Considering the fraction of frames where only task 'objects' were invisible to the robot, the LMM presented in Table 5.15 indicated a significant effect of the RCV factor. For participants who had access to the robot's camera view, the normalized task objects disappearance rate was significantly lower compared to other participants ($se = 0.007, t = -2.434, p < .05$). The model also indicated that the normalized task objects disappearance rate was significantly higher for participants who completed the GHAs ($se = 0.007, t = 2.975, p < .01$). This measure was also found to be affected by the type of task: when demonstrating both Tasks 2 and 3, task objects became invisible less often compared to when demonstrating Task 1 ($se = 0.003, t = -3.429, p < .001$ and $se = 0.003, t = -2.893, p < .01$, respectively). Lastly, we found a significant effect of demonstration round ($se = 0.001, t = 8.849, p < .001$) and ($se = 0.001, t = 4.126, p < .001$) task order on increasing normalized task objects disappearance rate.

5.3.9 [RQ4] Patterns to Find Better Teachers for the Robot

We examined our data at an individual level to study RQ4. We averaged the key measures and metrics we defined and used to assess the quality of the recorded demonstrations and robot learning outcomes previously studied to address RQ1.3 and RQ1.4 for all demonstrations each participant provided. In the same way, we also averaged the three characteristics

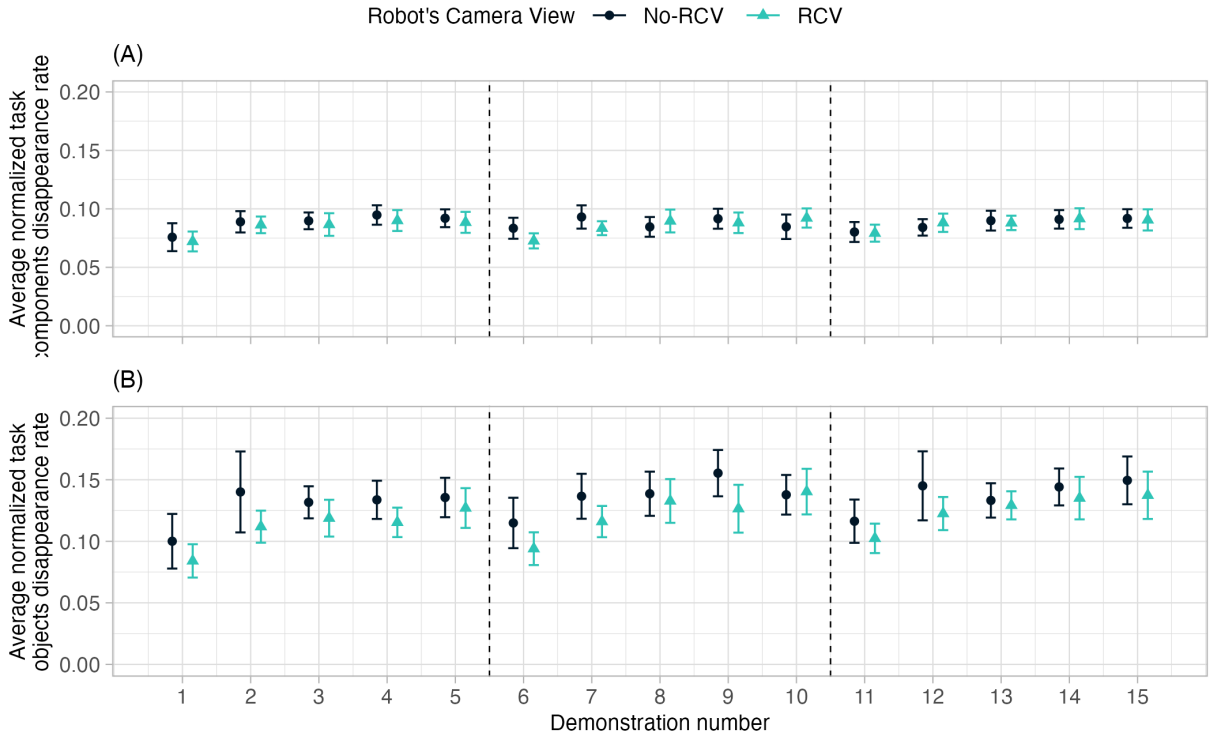


Figure 5.24: Average normalized (A) task components and (B) task objects disappearance rate in demonstrations over fifteen demonstrations (three tasks are separated by dashed lines) provided in a session for participants who did or did not have access to the robot's camera view. Error bars indicate 95% confidence intervals.

Table 5.14: Linear Mixed-effects Model predicting normalized task components disappearance rate in demonstration

Covariate	**Estimate**	**SE**	**t**	**p**
Goal-Highlighting Activity *GHA*	0.008	0.004	2.624	*
Demonstration round	0.003	0.000	6.190	***

No-GHA condition and Task *1* are baselines. * = $p < .05$; *** = $p < .001$.

Table 5.15: Linear Mixed-effects Model predicting normalized task objects disappearance rate in demonstration

Covariate	Estimate	SE	t	p
Robot Camera View *RCV*	-0.017	0.007	-2.434	*
Goal-Highlighting Activity *GHA*	0.020	0.007	2.975	**
Task *2*	-0.011	0.003	-3.429	***
Task *3*	-0.009	0.003	-2.893	**
Demonstration round	0.008	0.001	8.849	***
Task order	0.006	0.002	4.126	***

No-RCV and *No-GHA* conditions, and Task *1* are baselines. * = $p < .05$; ** = $p < .01$; *** = $p < .001$.

of demonstrations that were discussed when addressing RQ3. The results are presented in Figure 5.25. On the right side of this figure, a darker green colour indicates higher demonstration quality and better robot learning outcome when learning from each participant. On the left side, a darker blue colour means higher values in each characteristic of demonstrations. As can be seen, the average of measures for individual participants varied considerably, with no clear pattern visible to characterize good learning performance.

To gain a better understanding of this data, we ran multiple Pearson correlation tests between each of the three demonstration characteristics and each of the five demonstration and learning outcomes metrics. After the results were corrected using the Holm-Bonferroni method for 15 tests, we found a strong positive correlation between the task objects disappearance rates and the normalized sub-goal detection error rates ($r(70) = 0.62, p < .001$), visualized in Figure 5.26. It can also be seen in Figure 5.25 that darker and lighter colours are correlated in the third and the fifth columns.

5.4 Discussion

5.4.1 [RQ1.1] Attention to the Visibility of Demonstrations

Our results suggested that when human teachers for a robot had an extra resource that showed what the robot could see, i.e., the output of the robot's vision system displayed on a monitor, they reported paying more attention to ensuring task objects were visible to the robot in their demonstrations. This was in line with our initial expectation about the

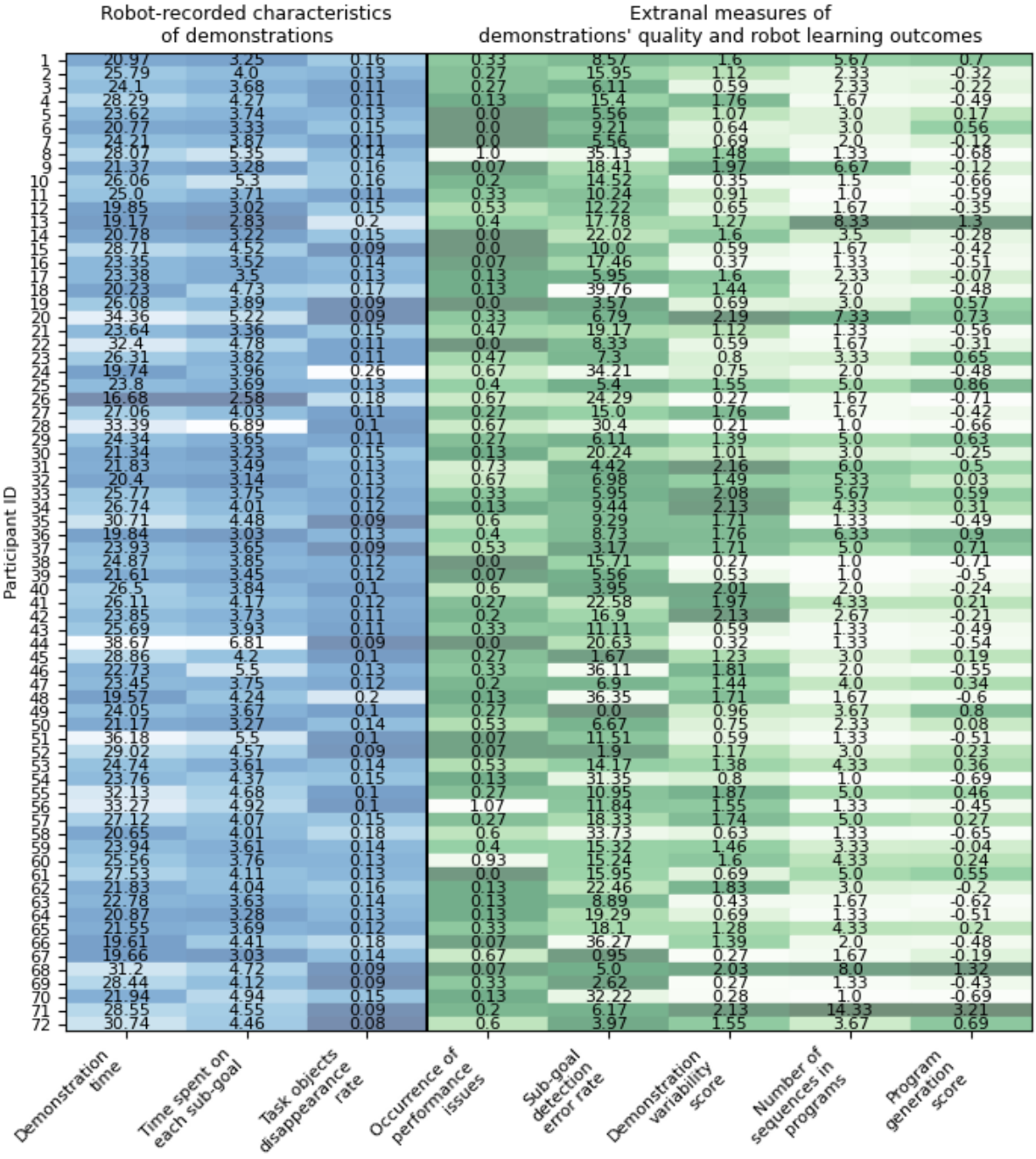


Figure 5.25: The average of demonstrations' characteristics (the first three columns) as well as metrics and measures assessing the quality of the recorded demonstrations and robot learning outcomes across all demonstrations provided by each participant. The vertical axis presents the ID assigned to each of the 72 participants.

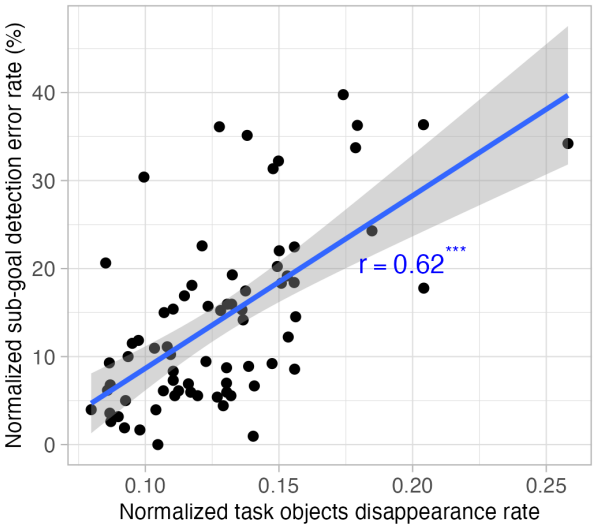


Figure 5.26: The linear correlation found between the normalized sub-goal detection error rates and normalized objects disappearance rates for all demonstrations each participant provided. *** = $p < .001$.

positive impact of offering the robot's camera view to human teachers.

On the other hand, with the completion of our goal-highlighting activities, participants reported paying less attention to objects' visibility in their demonstrations. A reason might be that after completing the GHAs, participants paid more attention to what is performed in their demonstration (e.g., the type of actions, the arrangement of sub-goals, etc.) compared to how the demonstrations are performed (i.e., visible to the robot).

Participants reported higher attention to the visibility of objects when placing task objects in front of the robot compared to when teaching the tasks. The difference in self-reported attention between the two phases may be because, in the *teaching* phase, participants had to think about how to perform the tasks, and therefore, their attention may have been directed to the tasks themselves. This explanation was supported by the results suggesting that participants in RCV condition checked the screen that displayed the robot's camera view more often and found it more useful when *placing* task objects compared to when *teaching* the robot. In principle, when placing task objects, participants could place the objects in an arbitrary way, see if they were detected by the robot after looking at the screen and checking for green markers, and then adjust the orientation of each object to make sure all of them are detectable by the robot.

The robot's camera view offered an opportunity for perspective-taking at both VPT

levels. Ensuring that task objects remain within the robot's sight was related to VPT Level 1, while recognizing that objects may not always appear visually useful to the robot (e.g., in case markers can be seen by the teacher but cannot be seen by the robot) is related to VPT Level 2. Note that the RCV also provided feedback to human teachers as it contained information on whether or not the robot had registered the presence of task objects.

5.4.2 [RQ1.2] Attention to Demonstrating Primitive Actions

We found no evidence to support that any of the extra resources provided to the participants, especially putting them in the position of the robot when labelling task actions by asking them to complete the GHAs, made an impact on participants' self-reported attention to demonstrating primitive actions so that the robot could learn effectively.

To measure participants' attention to this aspect of the robot's limitations and capabilities, the present study relied only on a question asking participants how much they paid attention to performing specific actions our robot could learn from, and no extra actions (Q3 in Table 5.2). This question might not have been able to accurately capture what was needed to be measured to study the research question and evaluate the impact of the experimental manipulations. Although what we implied by 'specific actions' was noted below the question (refer to Q3 in Table 5.2), non-expert participants might have understood the attention in question not as intended. In fact, participants still demonstrated 'the actions the robot could learn from' in cases where they combined the sub-goals, performed them very fast, etc. More detailed questions or other subjective and/or objective measures could replace this question in the future, so that participants' attention to demonstrating primitive actions could be better measured.

Another important point to note is that, based on participants' responses to the question asking them about the usefulness of the GHAs, the activities did not appear useful to nearly half of the participants who provided a comment. For many of these participants, the GHAs appeared irrelevant to the tasks of teaching the iCub robot. Nevertheless, as will be discussed in the rest of this section, GHAs were found effective in improving some aspects of the quality of demonstrations and robot learning outcomes.

5.4.3 [RQ1.3] Quality of Recorded Demonstrations

To analyze demonstrations' quality, we evaluated issues in what participants performed in demonstrations (i.e., performance issues), issues in detecting actions in demonstrations by the robot (i.e., sub-goal detection issues), and variations in each set of demonstrations

participants provided (because a robot with the PLI system is expected to better generalize tasks by observing variations in how a certain task is performed). We identified various types of performance issues in demonstrations. We found that, overall, performance issues in demonstrations were less often observed when participants completed the GHAs. This finding could be explained by potentially higher participants' attention to performing primitive actions in such a way that the robot could learn effectively (although it was not supported in RQ1.2). 'Being unaware of the system's limitation' was among the three identified sources of performance issues and included issues related to combining two sub-goals (Types #1 and #4) and achieving sub-goals in a manner that the robot was not programmed to be able to detect them properly (Types #2 and #3). Such issues appeared to be lower in demonstrations of participants who had completed the GHAs, according to a trend observed in Figure 5.13 (A). The GHAs could have increased participants' awareness of the system's limitations concerning the demonstrated actions and their relevance to robot primitives, and therefore, decreased the rate of performance issues overall.

Having access to the robot's camera view was also found to reduce performance issues in demonstrations, but only in the absence of GHAs. This result might imply that GHAs were effective in reducing such issues to a point where further reduction was difficult to achieve with non-expert participants, even after the RCV was also provided to them. Based on trends observed in Figure 5.13 (A), the reduction in performance issues with the source of 'actions inefficiencies' contributed the most to the overall performance issue reduction when completing the GHAs in the absence of the RCV. When the only extra resource provided to the participants was the GHAs, and the RCV did not direct participants' attention to itself, participants might have paid more attention to the actions and to demonstrating more efficient ways of performing the task to the robot, in particular, given that the examples in the GHAs did not include any extra, unnecessary, or repetitive actions.

We found that sub-goal detection error rates were significantly reduced with participants' access to the RCV and their increased self-reported attention to ensure the visibility of objects in the *placing* phase. Among the observed sub-goal detection errors, missing sub-goals could be directly caused if the robot was not able to see objects after participants placed them in the *placing* phase. Unsuitable initial placement could also lead to objects disappearing before they were actually moved to any box in the *teaching* phase.

The score we assigned to each set of demonstrations to measure the variability among them was not found to be affected by either having access to the RCV or completing the GHAs. Only a significant interaction between these two factors was observed.

All three metrics discussed to evaluate the quality of demonstrations varied by the nature of the task being taught to the robot. Among the three tasks in this experiment,

demonstrations of Task 3 had the lowest performance issues and the highest sub-goal detection error rates. Task 3 could be the easiest for participants to demonstrate, as sorting waste items was likely part of everyday tasks done by university students, and this task only included two kinds of actions. However, demonstrations of Task 3 could be the hardest for the robot to record, because it involved more objects and boxes than the other two tasks. As a result, sub-goal detection errors related to objects and boxes, such as invisible object placement in the *placing* phase and objects becoming invisible before moving to a box in the *teaching* phase, occurred more often compared to other tasks. While sub-goal detection issues were the lowest among the three tasks in Task 1, demonstrations of this task contained the highest performance issues. As observed in Figure 5.12, grabbing and moving two objects at the same time (related to being unaware of the system's limitations) and moving a box to an unnecessary place (related to actions inefficiency) occurred very often in demonstrations of Task 1. Participants often moved the two potatoes to the pot at the same time, which is also a strategy expected to be employed when boiling two potatoes in real life, presumably to increase the time efficiency of the actions. When teaching the robot, this strategy resulted in combining two sub-goals, which was considered an issue given the limitations of the robot's learning system. When demonstrating Task 1, participants also often moved the pot to an unnecessary place, i.e., to the countertop, to add potatoes. This step may be completely reasonable when carrying out such a task in a real-life setting, but could be skipped to show the robot the most efficient way of doing the task.

The variability among the demonstrations was the highest in Task 2 and the lowest in Task 3 among the three tasks. This finding could be due to the fact that Task 2 had two different completion strategies based on organizing yellow or orange objects first, which could considerably change the place where other sub-goals appeared in the sequences of actions. On the other hand, the strategies for completing Task 3 could be varied only based on the order in which the three waste items had to be moved to the bins, which could be hard to keep track of over five demonstrations to perform using a new sequence each round. As a result, participants might have demonstrated the same sequence multiple times.

Lastly, we observed an effect indicating that sub-goal detection issues happened more often when demonstrating tasks later in the sessions and in later demonstration rounds in each task. Fatigue or boredom may have caused participants to perform actions in a less clear manner that was more difficult for the robot to detect.

5.4.4 [RQ1.4] Robot Learning Outcomes

We evaluated robot learning outcomes based on objective measures and metrics derived from the PLI system, and also based on the subjective evaluation of participants as they watched the robot performing the learned tasks during the experiment. To begin with, the number of sequences in generated programs could indicate the generalizability of programs to new, unseen scenarios. The identified factors affecting this measure were similar to those impacting demonstration variability scores. This finding was anticipated because the PLI system is expected to generate programs with more sequences if the recorded demonstrations are more diverse. The only exception was that we did not observe any effect indicating that a larger number of sequences were in programs of Task 2 compared to Task 1, in line with the observed higher variability scores in Task 2 compared to Task 1. In Task 2, sub-goal detection errors could have made the variations among the demonstrations more difficult to detect by the robot and, therefore, prevented programs from having more sequences.

We proposed a metric called program generation score, based on the correctness and completeness of sequences within the programs, to proxy the success of the robot in performing the learned tasks. The results showed that higher program generation scores were achieved by learning from demonstrations of participants who had access to the robot's camera view. The program generation score is a complex metric. A high score could be due to a high number of generated sequences as a result of high variations in the order of actions in the demonstrations. While we did not find any evidence suggesting that participants who had access to the RCV may have included more variations in their demonstrations, when addressing RQ1.3, we found that sub-goal detection errors were reduced by offering the RCV to participants. In the absence of the RCV, such errors could have impacted program generation scores by causing missing or wrong target sub-goals to be part of the generated sequences, leading to lower scores. We also observed an order effect indicating lower program generation scores in tasks taught later in the session, which could be related to the observed increased frequency of sub-goal detection issues in demonstrations over time.

While analyzing program generation scores, we observed one instance when completing the GHAs could positively impact robot learning outcomes. When participants had access to the robot's camera view, completing GHAs significantly improved program generation scores of Task 2. This task was the most complicated to demonstrate (due to the presence of abstract constraints in the task). Future studies should investigate what factors contributed to the effectiveness of the GHAs in that particular situation of teaching such a task.

Subjective robot learning performance evaluations were found to be more positive when

participants had access to the RCV. It is important to note that the robot performed each task only once for each participant in the experiment, based on the estimated fastest sequence of actions in the generated program. In this case, the task performance of the robot could only give feedback to the teachers about the completeness of sequences in the program or the correctness of one particular sequence. Moreover, the task performance of the robot could have contained various errors related to the precision in performing primitive actions, accuracy of the vision system in detecting the location of task components, random environmental noises (e.g., unsuitable placement of task components), etc., which could have negatively impact to robot's performance and lowered participants' evaluations. Such negative impacts, however, are expected to have been equally present in all different experimental groups, as we tried to run the sessions uniformly. A human teacher would need to observe the robot's performance over multiple situations or have access to indicators that quantify the performance of robot learning in the PLI system to be able to better evaluate the robot's learning success.

5.4.5 [RQ1.5] Perceptions of Robot's Social Attributes

We found that participants who taught the robot with access to its camera view output evaluated the robot to be more competent. This effect was in line with the higher learning success of the robot (as indicated by a number of metrics discussed earlier) when being taught by participants in RCV groups: a robot that learned the tasks well appeared more competent to participants. As we could also expect, the two other social attributes of the robot considered in this study, i.e., warmth and discomfort, were in the same range across all experimental groups, as they were not directly related to the performance of a robot that learns household tasks with minimal social interaction (which was consistent across the experimental groups).

5.4.6 [RQ1.6] Perceptions of Teaching Workload

Our results suggested that participants who had access to the RCV found it easier to accomplish their level of performance in teaching the robot (lower 'effort' sub-scale in the NASA-TLX questionnaire) and were less frustrated during their experience of teaching the robot. These findings support the usefulness of offering the visual perspective of a learner robot to non-expert teachers as a support tool for improving human-robot teaching interaction and reducing aspects of teaching workload, without making the teaching process more mentally demanding.

5.4.7 [RQ2] Effectiveness of PLI when Non-experts Teach a Robot

To determine the effectiveness of using a PLI framework with non-expert humans teaching a robot, we studied participants' viewpoints on the learning performance of the robot (to determine whether or not the robot learned effectively from their perspective) and on the ease of the teaching process; and further investigated the quality of the generated programs, regardless of experimental groups. When participants evaluated the learning performance of the robot, in most cases, both negative and positive aspects were noted. However, only 5 (6.9%) out of 72 participants commented solely on the negative outcomes of robot learning, including issues, unexpected behaviours, etc.

We analyzed participants' comments on the teaching process using the TA method. The most occurring theme in the responses was positive teaching experience, with sub-themes of ease of teaching the robot and enjoyment of teaching being mentioned by at least 44% of participants, and only 5 participants (5.5%) commented negatively, citing difficult aspects of teaching. These subjective responses suggest that teaching an iCub robot with our implemented PLI framework via visually demonstrating tasks was positively evaluated by non-expert users. It should be noted that because the details of the learning system and underlying mechanisms were not shared with participants until the end of the experiment, their evaluations is likely to mostly reflect on the modality of the teaching used (i.e., observational learning), the procedure of the experiment (i.e., demonstration each task five times and then watching robot's performance), and the resources that were provided to some of participants (i.e., the RCV and the GHAs).

Taking into account the generated programs using demonstrations provided by each participant, the effectiveness of robot learning varied by the nature of the task. Learning how to perform Task 1, i.e., boiling two potatoes, was mostly successful in the sense that the robot could correctly perform the task in at least one way most of the time (when learning from 69% of participants). In learning how to organize waste items in Task 3, for example, the robot in most cases (62.8%) learned the task partially (i.e., moving between 0 to 2 waste items into the bins and closing the bins). With the selected parameters for the PLI system (refer to Section 5.2.3) in this experiment, each item should have been at least detected by the robot to be moved to the correct bin four times across the five demonstrations for correctly learning how to move a waste item to a bin. With the occurrence of sub-goal detection errors that resulted in sub-goals with wrong target for the waste items, at most three correct sub-goals were detected in those instances when the learning from five demonstrations resulted in incomplete sequences in the programs. As will be noted in Section 5.6, the learning outcomes of Task 3 could be improved in the future by enhancing the accuracy of detecting the sub-goals in the demonstrations.

Overall, across all three tasks, the PLI system failed to generate a program in only 3 (1.4%) out of 213 cases. In fact, in cases when learning was successful (i.e., sequences in the programs were all valid), we observed that between 44% to 58% of programs of the three tasks contained valid sequences that were not observed by the robot in the five demonstrations. In such cases, the PLI system was able to generalize the sequence of actions to complete the tasks, which could potentially help the robot perform more efficiently in a wider range of scenarios.

To further analyze the effectiveness of learning from non-expert humans, we compared learning tasks from the five demonstrations of a single participant with simulated cases of learning from five random demonstrations containing the same amount of performance and sub-goal detection issues. We found that learning Task 1 from random observations was significantly more effective, according to the higher average program generation score, while learning the other two tasks was not found to be different between the two cases. These results may imply that for some tasks, such as boiling potatoes (Task 1), individuals may have a particular preference in arranging their actions, which can be different from another person's preference. By watching multiple people, a robot can observe more diverse examples and therefore, generate programs containing more sequences, which can lead to better generalizability. On the other hand, for some other tasks such as Tasks 2 and 3, participants did not seem to display a strong preference and varied their demonstrations sufficiently so that learning from different individuals did not significantly improve robot learning outcomes.

With all these quantitative observations, we can still argue that observation learning with a PLI system was effective, depending on the task and also affected by the occurrence of errors in performing demonstrations by the participants and detecting demonstrations by the robot.

5.4.8 [RQ3] Characteristics of Demonstrations Over Time

We studied the overall time spent on demonstrating each task and the average time spent on demonstrating each sub-goal in the tasks to investigate whether such characteristics of demonstrations change over time (as was the case in our previous study presented in Chapter 3). As we expected based on our previous study, participants demonstrated tasks and also each sub-goal faster after gaining more robot teaching experience and adapting to teaching the robot, both across the three tasks in a session and across five demonstrations for each task. Note that the modality of teaching was different in this experiment compared to our previous work (kinesthetic versus observational learning).

The rates at which any of the task components or only task objects disappeared from the robot's view during demonstrations were also analyzed when addressing RQ3. In particular, we defined a metric called the normalized task objects disappearance rate, which was found to be increasing with more teaching experience (both across the tasks and across the demonstrations of a single task). This effect indicated that task objects more frequently disappeared later in the session and later within the five demonstrations of each task. The observed increase in the objects disappearance rate was in line with the observed increase in the sub-goal detection error rate in demonstrations and the decrease in the program generation scores after teaching more tasks, which could be due to participants' fatigue or boredom. Both these effects were opposed to our expectations, based on results presented in Chapter 3, that the quality of demonstrations could be improved with participants gaining more robot teaching experience. Note that our previous results were concerning the success of demonstrations, defined as the area of the tasks covered by kinesthetic demonstrations, which inherently differ from metrics defined in this chapter for the quality of visual task demonstrations. Future studies are needed to identify the underlying reason for these contradictory findings.

As a side note, in reference to the resources provided to participants, we observed that teaching times (both overall and per sub-goal) were reduced while the normalized disappearance rates of task components and task objects were increased when participants completed the GHAs. More importantly, the normalized task objects disappearance rates were reduced when participants had access to the RCV. This observation could be expected as these groups of participants had also noted higher attention to the visibility of objects, which could have resulted in preventing them from disappearing in the robot's field of view during the demonstrations.

5.4.9 [RQ4] Patterns to Identify Better Teachers for the Robot

While we discussed effects related to our experimental manipulations, we found that metrics including participants' self-reported attention, the characteristics of the provided demonstrations, and the robot learning outcomes varied considerably across individual participants in all the groups. Not all participants in one group taught the robot similarly. For example, while some participants in the No-RCV groups reported full attention to the visibility of objects or provided demonstrations that were recorded by the robot with minimum amount of errors, some other participants in the RCV groups reported that they did not check the robot's camera view often or provided demonstrations that were hard to be accurately detected by the robot.

As noted in Section 1.2, finding an answer to the question of 'who to imitate' in LfD requires establishing metrics that a robot can observe in human teachers to estimate the suitability of the teacher and potential for learning from them. We looked for correlations between the metrics that were recoded from the robot's perspective and the external measures of demonstration quality and robot learning outcomes that were defined in this study. We found a strong correlation between one of the metrics used to assess the quality of the recoded demonstrations, i.e., normalized sub-goal detection error rates, and normalized task objects disappearance rates that were captured using data from the robot itself. Using this relationship, a robot can potentially estimate the sub-goal detection error rate in demonstrations of a human teacher by knowing the rate at which task objects become invisible in the robot's field of view. This rate could be calculated by the robot itself by knowing how many objects were involved in the task at the beginning of receiving a demonstration. A robot can, e.g., ask for such information from the human demonstrator, or obtain it from other observed demonstrations of the same task. It is important to note that both these metrics and the correlation between them were reported with the presence of the experimental manipulations, i.e., at least one teaching resource was provided to 75% of participants.

5.5 Limitations

The conducted experiment and data analysis to address the research questions in this chapter had several limitations. In this work, in order to manipulate participants' awareness of the limitations and constraints faced by a learner robot, we provided some groups of participants with two additional resources to help them better understand the robot's capabilities and constraints. Our experimental manipulations were therefore concerned with only two aspects of limitations and constraints of a learner robot: its limited visual field of view of tasks, and its learning strategy that only allowed learning from certain primitive actions. These two constraints only exist for robots with a certain type of LfD system. Robots that learn from humans may have a variety of other constraints and limitations, where increased attention from human teachers may improve robot learning outcomes. While robots that learn from visual observations will need to clearly see the demonstrations within their visual field of view to effectively learn from them, a robot that uses kinesthetic learning may be limited to learn from trajectories that do not contain self-collisions [67], or a robot that learns by verbal instructions may be constrained to clear, non-affective instructions [109]. Moreover, the constraint about the ability of a robot to only learn from a limited set of sub-goals is specific to a robot that uses the current implementation of the PLI system.

There are also limitations about the way we manipulated participants' awareness of the capabilities and limitations of the robot that were considered in this experiment. We offered participants an 'opportunity' to take the visual perspective of a learner robot before and while teaching it the tasks. As our results also suggested, not all participants who had access to the RCV paid full attention to and made use of the provided VPT opportunity, possibly due to focusing on other matters, such as planning and performing actions in their demonstrations. While the GHAs that we designed for this experiment were intended to increase participants' awareness about clearly presenting primitive actions that a robot with a PLI system can learn from, as discussed in Section 5.4.2, our results did not indicate any significant effects suggesting an increase in this aspect of teachers' awareness. It is also important to note that the present study does not answer the question of whether or not human teachers for robots can naturally develop awareness of the limitations and constraints faced by a learner robot, or what factors contribute to higher awareness in certain participants without any extra resources.

Lastly, we studied RQ2, RQ3, and RQ4 using data from participants in all groups, including those who had access to the RCV and/or had completed the GHAs. Because only 18 participants taught the robot without any provided resources, considering all participants in the analysis could increase the sample size, with a wider range of variations among the performance of individual participants (note that the sample remained balanced with respect to the experimental conditions). This aspect may limit the generalizability of our findings regarding the research questions noted above to situations where teaching a robot occurs naturally, e.g., with no access to what the robot can see.

5.6 Future Directions

While we considered several metrics to evaluate robot learning outcomes, future work can still improve this aspect. Most importantly, a wider range of subjective and/or objective measures can be included in the future to assess participants' attention to and awareness of robot limitations. For instance, by tracking the gaze behaviour of human teachers, future studies can more accurately identify participants who paid more attention to the visual constraints of the robot. This can be done through evaluating participants' visual attention to the screen showing the RCV, or even by checking whether or not human teachers visually checked areas of the task that could be invisible to the robot. Wider subjective measures (e.g., more detailed questions, interviews, etc.) could also be utilized in the future to assess the amount of attention to highlighting primitive actions in demonstrations provided by the participants.

As noted in Section 5.5, the present study only considered limitations with respect to the vision system of a robot that learns via observation and the ability of the robot for learning only from certain basic actions. Future work can study human teachers' awareness of learner robots' capabilities and limitations from a broader perspective, e.g., when robots utilize other learning mechanisms, receive demonstrations through different modalities, etc. Moreover, humans may develop varying amounts of attention, knowledge, or awareness of the underlying learning system of a robot they use over time, and that may impact robot learning outcomes. In the future, a robot may try to increase these factors in human teachers by providing feedback in a social learning scenario. With respect to the limitations of the vision system of a robot, the robot may signal to human teachers when objects become invisible, when actions are not detectable, etc., through channels such as voice, lights, or even a robot's non-verbal behaviours (as discussed, e.g., in [14]). To support non-expert users to pay attention to constraints of the learning mechanism of a robot and understand the robots' learning strategies, the robot can also convey its internal states, beliefs, etc., which can allow cognitive perspective-taking in human teaching [86]

Future research can also focus on designing resources for non-expert users, improving what was used in this experiment, with the goal of enhancing the learning outcomes of the robots in human-robot teaching interactions. Examples of other modalities of feedback include a robot providing verbal remarks about the visibility of objects to human teachers (e.g., by saying "I cannot see this object anymore") or providing feedback through augmented reality (e.g., projecting the boundaries of its vision system on the task). The output of the robot's vision system might also include additional visual cues related to the detection of task components as well as actions in demonstrations. This feedback might be provided via a robot's own screen (if a robot is equipped with such a feature), or an external screen that might be placed in different locations. The placement of the screen and its distance to the teachers' lines of sight can change the required amount of gaze changes and, therefore, influence the usefulness of such a resource. Future work can study the impact of such design choices on the learning success of a robot, ease of teaching the robot by non-experts, teaching workload, etc.

5.7 Summary and Conclusion

Phase 3 of my research was primarily focused on factors in human teachers for the robots that can lead to better robot learning outcomes. I explored the role of human teachers' awareness of a learner robot's visual constraints and learning strategies in impacting human-robot teaching outcomes to address O3. As part of this experimental study, non-

expert human teachers were enabled to take the visual perspective of a learner robot, which was found to help them pay more attention to ensure higher visibility of demonstrations to the robot. Several aspects of human-robot teaching were improved when the robot learned tasks from participants who could take the visual perspective of the robot. Designers of robots that learn from visual inputs may consider using a screen (either placed next to the robot or, for certain robot embodiments, integrated into the robot's torso) to display a robot's visual perception as a feedback channel to human teachers. This approach can be feasible for many robots that are designed to operate in household environments or as assistants. These findings highlight the importance of human teachers' awareness of and attention to a robot's visual constraints in observational learning scenarios. The experiment also offered a virtual activity that put participants in the position of the robot when labelling task actions in some example scenarios, but the results did not indicate any evidence suggesting the activities could help them pay more attention to ensure primitive actions are demonstrated in a way that the robot can effectively learn from them.

Human teachers' awareness of and attention to the limitations and constraints faced by a learner robot and its indicators can be used by robotic systems to improve robot learning by choosing more suitable teachers and selectively learning from humans. To contribute to O4, this work suggested one such indicator, namely the rate at which objects that are known to the robot become invisible in demonstrations, and demonstrated its correlation with one of the metrics defining the quality of recorded demonstrations. More indicators of higher robot awareness in human teachers for the robots may be discovered further in future research.

The experiment presented in this chapter also contributed to addressing O1 by presenting results suggesting that non-expert participants can provide visual demonstrations of physical tasks faster after teaching the robot multiple times, even though their actions in the provided demonstrations were detected by the robot with a higher error rate. A part of O2 was also addressed with non-expert humans teaching a robot that was equipped with a PLI system in this experiment. The achieved robot learning outcome in such a scenario highlights that the PLI system can be used to allow robots to learn simple, everyday tasks efficiently.

Chapter 6

Conclusion

The research presented in this thesis aimed at enhancing aspects of social learning from non-expert human teachers in humanoid robots by contributing to overcoming challenges and addressing research gaps. The findings contributed to the understanding of how non-expert humans teach robots without any supervision or prior training, how robots can acquire knowledge about the structure of goal-oriented tasks, and what factors in human teachers lead to better social learning outcomes in robots, so that robots can potentially use such factors to identify the most suitable human teachers for their social learning.

This chapter concludes my thesis by summarizing the key findings related to each research objective (Section 6.1), highlighting the contributions of the work (Section 6.2), and describing the overall limitations of the research (Section 6.3).

6.1 Summary of Findings

The following discusses how the research objectives defined in Section 1.3 were addressed based on the obtained results:

O1 Understand how the behaviours of non-expert humans and/or characteristics of their demonstrations change when they teach tasks to a humanoid robot multiple times

The experiment conducted as Phase 1 of my research, presented in Chapter 3, investigated changes in the behaviour of human teachers for robots and studied certain characteristics of their demonstration over time, when providing 20 kinesthetic

demonstrations of basic physical tasks to a Pepper humanoid robot, without any external support and across five experimental sessions. The results indicated diversity in the kinesthetic teaching style of non-experts during repeated interactions. Three clusters of *high-achievers*, *improving*, and *low-achievers* were identified based on participants' success in providing demonstrations throughout the sessions. The key findings considering these clusters were:

- High-achievers (9 out of 28 participants) demonstrated the tasks with high success from the first session, as indicated by large overall areas of the tasks covered by their kinesthetic demonstrations, and they remained consistently successful. Improving participants (15 out of 28 participants) had initially lower teaching success that later improved, especially over the first three sessions. Low-achievers (only 4 out of 28 participants) provided less successful task demonstrations without any noticeable improvement over sessions.
- Participants spent significantly less time teaching the robot as they gained robot teaching experience throughout the sessions.
- Participants' gaze fixations on the robot's hand significantly decreased as they became more experienced robot teachers within the same session and over multiple sessions.
- Improving participants were more likely to use both hands for grabbing the robot's arm in their kinesthetic teaching and more often guided the robot's upper arm as they gained experience over the sessions and within the sessions. These observations may help explain why this cluster of participants showed improvement.
- Perceived effort significantly decreased with repeated kinesthetic teaching experience in multiple sessions.

Overall, the results suggested that multiple exposures to kinesthetic teaching of a robot led to providing more successful kinesthetic demonstrations. The experiment conducted later in Phase 3 also contributed to the understanding of changes in the characteristics of demonstrations provided by non-experts over time, by focusing on observation learning of an iCub robot. In Experiment 2, participants provided 15 visual demonstrations of abstract physical tasks, in a single session, with and without external resources to support their teaching. It was found that:

- Participants spent significantly less time teaching each task to the robot as they proceeded in the session.

– Participants also demonstrated each sub-goal in the tasks faster when they proceeded in the session.
– Overall, the quality of the recorded demonstrations and the quality of the generated task programs (indicated by metrics such as sub-goal detection error rate, program generation score, etc.) decreased as participants taught more tasks to the robot.

In these two experiments, it was observed that while in a certain context (i.e., kinesthetic teaching of how to clean/clear surfaces in multiple days), the majority of non-experts could become more successful teachers for a robot only by gaining more experience of teaching a robot, in another context (i.e., demonstrating basic household tasks in an abstract format in a single day), participants' teaching success decreased with time. When comparing these two cases, it should be taken into consideration that kinesthetic teaching required a new skill set in non-expert teachers that involves moving the arm of a humanoid robot and physically putting the robot through certain actions, which they had never done before, but was developed with gaining experience. On the other hand, providing visual demonstrations of household tasks primarily required non-expert teachers to carry out the tasks themselves in front of the robot[16]. Factors such as fatigue and boredom when repeatedly demonstrating tasks (three tasks, five times each, in a single session) might have played a role in Experiment 2 in decreasing the effectiveness of humans teaching a robot over time.

O2 Design, implement, and evaluate a learning framework for robots to allow them to understand the sequential structure of multi-step tasks and efficiently perform such tasks

In Phase 2 of my research (presented in Chapter 4), I designed a novel LfD framework for robots by taking inspiration from a certain type of social learning in primates, i.e., Program-Level Imitation (PLI). This approach enables robots to understand and perform complex actions by using high-level programs that incorporate sequential regularities between sub-goals. The PLI system processes human-provided demonstrations to discover multiple possible arrangements of sub-goals that achieve a task goal. Then, for performing a learned task, the robot evaluates available sequences through mental simulation based on user-defined criteria (e.g., speed, safety, etc.) to

[16]Later, when addressing O3 and O4, it will be highlight that improved robot learning outcomes can be achieved when a human teacher goes beyond only performing the tasks in front of a robot and pays attention to capabilities and limitations of the robot while performing actions. This awareness, however, may not naturally occur in non-experts, and therefore, human teachers of a robot may view providing a task demonstration to the robot as similar to performing the tasks for themselves.

find the optimal sequence, which is then executed using the hierarchical structure of actions embedded in the program.

I implemented the proposed system on an iCub humanoid robot to enable learning abstract physical tasks by observing visual demonstrations. The implementation process involved developing specific functionalities for each module of the PLI system, including programming visuomotor capabilities in the robot for detecting and performing three primitives of grasping objects and putting them inside boxes, pressing buttons, and sliding boxes from one place to another.

To conclude Phase 2, I evaluated the proposed PLI system by performing a number of tests in real-world and in simulation, which included learning two abstract tasks (i.e., boiling potatoes and organizing a space). Through some example scenarios when I took the role of a teacher for the iCub robot, I illustrated how the system could accommodate variations in human teaching styles and use inconsistencies in demonstrations to discover more ways of performing tasks and improving its performance; and then select optimal sequences based on different user preferences (e.g., prioritizing task completion time, safety, etc.). By conducting a series of simulated tests, I reported how adjusting system parameters could improve the learning performance of the robot when demonstrations contained errors (e.g., missing sub-goals, wrong order, incorrect target information, and incomplete demonstrations). Lastly, I presented how a robot with the proposed PLI system can achieve a task goal from varying initial task states, cope with interruptions and errors during execution by noticing changes in the task state and re-achieving necessary sub-goals based on learned relations, and skipping unnecessary actions.

Results from the experiment conducted in Phase 3 also contributed to this research objective by further evaluating the effectiveness of observational learning with the PLI system by involving non-expert participants to teach a robot physical tasks. Teaching the iCub robot via visual demonstrations using the PLI framework was positively evaluated by non-expert users based on the provided comments, highlighting the ease and enjoyment of teaching. The system exhibited high reliability, failing to generate a program in less than 2% of task demonstrations across all three tasks (i.e., abstract tasks used in Phase 2, plus a waste sorting task). The system generalized action sequences in many cases, where the generated programs included valid ways of performing the tasks that the robot did not explicitly observe in the demonstrations.

O3 Study how giving humans extra resources for understanding robot limitations and learning strategies when they teach a robot may impact robot learning and human-robot teaching

With the ultimate goal of answering the question of 'who to imitate' in imitation learning of robots, I proposed to use 'awareness of and attention to a learner robot's capabilities and constraints' as a factor in human teachers that might indicate more suitable teachers for robots. An experiment was conducted in Phase 3 of my research to study whether or not giving some groups of non-expert participants extra resources for understanding robot limitations and learning strategies, to increase their awareness of and attention to the robot's capabilities and limitations, impacts robot learning. While participants taught three abstract physical tasks to an iCub humanoid robot that was equipped with the proposed PLI framework, a screen next to the robot displayed its real-time Robot's Camera View (RCV) for some participants to help them understanding robot's visual limitations, and/or some participants completed a virtual Goal-Highlighting Activity (GHA) that put them in the position of a robot when recognizing and labelling observed task steps to help them understanding robot learning strategies related to learning only from specific primitive actions. The main findings in this experiment were:

- In line with the initial expectations, participants with access to the RCV reported higher attention to ensuring object visibility in their demonstration. However, neither of the two extra teaching resources was found to impact participants' self-reported attention to performing specific primitive actions that the robot could learn effectively from.
- Considering the quality of the recorded demonstrations, issues with what participants performed in demonstrations occurred less frequently when participants completed the GHAs, especially when they did not have access to the RCV. This finding suggests GHAs may have increased awareness of participants about the system's capabilities and limitations to more carefully perform their actions. Moreover, the error rates in detecting sub-goals in demonstrations by the robot were significantly reduced with participants' access to the RCV.
- In terms of robot learning outcomes when learning from the provided demonstrations, higher program generation scores (a metric that was defined to quantify the performance of robot learning with the PLI system) were achieved when learning from participants who had access to the RCV.
- Subjective ratings of the robot's competence were significantly higher when participants had access to the RCV.
- Participants with access to the RCV perceived significantly less effort and less frustration when teaching the robot.

Overall, the findings supported that the awareness of human teachers to the capabilities and constraints of a learner robot, especially to the point that visual demonstrations must be visible within the field of view of a robot, contributes to better robot learning outcomes.

O4 Detect patterns in non-experts' teaching that a robot can use to identify the most suitable teachers

While discussing results from Experiments 1 and 2 to address O1 and O3, some factors were highlighted in human teachers that led to robots learning tasks more effectively. In particular, the amount of experience of non-expert people with the kinesthetic teaching of a robot and their awareness of and attention to the visual limitations of a robot were found to improve aspects of robot learning (e.g., quality and success of demonstrations) in kinesthetic and observational robot LfD scenarios, respectively. Still, to be able to use them as an indicator to automatically identify more suitable human teachers to learn from, a robot will need to estimate such factors in human teachers from its own perspective.

For addressing this research objective, I looked for patterns in non-experts' teaching robots in the two experiments that could potentially help a robot identify the most suitable teachers. In Experiment 1, successful robot teachers tended to use both hands more and guide the robot's upper arm; while low-achievers used both hands approximately three times less than other clusters and rarely guided the robot's upper arm in their kinesthetic instructions. While these metrics were not captured by the robot itself in Experiment 1, a more advanced robot equipped with touch-sensitive skin on its arm could possibly detect the placement of the teacher's hand on its arm and estimate the metrics noted above to identify more suitable teachers.

More importantly, in Experiment 2, the data indicated a strong positive correlation between a metric called the normalized task objects disappearance rate and the normalized sub-goal detection error rates in the provided demonstrations. Normalized task objects disappearance rates were estimated using the data captured by the robot itself by tracking the frequency at which known items in the tasks became invisible to the robot. This metric can be a potential indicator that a robot that learns from visual observation can track in human teachers to estimate the potential for successful social learning.

6.2 Contributions to Original Knowledge

This thesis makes the following contribution to the original knowledge in the fields of robot learning and Human-Robot Interaction (HRI):

- **Studying the behaviour of non-expert humans when teaching physical tasks to a humanoid robot**

  The results of Experiment 1 contributed to the understanding of the behavioural characteristics of non-expert, inexperienced human teachers for robots, including changes in their gaze patterns and robot arm manipulation strategies over repeated kinesthetic teaching interactions. More importantly, the results demonstrated that the majority of non-experts can adapt to kinesthetic teaching of a humanoid robot and improve their teaching success and efficiency in an unsupervised setting, simply by repeatedly engaging in teaching tasks. These findings on human teaching behaviours are important because in many real-world applications of robots, users may find themselves in a situation where the only option for learning how to instruct a robot is simply exploring robot teaching on their own, without additional instructional tools.

- **Designing, implementing, and evaluating a novel, biologically-inspired robot learning architecture**

  I designed, implemented, and tested a novel, biologically inspired LfD architecture for robots, enabling them to learn how to execute complex tasks by using the hierarchical structure of sub-goals rather than merely copying observed actions. A significant contribution of this system is its ability to process diverse human-provided demonstrations, even those containing errors or inconsistencies, to identify multiple valid sequences of actions that can result in achieving a task's goals, to enhance flexibility and efficiency in task execution and adaptation to varying contexts. This ability in robots is important because, without adapting to new environments and the changing needs of different individuals, a robot cannot fulfill its purpose in the long term as a multi-functional tool. Moreover, because this system is biologically inspired and operates using intuitive logic that is employed by primates and humans in processing information to learn hierarchical actions, I anticipate that the proposed system will potentially facilitate non-expert users to understand the learner robots.

- **Exploring the role of human teachers' awareness of a robot's visual constraints and learning strategies in human-robot teaching outcomes**

The results obtained in Experiment 2 contributed to the understanding of factors in human teachers for robots that lead to better learning outcomes in robots. It was systematically investigated that the attention of human teachers to the capabilities and constraints of a learner robot, in particular, to the limited ability of a robot to observe visual demonstrations, leads to improved human-robot teaching outcomes. An innovation in this experiment was offering the visual perspective of a learner robot to non-expert human teachers, to increase their awareness of and attention to the limitations when the robot visually observed the demonstrations. While this was intended as an experimental manipulation, designers of robots that learn from visual inputs may consider using a screen to allow VPT in human teachers when demonstrating physical tasks to a robot, which was shown to improve the quality of demonstrations and robot learning.

- **Proposing metrics for robots to identify the most suitable human teachers**

  My research identified quantifiable patterns to be used by robots to assess the suitability of non-expert human teachers to learn from them. This work presented in this thesis is a first step towards answering the complex question of 'who to imitate' from a robot's perspective, contributing to the long-term goal of enabling future advanced robots to identify and select suitable human teachers for more effective social learning.

The findings of this thesis have been communicated to the research community via a number of journal articles and conference papers [16, 17, 18, 19, 20].

6.3 Overall Limitations and Future Work

Throughout this thesis, the individual chapters (Section 3.5, 3.6, 4.6, 4.7, 5.5, and 5.6) discussed specific limitations and future work of each study. Here, I point out the overall limitations of the work, which can be addressed through further research in the following directions in the future:

1. **Expanding the type of tasks involved:** The topics related to human-robot teaching were studied by considering the learning of only a limited set of tasks by robots. In particular, one type of task was considered in Experiment 1: wiping surfaces during cleaning activities. While both vertical and horizontal surfaces were included, and the difficulty of the tasks and types of particles that needed to be cleaned were varied

in the experiment, different types of tasks may produce different results. Experiment 2 also included only learning of three basic household tasks, in abstract form, in the domains of cooking and organizing. Studies in the future may explore social learning of other tasks by robots, with a wider range of motions and a broader set of actions, including robot navigation in addition to object manipulation, as needed to be carried out in realistic settings. Such tasks may even go beyond the domain of performing every household activity and include duties in healthcare (e.g., delivering medication to hospital patients), industry (e.g., helping with assembling products), etc.

2. **Improving diversity in participated human teachers:** Recruiting participants for HRI studies, especially for multiple-session experiments such as Experiment 1, is challenging. However, the presented studies could benefit from the participation of the general public rather than recruiting university students. University students, in particular those in a technical university such as the University of Waterloo, might have been positively biased toward using robots and are generally more familiar with new technologies. Moreover, as previously pointed out, most of the participants in the studies reported in this thesis were of a certain age range (i.e., between the ages of 18 and 30). This might have contributed to the result that only a few participants were in the low-achievers cluster in Experiment 1 and the observation learning of iCub using the PLI framework was successful. If the same studies had been conducted on a different population, the composition of clusters in Experiment 1 and the success of robot learning in Experiment 2 could have been different. The inclusion of people other than university students, e.g., older adults, could address this issue and more accurately reveal the possible effects of background and age differences on the behaviours of human teachers for robots, characteristics of their demonstrations, and the effectiveness of social learning in robots.

3. **Evaluating the PLI system in broader contexts**: In this work, the effectiveness of robot LfD using the proposed PLI approach was evaluated in simulation and in an experiment involving a limited population, as noted earlier. The effectiveness of the present PLI approach should be tested in a broader range, including a variety of human-robot interaction scenarios. Having a wider range of non-expert robot users demonstrate different tasks to teach the robot will help better identify the strengths and weaknesses of the system, its effectiveness in addressing users' preferences, challenges users may face, etc., and it is crucial to the future development of the system. The performance and effectiveness of the present approach should be quantitatively compared to other state-of-the-art robot learning techniques, e.g., LLM-based sys-

tems [12], from the perspective of non-expert robot users.

4. **Experimenting in the real world:** The reported experiments were performed only in laboratory settings. Future work should consider conducting human-robot teaching studies in real environments, for example, a home setting. In such settings, many other factors are present which may hinder the ability of robots to learn from humans and perform tasks. For instance, environmental noises such as changes in the lighting condition and the loudness of the surroundings may limit the sensing ability of a robot, which can impose more constraints and limitations for a robot that learns from examples. Human teachers' awareness of and attention to these new robot limitations can be considered in future studies to follow up on Experiment 2.

5. **Enhancing other aspects in social learning of robots:** The focus of the present thesis was to improve the learning ability of robots from non-expert humans by focusing on three main challenges noted in Section 1.2. There are still many other aspects of social learning of robots that can be improved in the future. To begin with, considering the proposed PLI system, more demonstration modalities could be included and tests within the framework (e.g., teaching using voice instructions, as was also suggested by participants in Experiment 2). Future implementations of the PLI architecture on an iCub robot or another robot can benefit from improved robot accuracy in performing actions and detecting sub-goals in demonstrations. In a more general context, future studies can focus on explainability in social learning of robots, which refers to the ability of AI systems to provide clear and understandable explanations about their actions and decisions. For adaptable social robots, providing multi-modal feedback to humans becomes highly important both when learning new tasks and when executing learned tasks. In the reported studies, the robot provided minimal feedback to human teachers, mostly via its task performance. Social learning of robots could be improved by facilitating access to the robot's internal state within the learning system. For example, a new module in the PLI framework could be developed to automatically process and translate this information into understandable verbal explanations, to fully realize the potential of social learning using a biologically inspired approach.

References

[1] Cartesian control API - Aldebaran 2.1.4.13 documentation. http://doc.aldebaran.com/2-1/naoqi/motion/control-cartesian-api.html, Accessed: Sept. 15, 2022.

[2] Choregraphe - NAO Software 1.14.5 documentation. http://doc.aldebaran.com/1-14/software/choregraphe/choregraphe_overview.html, Accessed: April 15, 2023.

[3] iCub Forward Kinematics. https://icub-tech-iit.github.io/documentation/icub_kinematics/icub-forward-kinematics/icub-forward-kinematics/, Accessed: Jan. 5, 2025.

[4] Mountain Gorilla of DRC stock photo. https://www.istockphoto.com/photo/mountain-gorilla-of-drc-gm1191746980-338346445, Accessed: April 21, 2025.

[5] Pepper the humanoid and programmable robot. https://www.aldebaran.com/en/pepper, Accessed: Sept. 15, 2022.

[6] Python Wrapper for Intel Realsense SDK 2.0. https://pypi.org/project/pyrealsense2/, Accessed: June. 15, 2023.

[7] Qualtrics experience management software. https://www.qualtrics.com/, Accessed: Sept. 10, 2024.

[8] The CSRT tracker in OpenCV. https://docs.opencv.org/4.6.0/d2/da2/classcv_1_1TrackerCSRT.html#details, Accessed: April 10, 2023.

[9] Tobii Pro Glasses 2 wearable eye tracker. https://www.tobiipro.com/product-listing/tobii-pro-glasses-2/, Accessed: March 15, 2023.

[10] Shahar Aberbach-Goodman and Roy Mukamel. Temporal hierarchy of observed goal-directed actions. *Scientific Reports*, 13, 2023.

[11] Mouad Abrini, Malika Auvray, and Mohamed Chetouani. Humans' Spatial Perspective-Taking When Interacting with a Robotic Arm. In *Proceedings of the 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN '23)*, pages 533–540. IEEE Press, 2023.

[12] Michael Ahn, Anthony Brohan, Noah Brown, Yevgen Chebotar, Omar Cortes, Byron David, et al. Do As I Can and Not As I Say: Grounding Language in Robotic Affordances. In *arXiv preprint arXiv:2204.01691*, 2022.

[13] Gopika Ajaykumar, Maia Stiber, and Chien-Ming Huang. Designing user-centric programming aids for kinesthetic teaching of collaborative robots. *Robotics and Autonomous Systems*, 145:103845, 2021.

[14] Pourya Aliasghari, Moojan Ghafurian, Chrystopher L. Nehaniv, and Kerstin Dautenhahn. Effects of Gaze and Arm Motion Kinesics on a Humanoid's Perceived Confidence, Eagerness to Learn, and Attention to the Task in a Teaching Scenario. In *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction (HRI '21)*, page 197–206. ACM, 2021.

[15] Pourya Aliasghari, Moojan Ghafurian, Chrystopher L. Nehaniv, and Kerstin Dautenhahn. Impact of nonverbal robot behaviour on human teachers' perceptions of a learner robot. *Interaction Studies*, 22(2):141–176, 2022.

[16] Pourya Aliasghari, Moojan Ghafurian, Chrystopher L. Nehaniv, and Kerstin Dautenhahn. Kinesthetic teaching of a robot over multiple sessions: Impacts on speed and success. In *Proceedings of the 14th International Conference of Social Robotics (ICSR '22)*, pages 160–170. Springer, 2023.

[17] Pourya Aliasghari, Moojan Ghafurian, Chrystopher L. Nehaniv, and Kerstin Dautenhahn. A Biologically Inspired Program-level Imitation Approach for Robots: Proof-of-Concept. In *Proceedings of the 33rd IEEE International Conference on Robot and Human Interactive Communication (ROMAN '24)*, pages 755–762. IEEE Press, 2024.

[18] Pourya Aliasghari, Moojan Ghafurian, Chrystopher L. Nehaniv, and Kerstin Dautenhahn. How Non-experts Kinesthetically Teach a Robot over Multiple Sessions: Diversity in Teaching Styles and Effects on Performance. *International Journal of Social Robotics*, 16(11):2079–2105, 2024. (Reproduced with permission from Springer Nature).

[19] Pourya Aliasghari, Moojan Ghafurian, Chrystopher L. Nehaniv, and Kerstin Dautenhahn. A Biologically Inspired Program-Level Imitation Approach for Robots. *IEEE Access*, 13:89481–89501, 2025. (© 2025 IEEE. Reprinted, with permission).

[20] Pourya Aliasghari, Chrystopher L. Nehaniv, Moojan Ghafurian, and Kerstin Dautenhahn. Improving Robot Learning Outcomes in Human-Robot Teaching: The Role of Human Teachers' Awareness of a Robot's Vision Constraints. In *Proceedings of the 34rd IEEE International Conference on Robot and Human Interactive Communication (ROMAN '25)*. IEEE Press, 2025. (© 2025 IEEE. Reprinted, with permission).

[21] Gerard F. Anderson and Peter Sotir Hussey. Population Aging: A Comparison Among Industrialized Countries. *Health Affairs*, 19(3):191–203, 2000.

[22] Brenna D. Argall, Sonia Chernova, Manuela Veloso, and Brett Browning. A survey of robot learning from demonstration. *Robotics and Autonomous Systems*, 57(5):469–483, 2009.

[23] Samy A. Azer. The Qualities of a Good Teacher: How can they be Acquired and Sustained? *Journal of the Royal Society of Medicine*, 98(2):67–69, 2005.

[24] Douglas Bates, Martin Mächler, Benjamin M. Bolker, and Steven C. Walker. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 2015.

[25] Tony Belpaeme, James Kennedy, Aditi Ramachandran, Brian Scassellati, and Fumihide Tanaka. Social robots for education: A review. *Science Robotics*, 3(21):5954, 2018.

[26] Muhammad Bilal, Nir Lipovetzky, Denny Oetomo, and Wafa Johal. Beyond Success: Quantifying Demonstration Quality in Learning from Demonstration. In *Proceedings of the 2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '24)*, pages 5120–5127. IEEE Press, 2024.

[27] Aude Billard, Sylvain Calinon, Rüdiger Dillmann, and Stefan Schaal. Robot Programming by Demonstration. In Bruno Siciliano and Oussama Khatib, editors, *Springer Handbook of Robotics*, pages 1371–1394. Springer, 2008.

[28] Susan A.J. Birch, Nazanin Akmal, and Kristen L. Frampton. Two-year-olds are vigilant of others' non-verbal cues to credibility. *Developmental Science*, 13(2):363–369, 2010.

[29] Hamparsum Bozdogan. Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions. *Psychometrika*, 52(3):345–370, 1987.

[30] Gary Bradski. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, 2000.

[31] Stein Bråten. *Intersubjective communication and emotion in early ontogeny.* Studies in emotion and social interaction, 2nd series. Cambridge University Press, 1998.

[32] Virginia Braun and Victoria Clarke. Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2):77–101, 2006.

[33] Cynthia Breazeal. Role of expressive behaviour for robots that learn from people. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1535):3527–3538, 2009.

[34] Cynthia Breazeal. Social robots for health applications. In *Proceedings of the 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 5368–5371. IEEE Press, 2011.

[35] Cynthia Breazeal, Matt Berlin, Andrew Brooks, Jesse Gray, and Andrea L. Thomaz. Using perspective taking to learn from ambiguous demonstrations. *Robotics and Autonomous Systems*, 54(5):385–393, 2006.

[36] Cynthia Breazeal and Brian Scassellati. Robots that imitate humans. *Trends in Cognitive Sciences*, 6(11):481–487, 2002.

[37] Diana Burkart and Barbara Bruno. Human-like Social Learning for Social Robots: A Systematic Review. In *Proceedings of the 33rd IEEE International Conference on Robot and Human Interactive Communication (ROMAN '24)*. IEEE Press, 2024.

[38] Richard. W. Byrne. The evolution of intelligence. In Peter J. B. Slater and Timothy R. Halliday, editors, *Behaviour and Evolution*, pages 223–265. Cambridge University Press, 1994.

[39] Richard W. Byrne. *The Thinking Ape: Evolutionary Origins of Intelligence.* Oxford University Press, 1995.

[40] Richard W. Byrne. Imitation of novel complex actions: What does the evidence from animals mean? *Advances in the Study of Behavior*, 31:77–105, 2002.

[41] Richard W. Byrne. Imitation as behaviour parsing. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 358(1431):529–536, 2003.

[42] Richard W. Byrne and Anne E. Russon. Learning by imitation: A hierarchical approach. *Behavioral and Brain Sciences*, 21(5):667–684, 1998.

[43] Maya Cakmak and Leila Takayama. Teaching People How to Teach Robots: The Effect of Instructional Materials and Dialog Design. In *Proceedings of the 2014 ACM/IEEE International Conference on Human-Robot Interaction (HRI '14)*, page 431–438. IEEE, 2014.

[44] Maya Cakmak and Andrea L. Thomaz. Eliciting good teaching from humans for machine learners. *Artificial Intelligence*, 217:198–215, 2014.

[45] Sylvain Calinon and Aude Billard. Learning of gestures by imitation in a humanoid robot. In Chrystopher L. Nehaniv and Kerstin Dautenhahn, editors, *Imitation and Social Learning in Robots, Humans and Animals: Behavioural, Social and Communicative Dimensions*, page 153–178. Cambridge University Press, 2007.

[46] Sylvain Calinon and Aude G. Billard. What is the teacher's role in robot programming by demonstration?: Toward benchmarks for improved learning. *Interaction Studies*, 8(3):441–464, 2007.

[47] Sylvain Calinon, Florent Guenter, and Aude Billard. Goal-Directed Imitation in a Humanoid Robot. In *Proceedings of the 2005 IEEE International Conference on Robotics and Automation (ICRA '05)*, pages 299–304. IEEE Press, 2005.

[48] Timothy Chen, Ola Shorinwa, Joseph Bruno, Aiden Swann, Javier Yu, Weijia Zeng, Keiko Nagami, Philip Dames, and Mac Schwager. Splat-Nav: Safe Real-Time Robot Navigation in Gaussian Splatting Maps. *IEEE Transactions on Robotics*, 41:2765–2784, 2025.

[49] Eugenio Chisari, Nick Heppert, Max Argus, Tim Welschehold, Thomas Brox, and Abhinav Valada. Learning Robotic Manipulation Policies from Point Clouds with Conditional Flow Matching. *arXiv preprint arXiv:2409.07343*, 2024.

[50] Sumin Cho and Sungho Jo. Incremental Online Learning of Robot Behaviors From Selected Multiple Kinesthetic Teaching Trials. *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans*, 43(3):730–740, 2013.

[51] Gergely Csibra and György Gergely. Social learning and social cognition: The case for pedagogy. In Yuko Munakata and Mark H. Johnson, editors, *Processes of Change in Brain and Cognitive Development: Attention and Performance XXI*, volume 21, pages 249–274. Oxford University Press, 2006.

[52] Joel Currie, Katrina Louise Mcdonough, Agnieszka Wykowska, Maria Elena Giannaccini, and Patric Bach. More Than Meets the Eye? An Experimental Design to Test Robot Visual Perspective-Taking Facilitators Beyond Mere-Appearance. In *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction (HRI '24)*, page 359–363. ACM, 2024.

[53] Deborah Custance, Andrew Whiten, and Tamar Fredman. Social learning of an artificial fruit task in capuchin monkeys (*Cebus apella*). *Journal of Comparative Psychology*, 113(1):13, 1999.

[54] Kerstin Dautenhahn. The art of designing socially intelligent agents: science, fiction, and the human in the loop. *Applied Artificial Intelligence*, 12(7-8):573–617, 1998.

[55] Kerstin Dautenhahn. Socially intelligent robots: dimensions of human–robot interaction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1480):679–704, 2007.

[56] Kerstin Dautenhahn and Chrystopher L. Nehaniv. The Agent-Based Perspective on Imitation. In Kerstin Dautenhahn and Chrystopher L. Nehaniv, editors, *Imitation in Animals and Artifacts*, page 1–40. MIT Press, 2002.

[57] Yiannis Demiris and Matthew Johnson. Distributed, predictive perception of actions: a biologically inspired robotics architecture for imitation and learning. *Connection Science*, 15(4):231–243, 2003.

[58] Yiannis Demiris and Bassam Khadhouri. Hierarchical attentive multiple models for execution and recognition of actions. *Robotics and Autonomous Systems*, 54(5):361–369, 2006.

[59] Munjal Desai, Kristen Stubbs, Aaron Steinfeld, and Holly Yanco. Creating trustworthy robots: Lessons and inspirations from automated systems. In *Proceedings of the AISB '09 Convention: New Frontiers in Human-Robot Interaction*, 2009.

[60] Thomas Eiband, Johanna Liebl, Christoph Willibald, and Dongheui Lee. Online task segmentation by merging symbolic and data-driven skill recognition during kinesthetic teaching. *Robotics and Autonomous Systems*, 162:104367, 2023.

[61] Staffan Ekvall and Danica Kragic. Robot Learning from Demonstration: A Task-level Planning Approach. *International Journal of Advanced Robotic Systems*, 5(3):33, 2008.

[62] Aviv Elor, Sri Kurniawan, and Leila Takayama. Human Experiences in Teaching Robots: Understanding Agent Expressivity and Learning Effects through a Virtual Robot Arm. In *Proceedings of the 2022 IEEE International Conference on Smart Computing (SMARTCOMP)*, pages 133–141. IEEE, 2022.

[63] Nathan J. Emery. The eyes have it: the neuroethology, function and evolution of social gaze. *Neuroscience and Biobehavioral Reviews*, 24:581–604, 2000.

[64] Oriel FeldmanHall and Matthew R. Nassar. The computational challenge of social learning. *Trends in Cognitive Sciences*, 25(12):1045–1057, 2021.

[65] Julia Fink, Valérie Bauwens, Frederic Kaplan, and Pierre Dillenbourg. Living with a Vacuum Cleaning Robot. *International Journal of Social Robotics*, 5:389–408, 2013.

[66] Roya Firoozi, Johnathan Tucker, Stephen Tian, Anirudha Majumdar, Jiankai Sun, Weiyu Liu, Yuke Zhu, Shuran Song, Ashish Kapoor, Karol Hausman, Brian Ichter, Danny Driess, Jiajun Wu, Cewu Lu, and Mac Schwager. Foundation models in robotics: Applications, challenges, and the future. *The International Journal of Robotics Research*, 44(5):701–739, 2025.

[67] Kerstin Fischer, Franziska Kirstein, Lars Christian Jensen, Norbert Krüger, Kamil Kukliński, Maria Vanessa aus der Wieschen, and Thiusius Rajeeth Savarimuthu. A Comparison of Types of Robot Control for Programming by Demonstration. In *Proceedings of the 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI '16)*, pages 213–220. IEEE Press, 2016.

[68] Tobias Fischer and Yiannis Demiris. Markerless perspective taking for humanoid robots in unconstrained environments. In *Proceedings of the 2016 IEEE International Conference on Robotics and Automation (ICRA '16)*, pages 3309–3316. IEEE Press, 2016.

[69] Tobias Fischer and Yiannis Demiris. Computational Modeling of Embodied Visual Perspective Taking. *IEEE Transactions on Cognitive and Developmental Systems*, 12(4):723–732, 2020.

[70] Terrence Fong, Illah Nourbakhsh, and Kerstin Dautenhahn. A survey of socially interactive robots. *Robotics and Autonomous Systems*, 42(3-4):143–166, 2003.

[71] Livia Freier, Richard P. Cooper, and Denis Mareschal. The planning and execution of natural sequential actions in the preschool years. *Cognition*, 144:58–66, 2015.

[72] Olivier Friard and Marco Gamba. BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in ecology and evolution*, 7(11):1325–1330, 2016.

[73] Bennett G. Galef. Social learning and traditions in animals: evidence, definitions, and relationship to human culture. *Wiley Interdisciplinary Reviews: Cognitive Science*, 3(6):581–592, 2012.

[74] Sergio Garrido-Jurado, Rafael Muñoz-Salinas, Francisco J. Madrid-Cuevas, and Rafael Medina-Carnicer. Generation of fiducial marker dictionaries using Mixed Integer Linear Programming. *Pattern Recognition*, 51:481–491, 2016.

[75] Nadine Glas and Catherine Pelachaud. Definitions of engagement in human-agent interaction. In *Proceedings of the 2015 International Conference on Affective Computing and Intelligent Interaction (ACII '15)*, pages 944–949. IEEE Press, 2015.

[76] Wonjoon Goo and Scott Niekum. One-Shot Learning of Multi-Step Tasks from Observation via Activity Localization in Auxiliary Video. In *Proceedings of the 2019 International Conference on Robotics and Automation (ICRA '19)*, pages 7755–7761. IEEE Press, 2019.

[77] Nakul Gopalan, Nina Moorman, Manisha Natarajan, and Matthew Gombolay. Negative Result for Learning from Demonstration: Challenges for End-Users Teaching Robots with Task And Motion Planning Abstractions. In *Proceedings of the RSS 2023 Workshop on Learning for Task and Motion Planning*, 2023.

[78] Scott T. Grafton and Antonia F. de C. Hamilton. Evidence for a distributed hierarchy of action representation in the brain. *Human Movement Science*, 26(4):590–616, 2007.

[79] Daniel H. Grollman and Aude Billard. Donut as I do: Learning from failed demonstrations. In *Proceedings of the 2011 IEEE International Conference on Robotics and Automation (ICRA '11)*, pages 3804–3809. IEEE Press, 2011.

[80] Daniel H. Grollman and Odest Chadwicke Jenkins. Incremental learning of subtasks from unsegmented demonstration. In *Proceedings of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '10)*, pages 261–266. IEEE Press, 2010.

[81] Hyowon Gweon. Inferential social learning: cognitive foundations of human social learning and teaching. *Trends in Cognitive Sciences*, 25(10):896–910, 2021.

[82] Antonia F.de C. Hamilton, Rachel Brindley, and Uta Frith. Visual perspective taking impairment in children with autistic spectrum disorder. *Cognition*, 113(1):37–44, 10 2009.

[83] Paul L. Harris and Kathleen H. Corriveau. Young children's selective trust in informants. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1567):1179–1187, 2011.

[84] Sandra G. Hart. Nasa-Task Load Index (NASA-TLX); 20 Years Later. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 50, page 904–908. Sage Publications, 2006.

[85] Sandra G. Hart and Lowell E. Staveland. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Advances in Psychology*, volume 52, pages 139–183. Elsevier, 1988.

[86] Meghan L. Healey and Murray Grossman. Cognitive and Affective Perspective-Taking: Evidence for Shared and Dissociable Anatomical Substrates. *Frontiers in Neurology*, 9, 2018.

[87] Erin Hedlund, Michael Johnson, and Matthew Gombolay. The Effects of a Robot's Performance on Human Teachers for Learning from Demonstration Tasks. In *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction (HRI '21)*, page 207–215. ACM, 2021.

[88] Joseph Henrich and James Broesch. On the nature of cultural transmission networks: Evidence from Fijian villages for adaptive learning biases. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1567):1139–1148, 2011.

[89] Cecilia M. Heyes. Social learning in animals: Categories and mechanisms. *Biological Reviews of the Cambridge Philosophical Society*, 69(2):207–231, 1994.

[90] Heather C. Hill, Deborah Loewenberg Ball, and Stephen G. Schilling. Unpacking Pedagogical Content Knowledge: Conceptualizing and Measuring Teachers' Topic-Specific Knowledge of Students. *Journal for Research in Mathematics Education*, 39(4):372–400, 2008.

[91] Heather C. Hill and Mark Chin. Connections Between Teachers' Knowledge of Students, Instruction, and Achievement Outcomes. *American Educational Research Journal*, 55(5):1076–1112, 2018.

[92] Stefanie Hoehl, Stefanie Keupp, Hanna Schleihauf, Nicola McGuigan, David Buttelmann, and Andrew Whiten. 'Over-imitation': A review and appraisal of a decade of research. *Developmental Review*, 51:90–108, 2019.

[93] William Hoppitt and Kevin N. Laland. Social processes influencing learning in animals: a review of the evidence. *Advances in the Study of Behavior*, 38:105–165, 2008.

[94] William Hoppitt and Kevin N. Laland. *Social Learning: An Introduction to Mechanisms, Methods, and Models*. Princeton University Press, 2013.

[95] Takayuki Hoshi and Hiroyuki Shinoda. Robot skin based on touch-area-sensitive tactile element. In *Proceedings of the 2006 IEEE International Conference on Robotics and Automation (ICRA '06)*, pages 3463–3468. IEEE, 2006.

[96] Sandy H. Huang, Isabella Huang, Ravi Pandya, and Anca D. Dragan. Nonverbal Robot Feedback for Human Teachers. In *Proceedings of the 3rd Conference on Robot Learning (CoRL '19)*, pages 1038–1051. PMLR, 2019.

[97] Wenlong Huang, Pieter Abbeel, Deepak Pathak, and Igor Mordatch. Language models as zero-shot planners: Extracting actionable knowledge for embodied agents. In *Proceedings of the 39th International Conference on Machine Learning*, pages 9118–9147. PMLR, 2022.

[98] Masato Ito, Kuniaki Noda, Yukiko Hoshino, and Jun Tani. Dynamic and interactive generation of object handling behaviors by a small humanoid robot using a dynamic neural network model. *Neural Networks*, 19(3):323–337, 2006.

[99] Hyeon-Ae Jeon. Hierarchical processing in the prefrontal cortex in a variety of cognitive domains. *Frontiers in Systems Neuroscience*, 8, 2014.

[100] Matthew Johnson and Yiannis Demiris. Perceptual Perspective Taking and Action Recognition. *International Journal of Advanced Robotic Systems*, 2(4):32, 2005.

[101] Michael Kaiser, Holger Friedrich, and Udiger Dillmann. Obtaining Good Performance From a Bad Teacher. In *Programming by Demonstration vs. Learning from Examples Workshop at ML'95*, 1995.

[102] Weixi Kang, Sònia Pineda Hernández, and Jie Mei. Neural Mechanisms of Observational Learning: A Neural Working Model. *Frontiers in Human Neuroscience*, 14, 2021.

[103] Tasneem Kaochar, Raquel Torres Peralta, Clayton T. Morrison, Ian R. Fasel, Thomas J. Walsh, and Paul R. Cohen. Towards Understanding How Humans Teach Robots. In *User Modeling, Adaption and Personalization: 19th International Conference (UMAP '11)*, pages 347–352. Springer, 2011.

[104] Rachel Kendal, Lydia M. Hopper, Andrew Whiten, Sarah F. Brosnan, Susan P. Lambeth, Steven J. Schapiro, and Will Hoppitt. Chimpanzees copy dominant and knowledgeable individuals: Implications for cultural diversity. *Evolution and Human Behavior*, 36(1):65–72, 2015.

[105] Rachel L. Kendal, Neeltje J. Boogert, Luke Rendell, Kevin N. Laland, Mike Webster, and Patricia L. Jones. Social Learning Strategies: Bridge-Building between Fields. *Trends in Cognitive Sciences*, 22(7):651–665, 2018.

[106] Leonid Keselman, John Iselin Woodfill, Anders Grunnet-Jepsen, and Achintya Bhowmik. Intel(R) RealSense(TM) Stereoscopic Depth Cameras. In *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 1267–1276. IEEE Press, 2017.

[107] Faisal Khan, Bilge Mutlu, and Jerry Zhu. How Do Humans Teach: On Curriculum Learning and Teaching Dimension. In *Advances in Neural Information Processing Systems (NIPS '11)*, 2011.

[108] Cory D. Kidd and Cynthia Breazeal. Robots at home: Understanding long-term human-robot interaction. In *Proceedings of the 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '08)*, pages 3230–3235. IEEE, 2008.

[109] Elizabeth S. Kim, Dan Leyzberg, Katherine M. Tsui, and Brian Scassellati. How people talk when teaching a robot. In *Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction (HRI '09)*, page 23–30. ACM, 2009.

[110] Robert M. Klassen, Sündüs Yerdelen, and Tracy L. Durksen. Measuring Teacher Engagement: Development of the Engaged Teachers Scale (ETS). *Frontline Learning Research*, 1(2):33–52, 2013.

[111] Moritz Knaust and Dorothea Koert. Guided Robot Skill Learning: A User-Study on Learning Probabilistic Movement Primitives with Non-Experts. In *Proceedings of*

the 2020 IEEE-RAS 20th International Conference on Humanoid Robots (Humanoids '20), pages 514–521. IEEE, 2021.

[112] Melissa A. Koenig and Mark A. Sabbagh. Selective Social Learning: New Perspectives on Learning From Others. *Developmental Psychology*, 49(3):399–403, 2013.

[113] Nathan Koenig, Leila Takayama, and Maja Matarić. Communication and knowledge sharing in human-robot interaction and learning from demonstration. *Neural Networks*, 23(8-9):1104–1112, 2010.

[114] Terry K. Koo and Mae Y. Li. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, 15(2):155–163, 2016.

[115] Fred A. J. Korthagen. In search of the essence of a good teacher: towards a more holistic approach in teacher education. *Teaching and Teacher Education*, 20(1):77–97, 2004.

[116] Iuliia Kotseruba and John K. Tsotsos. 40 Years of Cognitive Architectures: Core Cognitive Abilities and Practical Applications. *Artificial Intelligence Review*, 53(1):17–94, 2020.

[117] K. Krishna and M. Narasimha Murty. Genetic K-Means Algorithm. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 29(3):433–439, 1999.

[118] Sanjay Krishnan, Roy Fox, Ion Stoica, and Ken Goldberg. DDCO: Discovery of Deep Continuous Options for Robot Learning from Demonstrations. In *Proceedings of the 1st Conference on Robot Learning (CoRL '17)*, pages 418–437. PMLR, 2017.

[119] Oliver Kroemer, Scott Niekum, and George Konidaris. A Review of Robot Learning for Manipulation: Challenges, Representations, and Algorithms. *Journal of Machine Learning Research*, 22(1), 2021.

[120] Sau lai Lee, Ivy Yee man Lau, S. Kiesler, and Chi-Yue Chiu. Human Mental Models of Humanoid Robots. In *Proceedings of the 2005 IEEE International Conference on Robotics and Automation (ICRA '05)*, pages 2767–2772. IEEE Press, 2005.

[121] John E. Laird, Kevin Gluck, John Anderson, Kenneth D. Forbus, Odest Chadwicke Jenkins, and Christian Lebiere. Interactive task learning. *IEEE Intelligent Systems*, 32(4):6–21, 2017.

[122] Ehsan Latif, Ramviyas Parasuraman, and Xiaoming Zhai. PhysicsAssistant: An LLM-Powered Interactive Learning Robot for Physics Lab Investigations. In *Proceedings of the 33rd IEEE International Conference on Robot and Human Interactive Communication (ROMAN '24)*, pages 864–871. IEEE Press, 2024.

[123] Kyuhwa Lee, Yanyu Su, Tae-Kyun Kim, and Yiannis Demiris. A syntactic approach to robot imitation learning using probabilistic activity grammars. *Robotics and Autonomous Systems*, 61(12):1323–1334, 2013.

[124] Alan M. Leslie, Ori Friedman, and Tim P. German. Core mechanisms in 'theory of mind'. *Trends in Cognitive Sciences*, 8(12):528–533, 2004.

[125] Jacky Liang, Wenlong Huang, Fei Xia, Peng Xu, Karol Hausman, Brian Ichter, Pete Florence, and Andy Zeng. Code as Policies: Language Model Programs for Embodied Control. In *arXiv preprint arXiv:2209.07753*, 2022.

[126] Jacky Liang, Fei Xia, Wenhao Yu, Andy Zeng, Montserrat Gonzalez Arenas, Maria Attarian, et al. Learning to Learn Faster from Human Feedback with Language Model Predictive Control. In *arXiv preprint arXiv:2402.11450*, 2024.

[127] Junchi Liang, Bowen Wen, Kostas Bekris, and Abdeslam Boularias. Learning Sensorimotor Primitives of Sequential Manipulation Tasks from Visual Demonstrations. In *Proceedings of the 2022 IEEE International Conference on Robotics and Automation (ICRA '22)*, pages 8591–8597. IEEE Press, 2022.

[128] Cynthia Lightfoot, Sheila R. Cole, and Michael Cole. The Study of Human Development. In *The Development of Children (7th ed.)*, page 1–44. Worth Publishers, 2012.

[129] Chenghao Liu, Fei Zhu, Quan Liu, and Yuchen Fu. Hierarchical Reinforcement Learning With Automatic Sub-Goal Identification. *IEEE/CAA Journal of Automatica Sinica*, 8(10):1686–1696, 2021.

[130] Zeyi Liu, Cheng Chi, Eric Cousineau, Naveen Kuppuswamy, Benjamin Burchfiel, and Shuran Song. ManiWAV: Learning Robot Manipulation from In-the-Wild Audio-Visual Data. In *In Proceedings of the 8th Annual Conference on Robot Learning (CoRL '24)*, 2024.

[131] Manuel Lopes and José Santos-Victor. A Developmental Roadmap for Learning by Imitation in Robots. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 37(2):308–321, 4 2007.

[132] Wing-Yue Geoffrey Louie and Goldie Nejat. A learning from demonstration system architecture for robots learning social group recreational activities. In *Proceedings of the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '16)*, pages 808–814. IEEE Press, 2016.

[133] Alan Lukežic, Tomáš Vojír, Luka Cehovin Zajc, Jirí Matas, and Matej Kristan. Discriminative Correlation Filter with Channel and Spatial Reliability. In *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR '17)*, pages 6309–6318. IEEE, 2017.

[134] Miguel Lázaro-Gredilla, Dianhuan Lin, J. Swaroop Guntupalli, and Dileep George. Beyond imitation: Zero-shot task transfer on robots by learning concepts as cognitive programs. *Science Robotics*, 4(26), 2019.

[135] Andrew K. Martin, Garon Perceval, Islay Davies, Peter Su, Jasmine Huang, and Marcus Meinzer. Visual Perspective Taking in Young and Older Adults. *Journal of Experimental Psychology: General*, 148(11):2006–2026, 2019.

[136] Giorgio Metta, Paul Fitzpatrick, and Lorenzo Natale. YARP: Yet Another Robot Platform. *International Journal of Advanced Robotic Systems*, 3(1):43–48, 2006.

[137] Giorgio Metta, Lorenzo Natale, Francesco Nori, Giulio Sandini, David Vernon, Luciano Fadiga, Claes von Hofsten, Kerstin Rosander, Manuel Lopes, José Santos-Victor, Alexandre Bernardino, and Luis Montesano. The iCub humanoid robot: An open-systems platform for research in cognitive development. *Neural Networks*, 23(8):1125–1134, 2010.

[138] Nina Moorman, Nakul Gopalan, Aman Singh, Erin Hedlund-Botti, Mariah L Schrum, Chuxuan Yang, Lakshmi Seelam, and Matthew Gombolay. Investigating the Impact of Experience on a User's Ability to Perform Hierarchical Abstraction. In *Proceedings of the RSS 2023 Workshop on Learning for Task and Motion Planning*, 2023.

[139] Pietro Morasso. Mental simulation of actions for learning optimal poses. *Cognitive Robotics*, 3:185–200, 2023.

[140] Chrystopher L. Nehaniv and Kerstin Dautenhahn. Correspondence problems and mechanisms. In Kerstin Dautenhahn and Chrystopher L. Nehaniv, editors, *Imitation and Social Learning in Robots, Humans and Animals: Behavioural, Social and Communicative Dimensions*, page 19–66. Cambridge University Press, 2009.

[141] Ori Novanda. *Metrics to Evaluate Human Teaching Engagement from a Robot's Point of View*. PhD thesis, University of Hertfordshire, 2017.

[142] Andreas Olsson, Ewelina Knapska, and Björn Lindström. The neural and computational systems of social learning. *Nature Reviews Neuroscience*, 21(4):197–212, 2020.

[143] Eric M. Orendt, Myriel Fichtner, and Dominik Henrich. Robot Programming by Non-Experts: Intuitiveness and Robustness of One-Shot Robot Programming. In *Proceedings of the 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN '16)*, page 192–199. IEEE Press, 2016.

[144] Eduardo B. Ottoni, Briseida Dogo De Resende, and Patrícia Izar. Watching the best nutcrackers: What capuchin monkeys (*Cebus apella*) know about others' tool-using skills. *Animal Cognition*, 8(4):215–219, 2005.

[145] Ana-Lucia Pais Ureche and Aude Billard. Metrics for Assessing Human Skill When Demonstrating a Bimanual Task to a Robot. In *Proceedings of the 10th Annual ACM/IEEE International Conference on Human-Robot Interaction (HRI '15), Extended Abstracts*, page 37–38. ACM, 2015.

[146] Matthew K.X.J. Pan, Elizabeth A. Croft, and Günter Niemeyer. Evaluating Social Perception of Human-to-Robot Handovers Using the Robot Social Attributes Scale (RoSAS). In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction (HRI '18)*, pages 443–451. ACM, 2018.

[147] Katerina Pastra and Yiannis Aloimonos. The minimalist grammar of action. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1585):103–117, 2012.

[148] David Paulius, Alejandro Agostini, and Dongheui Lee. Long-Horizon Planning and Execution With Functional Object-Oriented Networks. *IEEE Robotics and Automation Letters*, 8(8):4513–4520, 2023.

[149] David Paulius, Yongqiang Huang, Roger Milton, William D Buchanan, Jeanine Sam, and Yu Sun. Functional Object-Oriented Network for Manipulation Learning. In *Proceedings of the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '16)*, pages 2655–2662. IEEE, 2016.

[150] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent

Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, and Édouard Duchesnay. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12(85):2825–2830, 2011.

[151] Penelope L. Peterson, Thomas Carpenter, and Elizabeth Fennema. Teachers' Knowledge of Students' Knowledge in Mathematics Problem Solving: Correlational and Case Analyses. *Journal of Educational Psychology*, 81(4):558–569, 1989.

[152] Diane Poulin-Dubois and Patricia Brosseau-Liard. The Developmental Origins of Selective Social Learning. *Current Directions in Psychological Science*, 25(1):60–64, 2016.

[153] David Premack and Guy Woodruff. Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1(4):515–526, 1978.

[154] Preeti Ramaraj, Charles L. Ortiz, and Shiwali Mohan. Unpacking human teachers' intentions for natural interactive task learning. In *Proceedings of the 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN '21)*, pages 1173–1180. IEEE, 2021.

[155] Harish Ravichandar, Athanasios S. Polydoros, Sonia Chernova, and Aude Billard. Recent Advances in Robot Learning from Demonstration. *Annual Review of Control, Robotics, and Autonomous Systems*, 3(1):297–330, 2020.

[156] Rania Rayyes, Heiko Donat, Jochen Steil, and Michael Spranger. Interest-Driven Exploration With Observational Learning for Developmental Robots. *IEEE Transactions on Cognitive and Developmental Systems*, 15(2):373–384, 2023.

[157] Phillip Richter, Heiko Wersing, and Anna-Lisa Vollmer. Improving Human-Robot Teaching by Quantifying and Reducing Mental Model Mismatch. In *arXiv preprint arXiv:2501.04755*, 2025.

[158] Francisco J. Romero-Ramirez, Rafael Muñoz-Salinas, and Rafael Medina-Carnicer. Speeded up detection of squared fiducial markers. *Image and Vision Computing*, 76:38–47, 2018.

[159] David A. Rosenbaum, Lanyun Gong, and Cory Adam Potts. Pre-Crastination: Hastening Subgoal Completion at the Expense of Extra Physical Effort. *Psychological Science*, 25(7):1487–1496, 2014.

[160] Ken Rowe. The importance of teacher quality as a key determinant of students' experiences and outcomes of schooling. In *Building teacher quality: research conference proceedings*. ACER, 2003.

[161] Robert A. Rutter and Jeffrey D. Jacobson. *Facilitating Teacher Engagement*. National Center on Effective Secondary Schools, 1986.

[162] Md. Sadman Sakib, David Paulius, and Yu Sun. Approximate Task Tree Retrieval in a Knowledge Network for Robotic Cooking. *IEEE Robotics and Automation Letters*, 7(4):11492–11499, 2022.

[163] Maram Sakr, Martin Freeman, H.F. Machiel Van der Loos, and Elizabeth Croft. Training Human Teacher to Improve Robot Learning from Demonstration: A Pilot Study on Kinesthetic Teaching. In *Proceedings of the 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN '20)*, pages 800–806. IEEE Press, 2020.

[164] Maram Sakr, Zexi Jesse Li, H.F. Machiel Van der Loos, Dana Kulić, and Elizabeth A Croft. Quantifying Demonstration Quality for Robot Learning and Generalization. *IEEE Robotics and Automation Letters*, 7(4):9659–9666, 2022.

[165] Maram Sakr, Zhikai Zhang, Benjamin Li, Haomiao Zhang, H.F. Machiel Van der Loos, Dana Kulić, and Elizabeth Croft. How Can Everyday Users Efficiently Teach Robots by Demonstration? *ACM Transactions on Human-Robot Interaction*, 2025.

[166] Claude Sammut, Scott Hurst, Dana Kedzier, and Donald Michie. Learning to Fly. In Kerstin Dautenhahn and Chrystopher L. Nehaniv, editors, *Imitation in Animals and Artifacts*, page 171–190. MIT press, 2002.

[167] Akanksha Saran, Kush Desai, Mai Lee Chang, Rudolf Lioutikov, Andrea Thomaz, and Scott Niekum. Understanding Acoustic Patterns of Human Teachers Demonstrating Manipulation Tasks to Robots. In *Proceedings of the 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '22)*. IEEE, 2022.

[168] Akanksha Saran, Elaine Schaertl Short, Andrea Thomaz, and Scott Niekum. Understanding Teacher Gaze Patterns for Robot Learning. In *Proceedings of the 3rd Conference on Robot Learning (CoRL '19)*. PMLR, 2019.

[169] Joe Saunders, Chrystopher L. Nehaniv, and Kerstin Dautenhahn. Teaching robots by moulding behavior and scaffolding the environment. In *Proceedings of the 1st*

ACM SIGCHI/SIGART conference on Human-robot interaction (HRI '06), pages 118–125. ACM, 2006.

[170] Joe Saunders, Chrystopher L. Nehaniv, Kerstin Dautenhahn, and Aris Alissandrakis. Self-Imitation and Environmental Scaffolding for Robot Teaching. *International Journal of Advanced Robotic Systems*, 4(1):109–124, 2007.

[171] Lisa Scherf, Aljoscha Schmidt, Suman Pal, and Dorothea Koert. Interactively learning behavior trees from imperfect human demonstrations. *Frontiers in Robotics and AI*, 10, 2023.

[172] Mariah L. Schrum, Erin Hedlund-Botti, and Matthew Gombolay. Reciprocal MIND MELD: Improving Learning From Demonstration via Personalized, Reciprocal Teaching. In *Proceedings of the 6th Conference on Robot Learning (CoRL '23)*, pages 956–966. PMLR, 2023.

[173] Caroline Schuppli, Ellen J.M. Meulman, Sofia I.F. Forss, Fikty Aprilinayati, Maria A. van Noordwijk, and Carel P. van Schaik. Observational social learning and socially induced practice of routine skills in immature wild orang-utans. *Animal Behaviour*, 119:87–98, 2016.

[174] Matthias Schurz, Markus Aichhorn, Anna Martin, and Josef Perner. Common brain areas engaged in false belief reasoning and visual perspective taking: a meta-analysis of functional brain imaging studies. *Frontiers in Human Neuroscience*, 7, 2013.

[175] Aran Sena and Matthew Howard. Quantifying teaching behavior in robot learning from demonstration. *The International Journal of Robotics Research*, 39(1):54–72, 2019.

[176] Karim Shabani, Mohamad Khatib, and Saman Ebadi. Vygotsky's Zone of Proximal Development: Instructional Implications and Teachers' Professional Development. *English Language Teaching*, 3(4):237–248, 2010.

[177] Pourya Shahverdi, Alexander Tyshka, Madeline Trombly, and Wing-Yue Geoffrey Louie. Learning Turn-Taking Behavior from Human Demonstrations for Social Human-Robot Interactions. In *Proceedings of the 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '22)*, pages 7643–7649. IEEE Press, 2022.

[178] Pratyusha Sharma, Balakumar Sundaralingam, Valts Blukis, Chris Paxton, Tucker Hermans, Antonio Torralba, Jacob Andreas, and Dieter Fox. Correcting Robot Plans with Natural Language Feedback. In *arXiv preprint arXiv:2204.05186*, 2022.

[179] Ibrahim Shehi Shehu, Yafei Wang, Athuman Mohamed Athuman, and Xianping Fu. Remote Eye Gaze Tracking Research: A Comparative Evaluation on Past and Recent Progress. *Electronics*, 10(24), 2021.

[180] Remy Siegfried, Bozorgmehr Aminian, and Jean-Marc Odobez. ManiGaze: A Dataset for Evaluating Remote Gaze Estimator in Object Manipulation Situations. In *Proceedings of the ACM Symposium on Eye Tracking Research and Applications (ETRA '20), Short Papers*. ACM, 2020.

[181] Chan Hee Song, Jiaman Wu, Clayton Washington, Brian M. Sadler, Wei-Lun Chao, and Yu Su. LLM-Planner: Few-Shot Grounded Planning for Embodied Agents with Large Language Models. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV '23)*. IEEE Press, 2023.

[182] Sotiris Stavridis, Dimitrios Papageorgiou, and Zoe Doulgeri. Kinesthetic teaching of bi-manual tasks with known relative constraints. In *Proceedings of the 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '22)*, pages 11796–11801. IEEE Press, 2022.

[183] Mary Kay Stein, Randi A. Engle, Margaret S. Smith, and Elizabeth K. Hughes and. Orchestrating Productive Mathematical Discussions: Five Practices for Helping Teachers Move Beyond Show and Tell. *Mathematical Thinking and Learning*, 10(4):313–340, 2008.

[184] Gavin Suddrey, Christopher Lehnert, Markus Eich, Frederic Maire, and Jonathan Roberts. Teaching Robots Generalizable Hierarchical Tasks Through Natural Language Instruction. *IEEE Robotics and Automation Letters*, 2(1):201–208, 2017.

[185] Douglas Summers-Stay, Ching L. Teo, Yezhou Yang, Cornelia Fermüller, and Yiannis Aloimonos. Using a Minimal Action Grammar for Activity Understanding in the Real World. In *Proceedings of the 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '12)*, pages 4104–4111. IEEE, 2012.

[186] John Sweller. Cognitive load theory, learning difficulty, and instructional design. *Learning and instruction*, 4(4):295–312, 1994.

[187] Ajay Kumar Tanwani, Nitesh Mor, John Kubiatowicz, Joseph E. Gonzalez, and Ken Goldberg. A Fog Robotics Approach to Deep Robot Learning: Application to Object Recognition and Grasp Planning in Surface Decluttering. In *Proceedings of the 2019 International Conference on Robotics and Automation (ICRA '19)*, pages 4559–4566. IEEE Press, 2019.

[188] Ajay Kumar Tanwani, Andy Yan, Jonathan Lee, Sylvain Calinon, and Ken Goldberg. Sequential robot imitation learning from observations. *The International Journal of Robotics Research*, 40(10-11):1306–1325, 2021.

[189] Konstantinos Theofilis, Katrin Solveig Lohan, Chrystopher L. Nehaniv, Kerstin Dautenhahn, and Britta Werde. Temporal emphasis for goal extraction in task demonstration to a humanoid robot by naive users. In *Proceedings of the 2013 IEEE Third Joint International Conference on Development and Learning and Epigenetic Robotics (ICDL '13)*. IEEE Press, 2013.

[190] Michael Tomasello. Cultural transmission in the tool use and communicatory signaling of chimpanzees? In Sue Taylor Parker and Kathleen Rita Gibson, editors, *'Language' and Intelligence in Monkeys and Apes, Comparative Developmental Perspectives*, pages 274–311. Cambridge University Press, 1990.

[191] Michael Tomasello. *The Cultural Origins of Human Cognition*. Harvard University Press, 1999.

[192] J. Gregory. Trafton, Nicholas L. Cassimatis, Magdalena D. Bugajska, Derek P. Brock, Farilee E. Mintz, and Alan C. Schultz. Enabling Effective Human–Robot Interaction Using Perspective-Taking in Robots. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 35(4):460–470, 2005.

[193] Kristen Swan Tummeltshammer, Rachel Wu, David M. Sobel, and Natasha Z. Kirkham. Infants Track the Reliability of Potential Informants. *Psychological Science*, 25(9):1730–1738, 2014.

[194] Alexander Tyshka and Wing-Yue Geoffrey Louie. Interactive Task Learning for Social Robots: A Pilot Study. In *Proceedings of the 2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '23)*, pages 4045–4051. IEEE Press, 2023.

[195] Constantin Uhde, Nicolas Berberich, Karinne Ramirez-Amaro, and Gordon Cheng. The Robot as Scientist: Using Mental Simulation to Test Causal Hypotheses Extracted from Human Activities in Virtual Reality. In *Proceedings of the 2020*

IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '20), pages 8081–8086. IEEE Press, 2020.

[196] Ana Lucia Pais Ureche, Keisuke Umezawa, Yoshihiko Nakamura, and Aude Billard. Task Parameterization Using Continuous Constraints Extracted from Human Demonstrations. *IEEE Transactions on Robotics*, 31(6):1458–1471, 2015.

[197] Christina Visconti and Elena Neiterman. Shifting to Primary Prevention for an Aging Population: A Scoping Review of Health Promotion Initiatives for Community-Dwelling Older Adults in Canada. *Cureus*, 13(8), 2021.

[198] Rachel L. VonderHaar, Dawn M. McBride, and David A. Rosenbaum. Task order choices in cognitive and perceptual-motor tasks: The cognitive-load-reduction (CLEAR) hypothesis. *Attention, Perception, Psychophysics*, 81(7):2517–2525, 2019.

[199] Basil Wahn, Leda Berio, Matthias Weiß, and Albert Newen. Try to see it my way: Humans take the level-1 visual perspective of humanoid robot avatars. *International Journal of Social Robotics*, 17(3):523–534, 2025.

[200] Jonas Werner, Kun Chu, Cornelius Weber, and Stefan Wermter. LLM-based Interactive Imitation Learning for Robotic Manipulation. In *arXiv preprint arXiv:2504.21769*, 2025.

[201] Andrew Whiten. Imitation of the Sequential Structure of Actions by Chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, 112(3):270–281, 1998.

[202] Andrew Whiten. The burgeoning reach of animal culture. *Science*, 372(6537), 2021.

[203] Andrew Whiten, Emma Flynn, Katy Brown, and Tanya Lee. Imitation of hierarchical action structure by young children. *Developmental Science*, 9(6):574–582, 2006.

[204] Andrew Whiten, Nicola McGuigan, Sarah Marshall-Pescini, and Lydia M. Hopper. Emulation, imitation, over-imitation and the scope of culture for child and chimpanzee. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1528):2417–2428, 2009.

[205] Andrew Whiten and Christian Rutz. The growing methodological toolkit for identifying and studying social learning and culture in non-human animals. *Philosophical Transactions B*, 380(1925):20240140, 2025.

[206] Andrew Whiten and Erica van de Waal. The pervasive role of social learning in primate lifetime development. *Behavioral Ecology and Sociobiology*, 72(5):80, May 2018.

[207] Felix Wohlgemuth, Iori Mizutani, Lukas Eichelberger, and Simon Mayer. Electromyography-based Kinesthetic Teaching of Industrial Collaborative Robots. In *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction (HRI '24)*, page 1124–1128. ACM, 2024.

[208] Luke Wood, Kerstin Dautenhahn, Ben Robins, and Abolfazl Zaraki. Developing Child-Robot Interaction Scenarios with a Humanoid Robot to Assist Children with Autism in Developing Visual Perspective Taking Skills. In *Proceedings of the 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN '17)*, pages 1055–1060. IEEE Press, 2017.

[209] Britta Wrede, Katharina J. Rohlfing, Marc Hanheide, and Gerhard Sagerer. Towards learning by interacting. In *Creating Brain-Like Intelligence*, pages 139–150. Springer, 2009.

[210] Sebastian Wrede, Christian Emmerich, Ricarda Grünberg, Arne Nordmann, Agnes Swadzba, and Jochen Steil. A User Study on Kinesthetic Teaching of Redundant Robots in Task and Configuration Space. *Journal of Human-Robot Interaction*, 2(1):56–81, 2013.

[211] Chengli Xiao, Ya Fan, Jingyu Zhang, and Renlai Zhou. People Do not Automatically Take the Level-1 Visual Perspective of Humanoid Robot Avatars. *International Journal of Social Robotics*, 14(1):165–176, 2022.

[212] Xuan Xiao, Jiahang Liu, Zhipeng Wang, Yanmin Zhou, Yong Qi, Shuo Jiang, Bin He, and Qian Cheng. Robot learning in the era of foundation models: a survey. *Neurocomputing*, 638:129963, 2025.

[213] Danfei Xu, Suraj Nair, Yuke Zhu, Julian Gao, Animesh Garg, Li Fei-Fei, and Silvio Savarese. Neural Task Programming: Learning to Generalize Across Hierarchical Tasks. In *Proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA '18)*. IEEE Press, 2018.

[214] Yezhou Yang, Cornelia Fermüller, and Yiannis Aloimonos. Detection of Manipulation Action Consequences (MAC). In *Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition*, pages 2563–2570. IEEE Press, 2013.

[215] Yezhou Yang, Yi Li, Cornelia Fermuller, and Yiannis Aloimonos. Robot Learning Manipulation Action Plans by "Watching" Unconstrained Videos from the World Wide Web. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence*, volume 29. AAAI Press, 2015.

[216] Tian Ye, Takashi Minato, Kurima Sakai, Hidenobu Sumioka, Antonia Hamilton, and Hiroshi Ishiguro. Human-like interactions prompt people to take a robot's perspective. *Frontiers in Psychology*, 14, 2023.

[217] Lotfi A. Zadeh. Fuzzy logic. *Computer*, 21(4):83–93, 1988.

[218] Robert B. Zajonc. Social facilitation. *Science*, 149(3681):269–274, 1965.

[219] Martina Zambelli and Yiannis Demiris. Online Multimodal Ensemble Learning Using Self-Learned Sensorimotor Representations. *IEEE Transactions on Cognitive and Developmental Systems*, 9(2):113–126, 2017.

[220] Abolfazl Zaraki, Luke Wood, Ben Robins, and Kerstin Dautenhahn. Development of a Semi-Autonomous Robotic System to Assist Children with Autism in Developing Visual Perspective Taking Skills. In *Proceedings of the 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN '18)*, pages 969–976. IEEE Press, 2018.

[221] Maryam Zare, Parham M. Kebria, Abbas Khosravi, and Saeid Nahavandi. A Survey of Imitation Learning: Algorithms, Recent Developments, and Challenges. *IEEE Transactions on Cybernetics*, 54(12):7173–7186, 2024.

[222] Xuan Zhao, Bertram Malle, and Hyowon Gweon. Is it a nine, or a six? Prosocial and selective perspective taking in four-year-olds. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 38, 2016.

[223] Xuan Zhao and Bertram F. Malle. Spontaneous perspective taking toward robots: The unique impact of humanlike appearance. *Cognition*, 224, 7 2022.

[224] Stephan Zheng, Yisong Yue, and Patrick Lucey. Generating Long-term Trajectories Using Deep Hierarchical Networks. In *Proceedings of the 30th Conference on Neural Information Processing Systems (NIPS '16)*, page 1551–1559. Curran Associates Inc., 2016.

APPENDICES

Appendix A

Ethics Clearance Certificates

UNIVERSITY OF WATERLOO

Notification of Ethics Clearance to Conduct Research with Human Participants

---

Principal Investigator: Kerstin Dautenhahn (Electrical and Computer Engineering)

Student investigator: Pourya Aliasghari (Electrical and Computer Engineering)

Co-Investigator: Moojan Ghafurian (Electrical and Computer Engineering)

Co-Investigator: Chrystopher Nehaniv (Systems Design Engineering)

File #: 43860

Title: Understanding the behaviour of human teachers when teaching robots (revised)

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

Initial Approval Date: 01/05/22 (m/d/y)

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

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

Expiry Date: 01/06/23 (m/d/y)

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

Level of review: Delegated Review

Signed on behalf of the Human Research Ethics Board



Vanessa Buote, Ethics Advisor, vbuote@uwaterloo.ca, 519-888-4567, ext. 30321

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: SafetyPlan.pdf

file: ResumeResearch.pdf

file: TobiiProGlasses2UserManual_v1_20200212.pdf

file: Shimmer3GSR+UserManual_v1_20200214.pdf

file: PepperUsersManual_v1_20200212.pdf

file: ScreeningQuestionnaire_v2_202112.docx

file: StudyFlyer_v2_202112.docx

file: ExperimentalSetup_v1_202110.pdf

file: AccessLab_v1_202110.pdf

file: Questionnaire_v1_202110.pdf

file: EmailScripts_v2_202112.docx

file: DemographicInformationForm_v1_202110.pdf

file: ConsentForm_v2_202112.docx

file: InformationSheet_v3_202112.docx

file: COVID_ConsentForm_v2_202112.docx

file: AppreciationLetter_v3_202201.pdf

file: AppreciationLetter_v3_202201.docx

file: DebriefingSheet_v3_202112.docx

file: VerbalDebriefingScript_v2_202112.docx

Approved Protocol Version 4 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.

UNIVERSITY OF WATERLOO

Notification of Ethics Clearance to Conduct Research with Human Participants

---

Principal Investigator: Kerstin Dautenhahn (Electrical and Computer Engineering)

Student investigator: Pourya Aliasghari (Electrical and Computer Engineering)

Co-Investigator: Chrystopher Nehaniv (Systems Design Engineering)

Co-Investigator: Moojan Ghafurian (Systems Design Engineering)

File #: 47048

Title: Exploring the Impact of Teachers' Awareness of Robots' Capabilities and Constraints on Robot Learning Performance

---

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

Initial Approval Date: 12/17/24 (m/d/y)

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

Expiry Date: 12/18/25 (m/d/y)

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

Level of review: Delegated Review

Signed on behalf of the Human Research Ethics Board

UNIVERSITY OF WATERLOO

Notification of Ethics Clearance to Conduct Research with Human Participants

---

Principal Investigator: Kerstin Dautenhahn (Electrical and Computer Engineering)

Student investigator: Pourya Aliasghari (Electrical and Computer Engineering)

Co-Investigator: Chrystopher Nehaniv (Systems Design Engineering)

Co-Investigator: Moojan Ghafurian (Systems Design Engineering)

File #: 47048

Title: Exploring the Impact of Teachers' Awareness of Robots' Capabilities and Constraints on Robot Learning Performance

---

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

Initial Approval Date: 12/17/24 (m/d/y)

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

Expiry Date: 12/18/25 (m/d/y)

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

Level of review: Delegated Review

Signed on behalf of the Human Research Ethics Board

Appendix B

Experiment 1 Additional Information

This appendix includes additional photos and information related to experimental material and the setup for Experiment 1. A video of an example experimental session can be found at: https://youtu.be/Ds2RUq5-rI8.

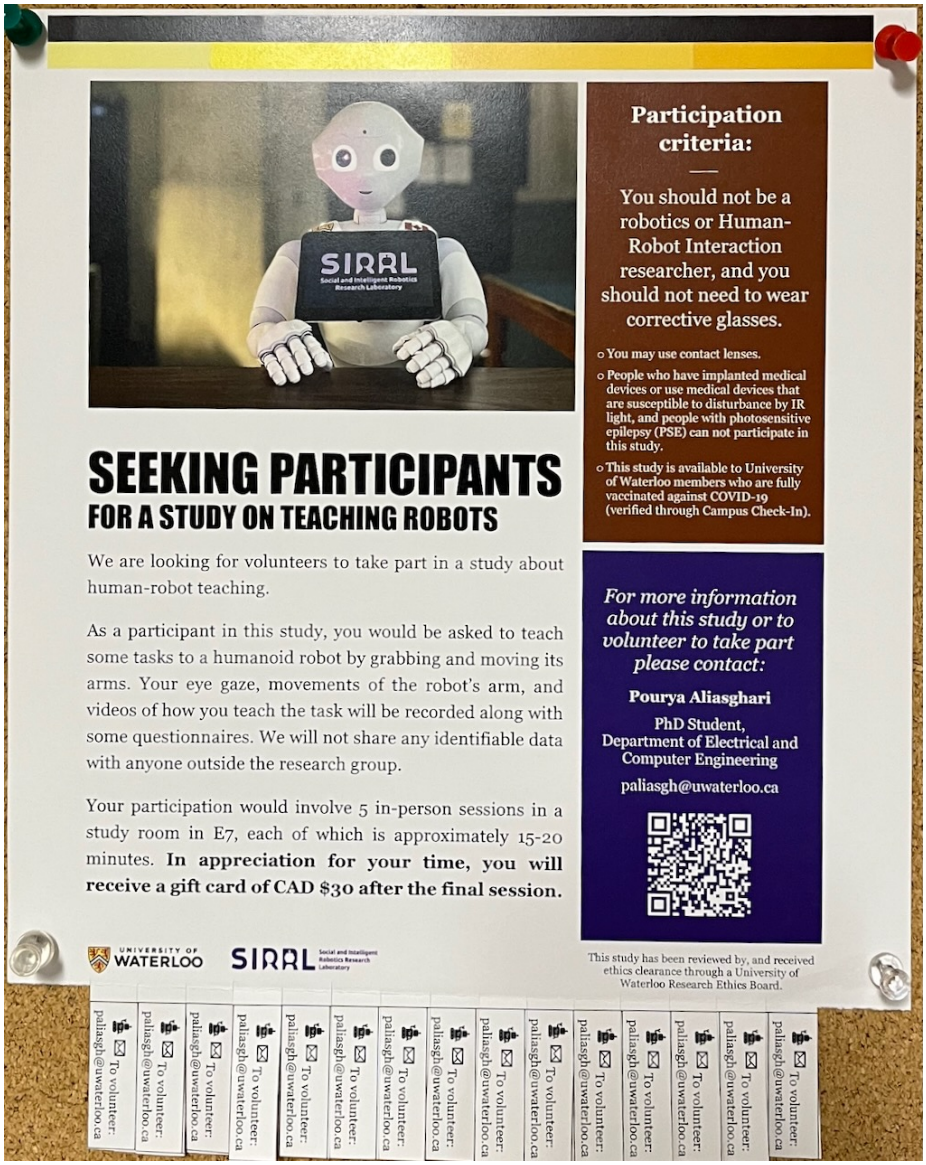


Figure B.1: The flyer used for advertising the experiment, posted on a public board. Students interested in participating could send an email to myself by scanning the QR code.

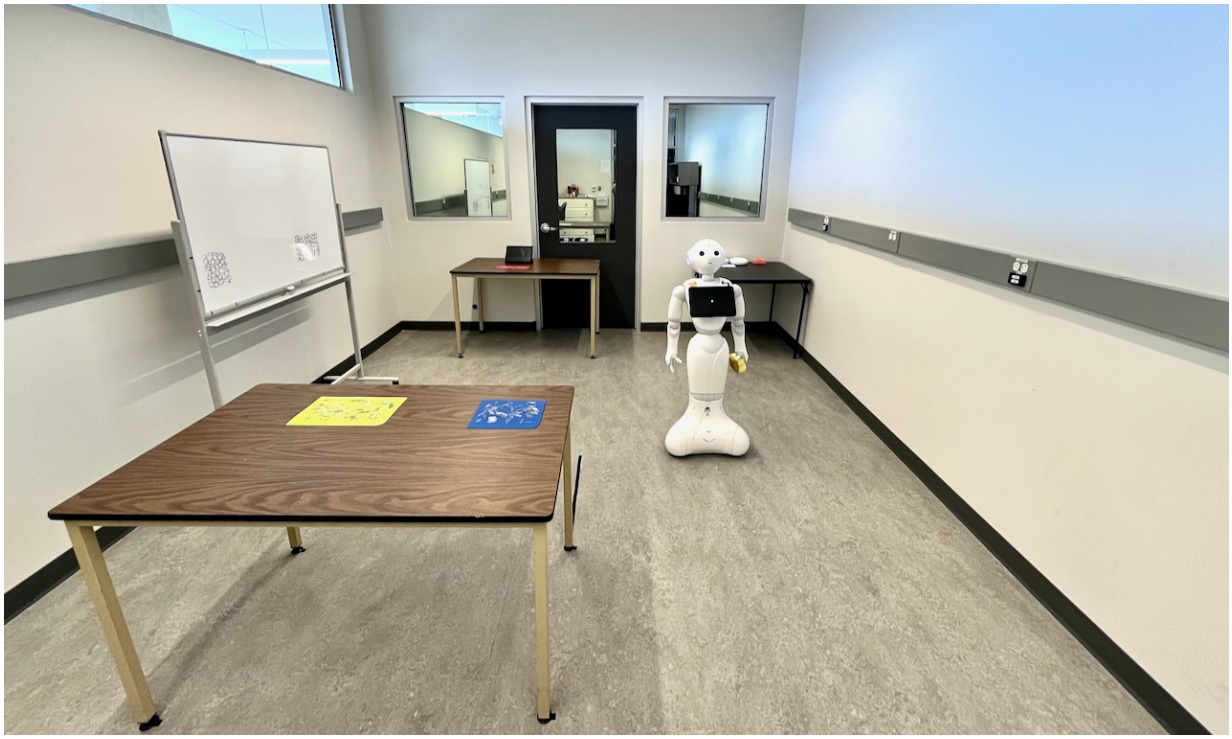


Figure B.2: The setup in the experimental room, before the arrival of participants. On the left side, the setup for task learning of the Pepper robot can be found, as was seen in Figure 3.1. In the middle, the door connecting the experimental room to the control room is visible. On the right side, materials for the cleaning tasks are stored for easy switching between the sessions.



Figure B.3: Different types of material that were spread on the cutting boards for the cleaning tasks. We used pieces of paper, paper clips, safety pins, binder clips, and buttons.

Appendix C

Experiment 2 Additional Information

This appendix includes snapshots from all the pages in each of the GHAs (Section C.1) as well as the designed online interface used for collecting questionnaire data from the participants, hosted on Qualtrics [7], before and during Experiment 2. (Section C.2). This appendix also includes additional photos and information related to the experimental material and the setup (Section C.3).

C.1 The Goal-Highlighting Activities

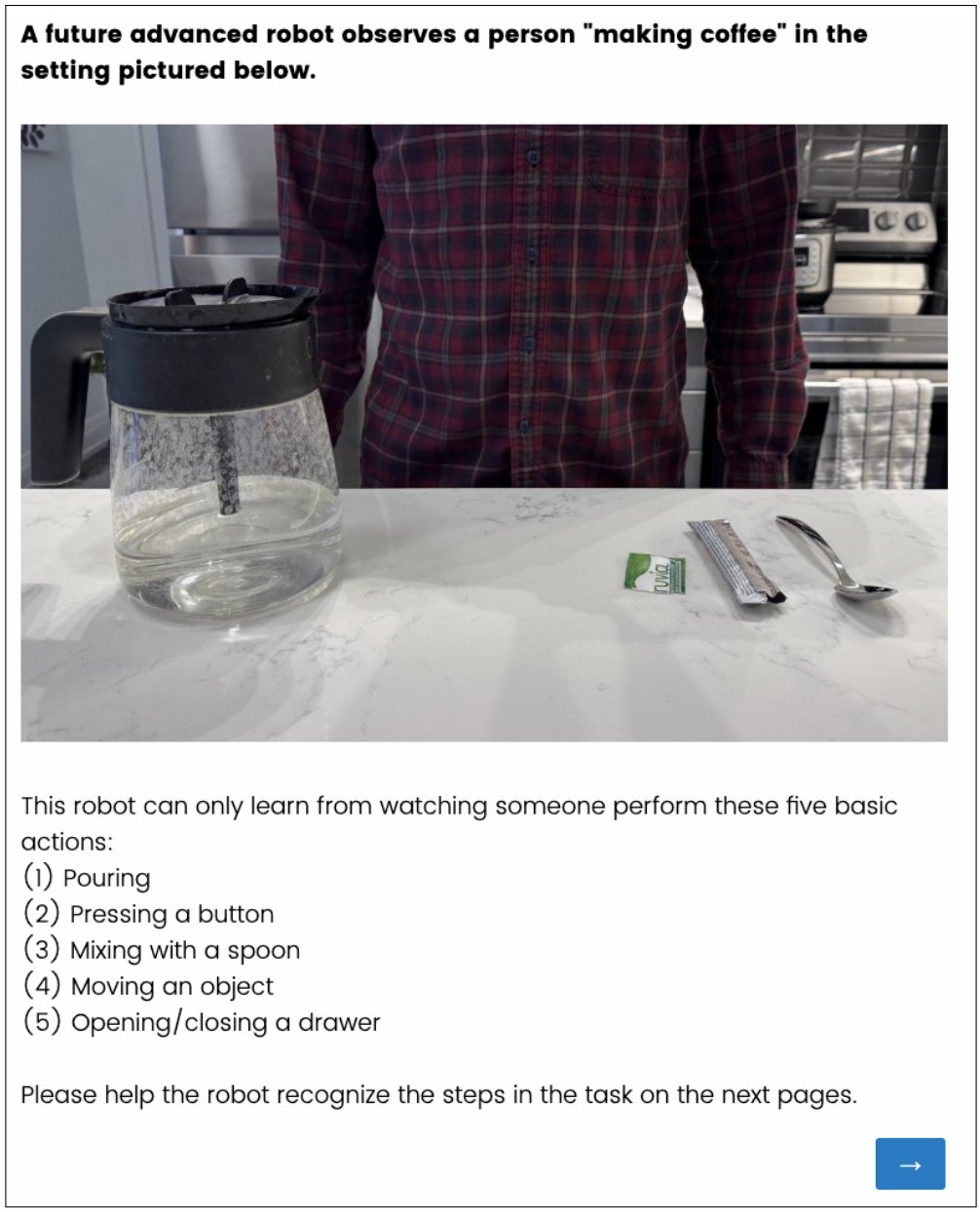


Figure C.1: The GHA about the scenario of a future advanced robot learning how to make coffee - Starting page

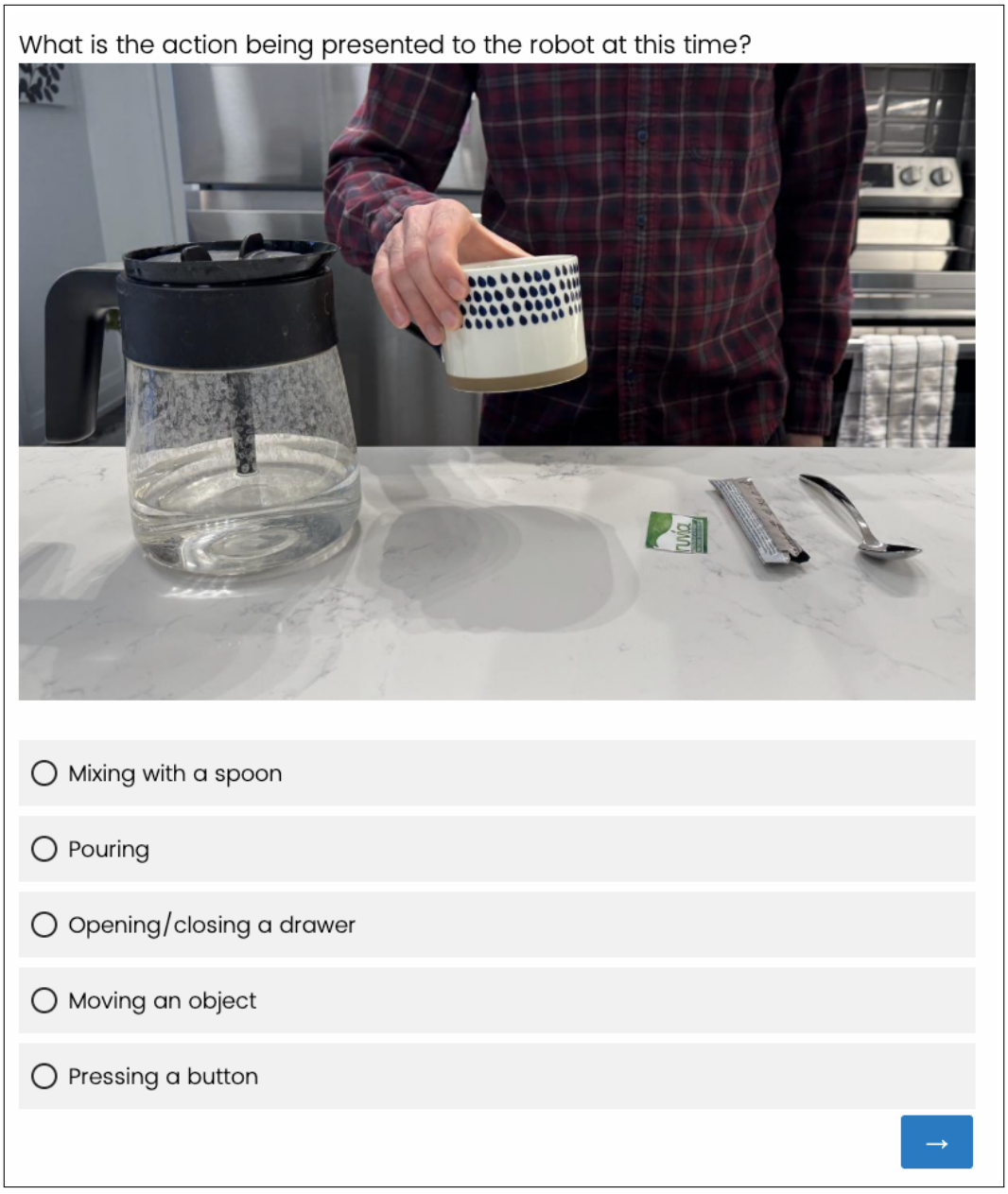


Figure C.1: [Cont.] The GHA about the scenario of a future advanced robot learning how to make coffee - Page 1

What is the action being presented to the robot at this time?



- ○ Mixing with a spoon
- ○ Moving an object
- ○ Pouring
- ○ Opening/closing a drawer
- ○ Pressing a button

→

Figure C.1: [Cont.] The GHA about the scenario of a future advanced robot learning how to make coffee - Page 2

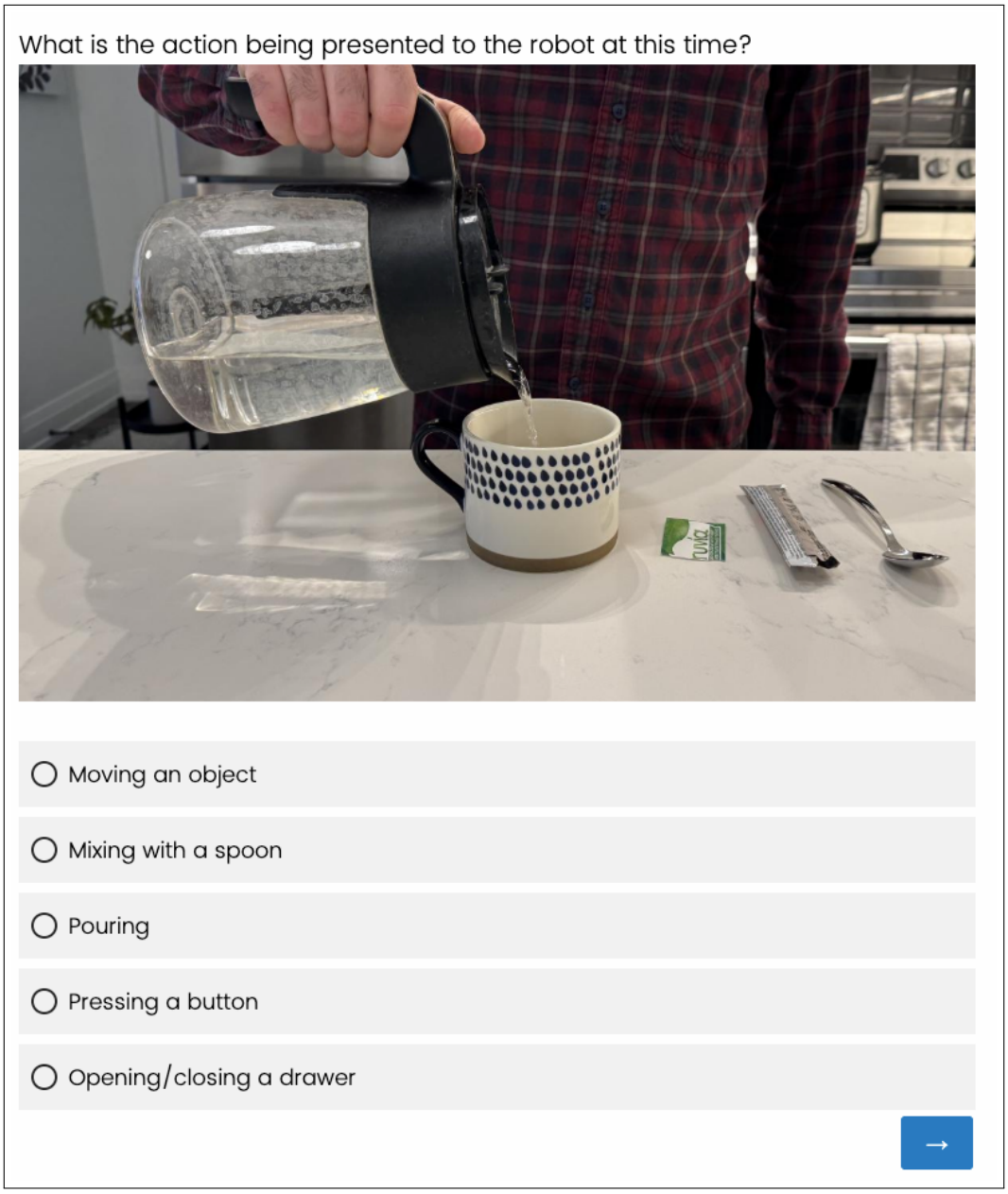


Figure C.1: [Cont.] The GHA about the scenario of a future advanced robot learning how to make coffee - Page 3

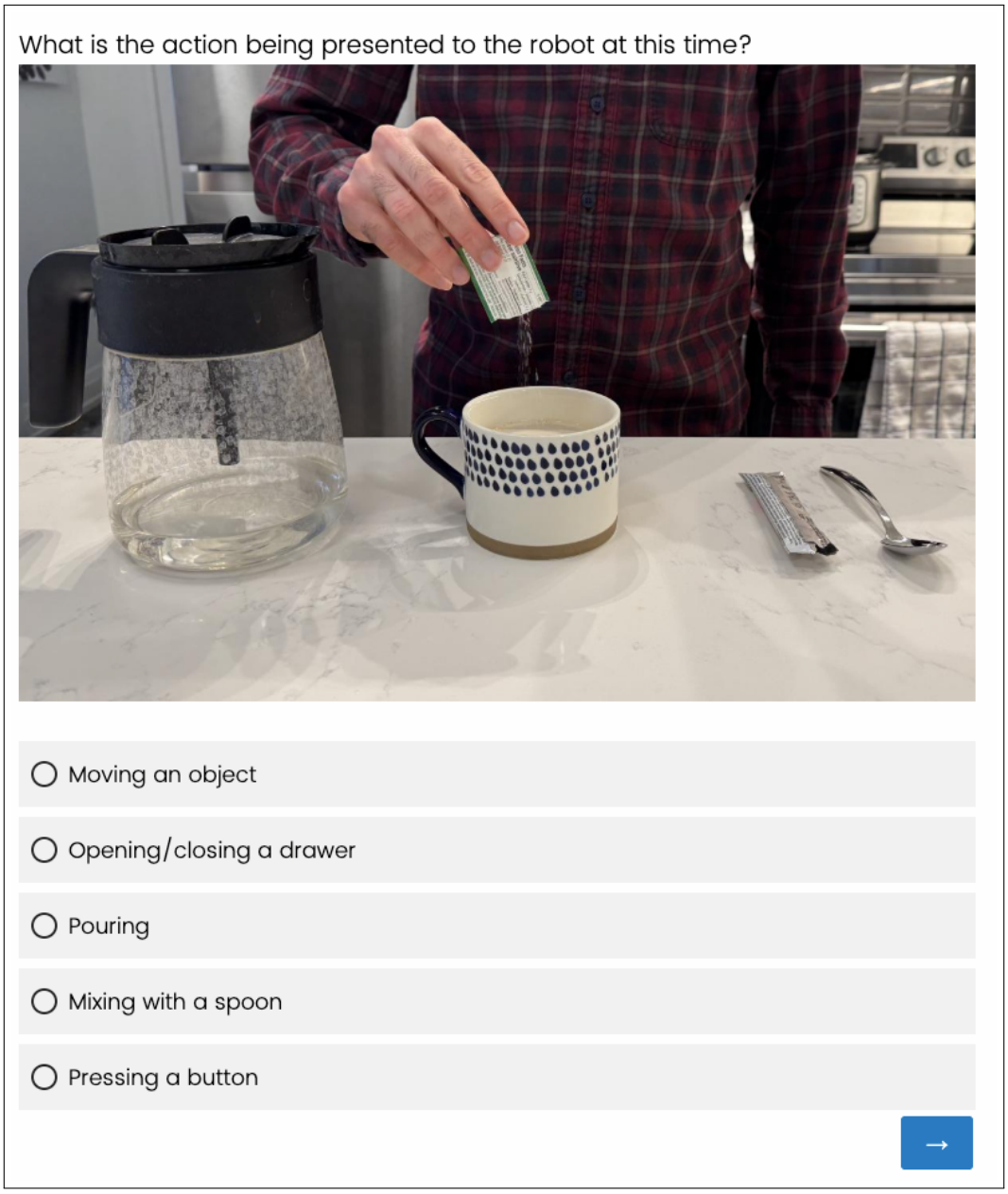


Figure C.1: [Cont.] The GHA about the scenario of a future advanced robot learning how to make coffee - Page 4

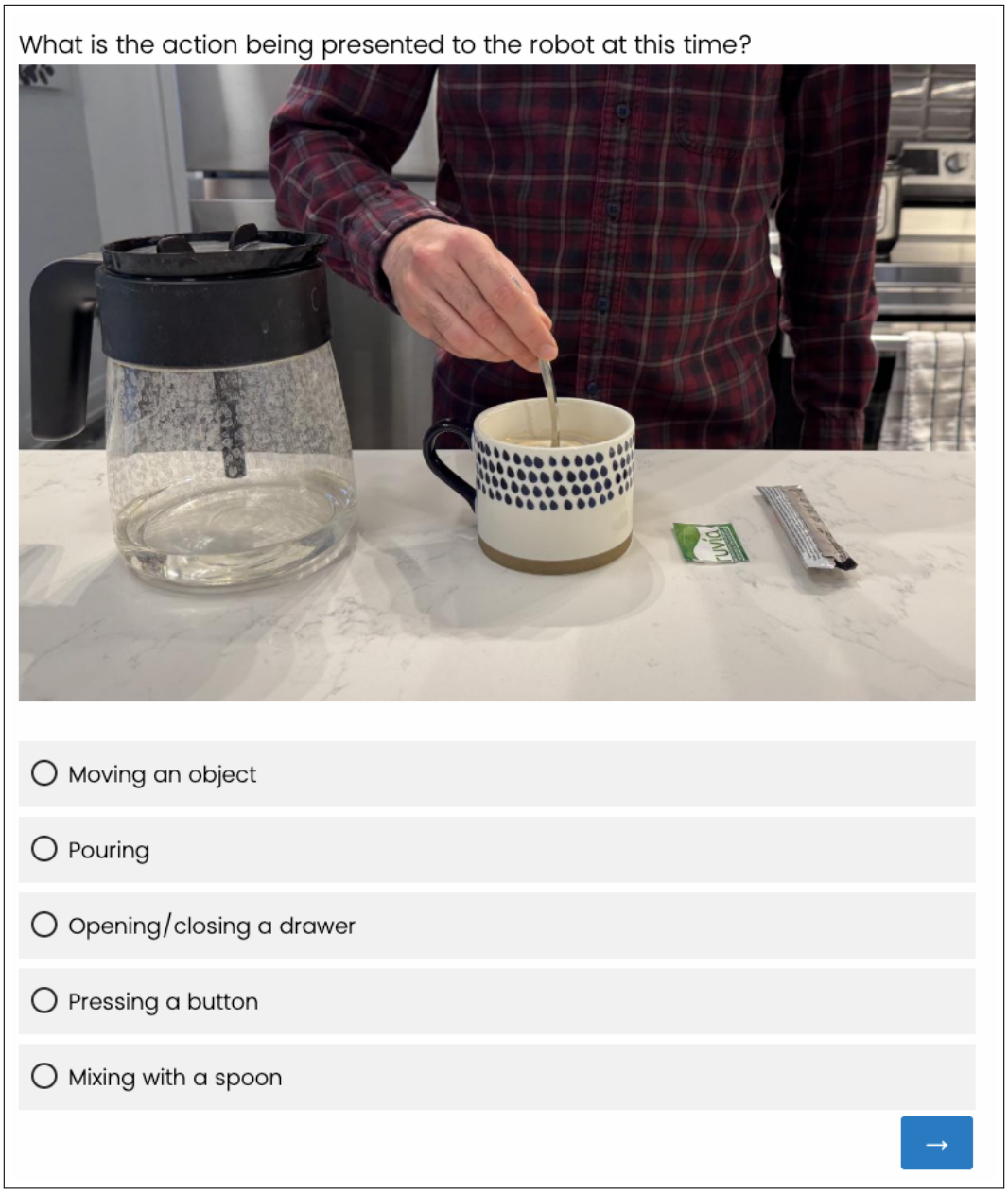


Figure C.1: [Cont.] The GHA about the scenario of a future advanced robot learning how to make coffee - Page 5

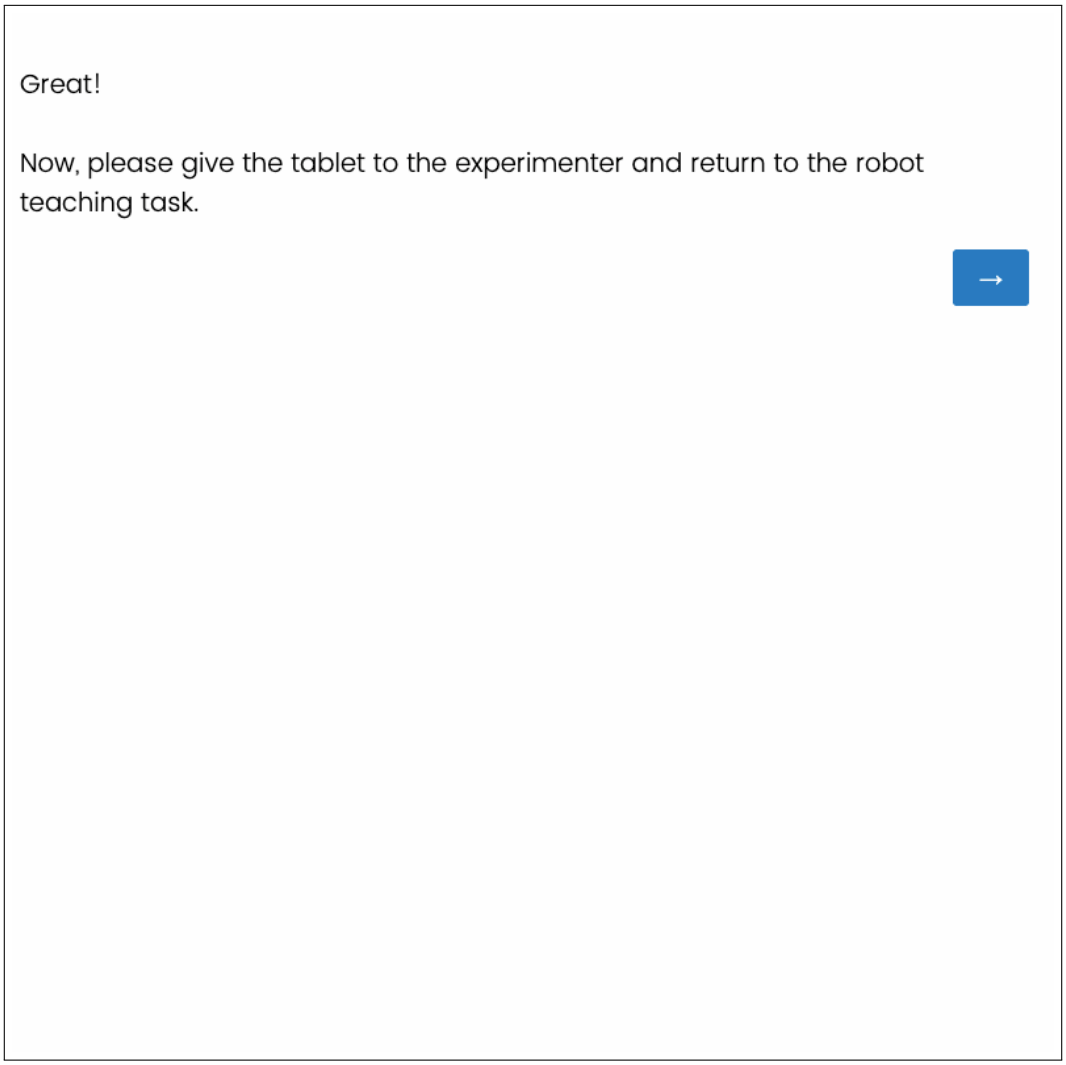


Figure C.1: [Cont.] The GHA about the scenario of a future advanced robot learning how to make coffee - Ending page

A future advanced robot observes a person "starting a washing machine" in the setting pictured below.



This robot can only learn from watching someone perform these five basic actions:
(1) Pouring
(2) Pressing a button
(3) Mixing with a spoon
(4) Moving an object
(5) Opening/closing a drawer

Please help the robot recognize the steps in the task on the next pages.

→

Figure C.2: The GHA about the scenario of a future advanced robot learning how to start a washing machine - Starting page

What is the action being presented to the robot at this time?



- ○ Opening/closing a drawer
- ○ Pouring
- ○ Mixing with a spoon
- ○ Moving an object
- ○ Pressing a button

→

Figure C.2: [Cont.] The GHA about the scenario of a future advanced robot learning how to start a washing machine - Page 1

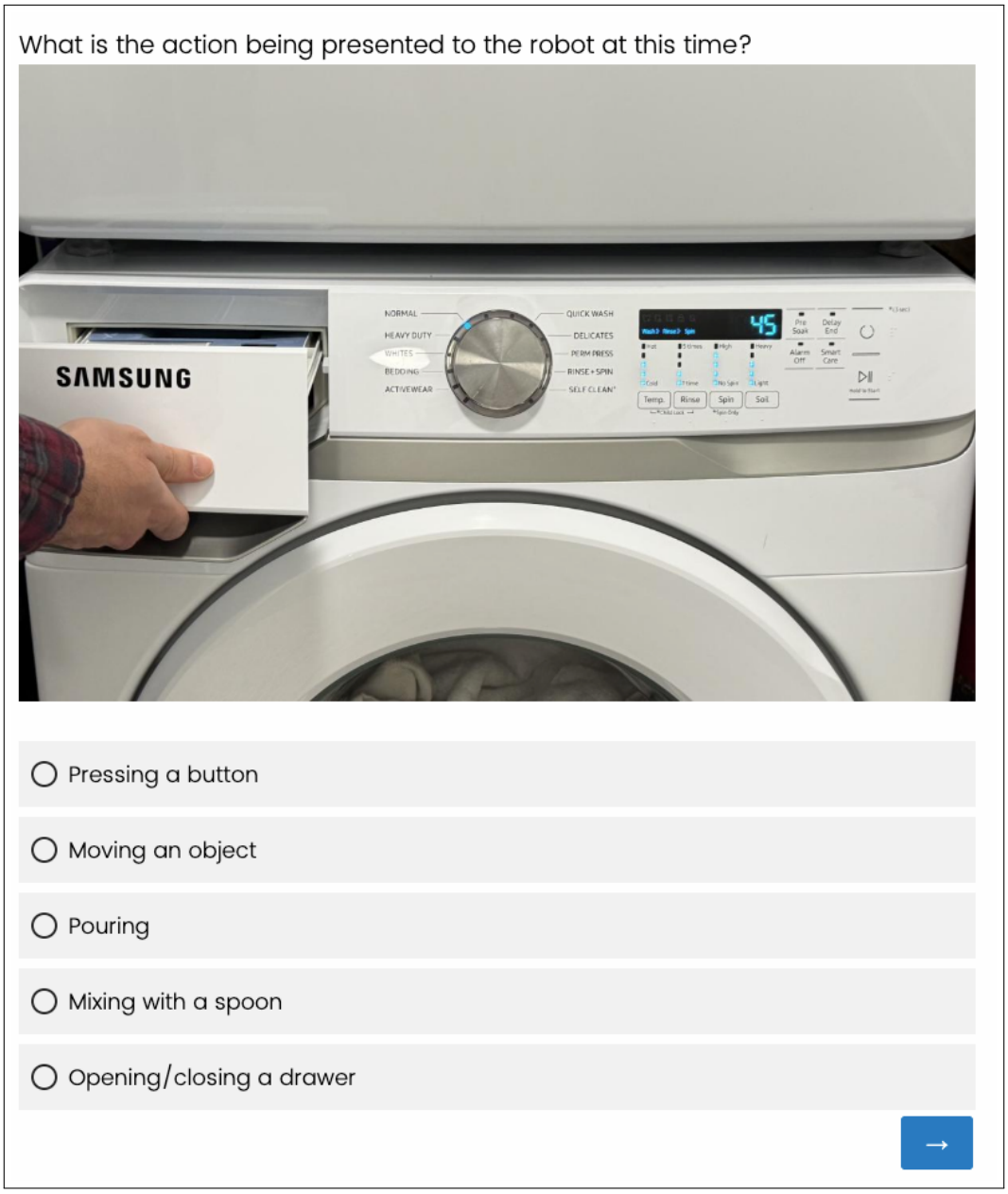


Figure C.2: [Cont.] The GHA about the scenario of a future advanced robot learning how to start a washing machine - Page 2

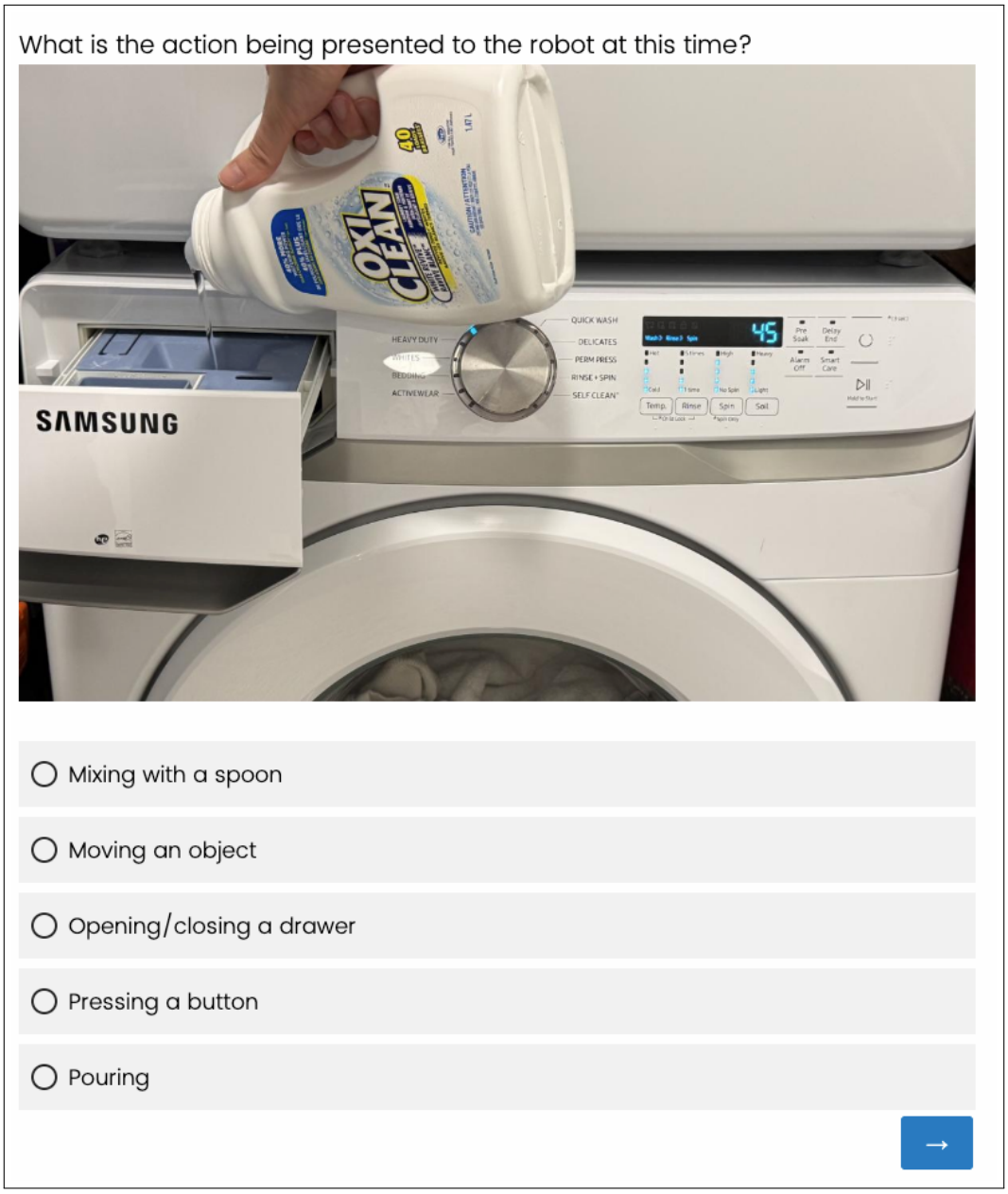


Figure C.2: [Cont.] The GHA about the scenario of a future advanced robot learning how to start a washing machine - Page 3

What is the action being presented to the robot at this time?



- ○ Moving an object
- ○ Pressing a button
- ○ Pouring
- ○ Mixing with a spoon
- ○ Opening/closing a drawer

→

Figure C.2: [Cont.] The GHA about the scenario of a future advanced robot learning how to start a washing machine - Page 4

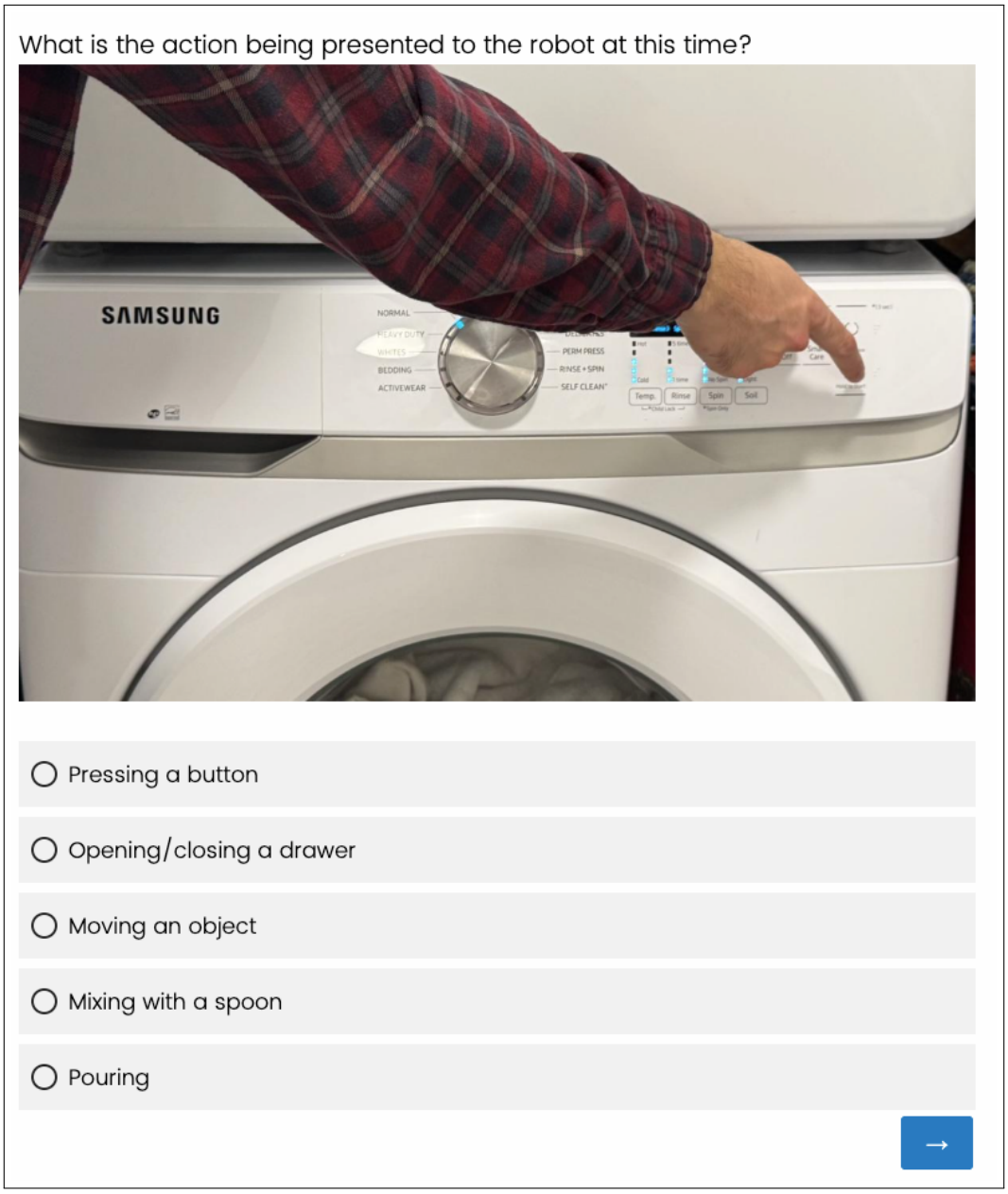


Figure C.2: [Cont.] The GHA about the scenario of a future advanced robot learning how to start a washing machine - Page 5

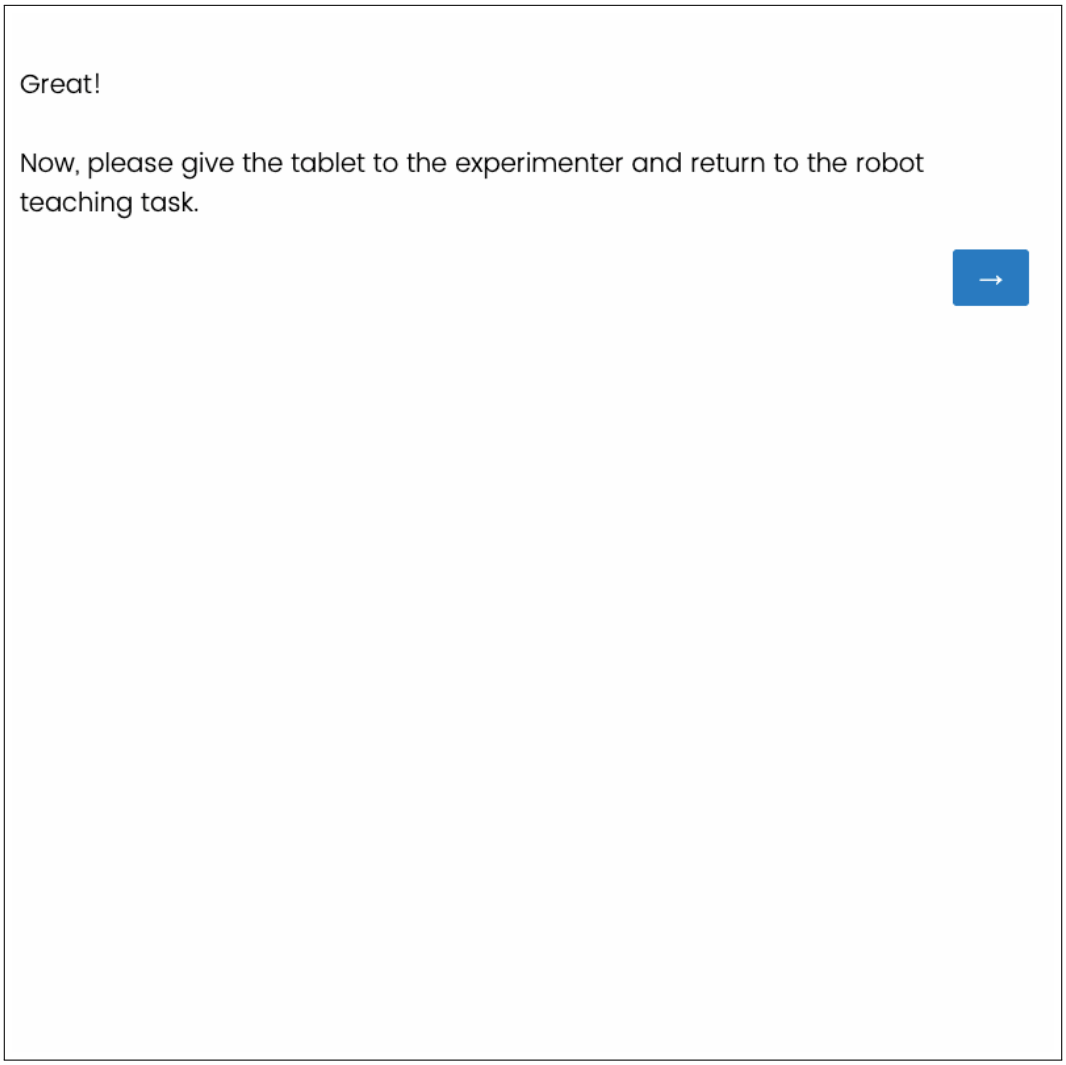


Figure C.2: [Cont.] The GHA about the scenario of a future advanced robot learning how to start a washing machine - Ending page

A future advanced robot observes a person "decluttering a kitchen countertop" in the setting pictured below.



This robot can only learn from watching someone perform these five basic actions:
(1) Pouring
(2) Pressing a button
(3) Mixing with a spoon
(4) Moving an object
(5) Opening/closing a drawer

Please help the robot recognize the steps in the task on the next pages.

→

Figure C.3: The GHA about the scenario of a future advanced robot learning how to declutter a kitchen space - Starting page

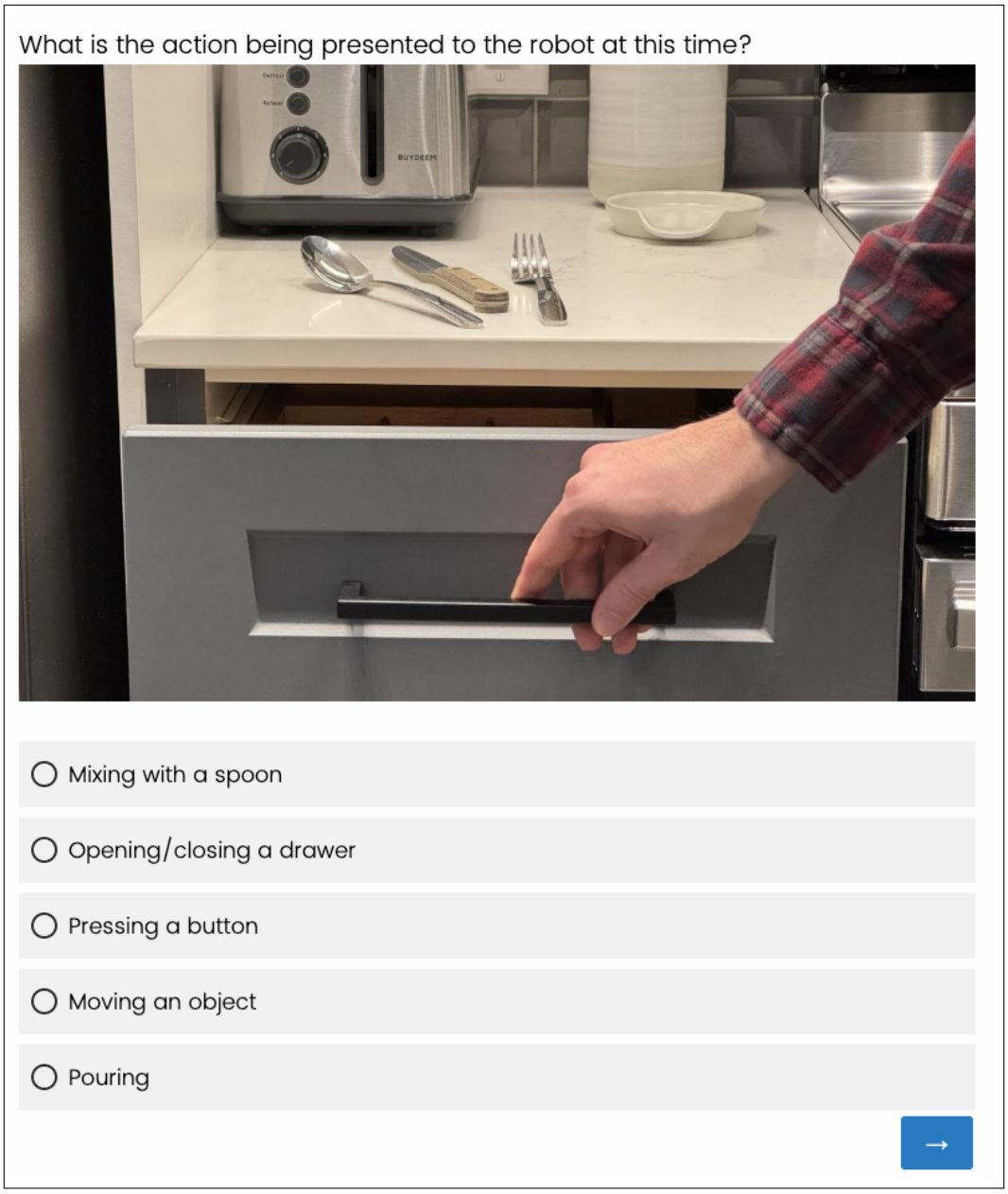


Figure C.3: [Cont.] The GHA about the scenario of a future advanced robot learning how to declutter a kitchen space - Page 1

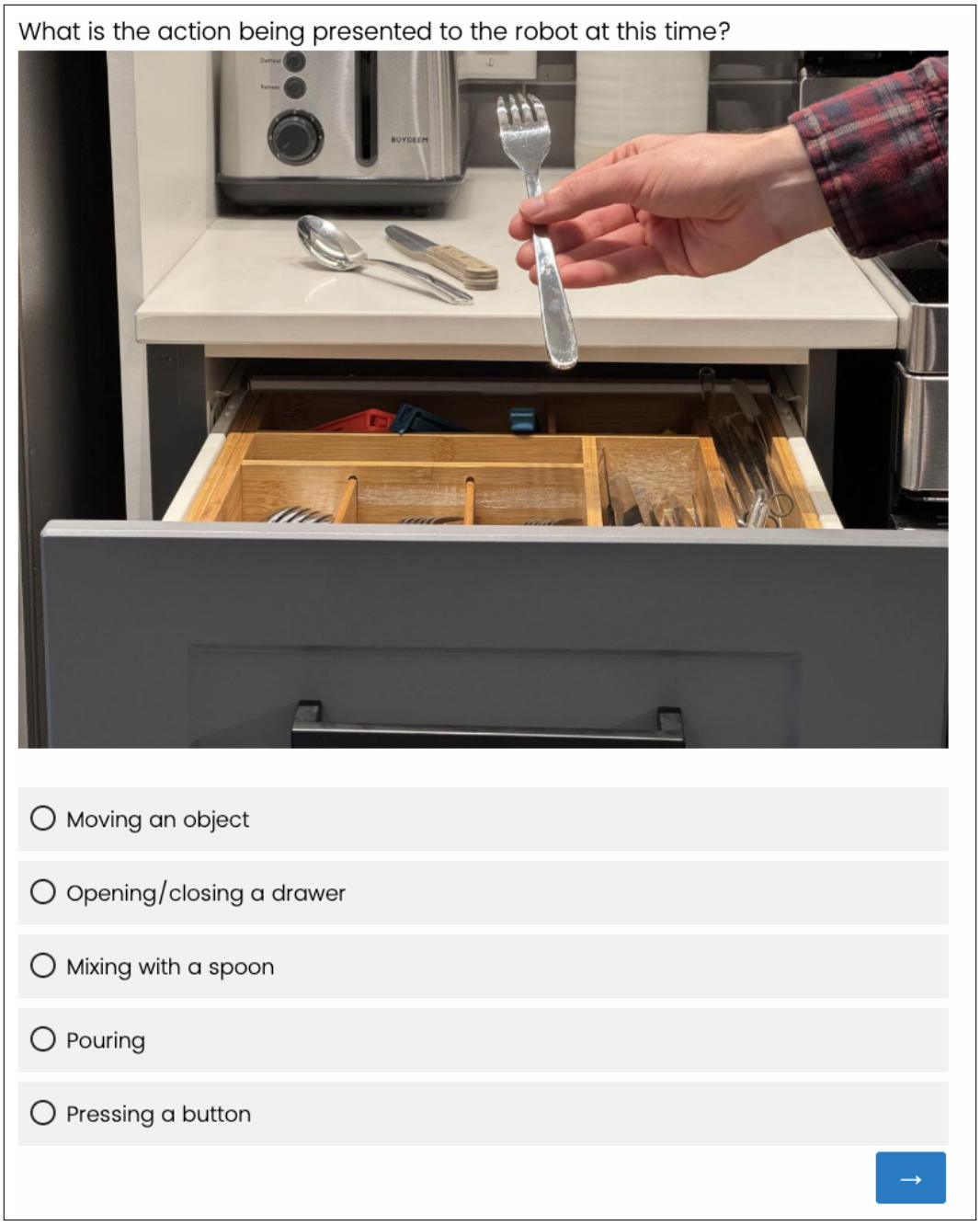


Figure C.3: [Cont.] The GHA about the scenario of a future advanced robot learning how to declutter a kitchen space - Page 2

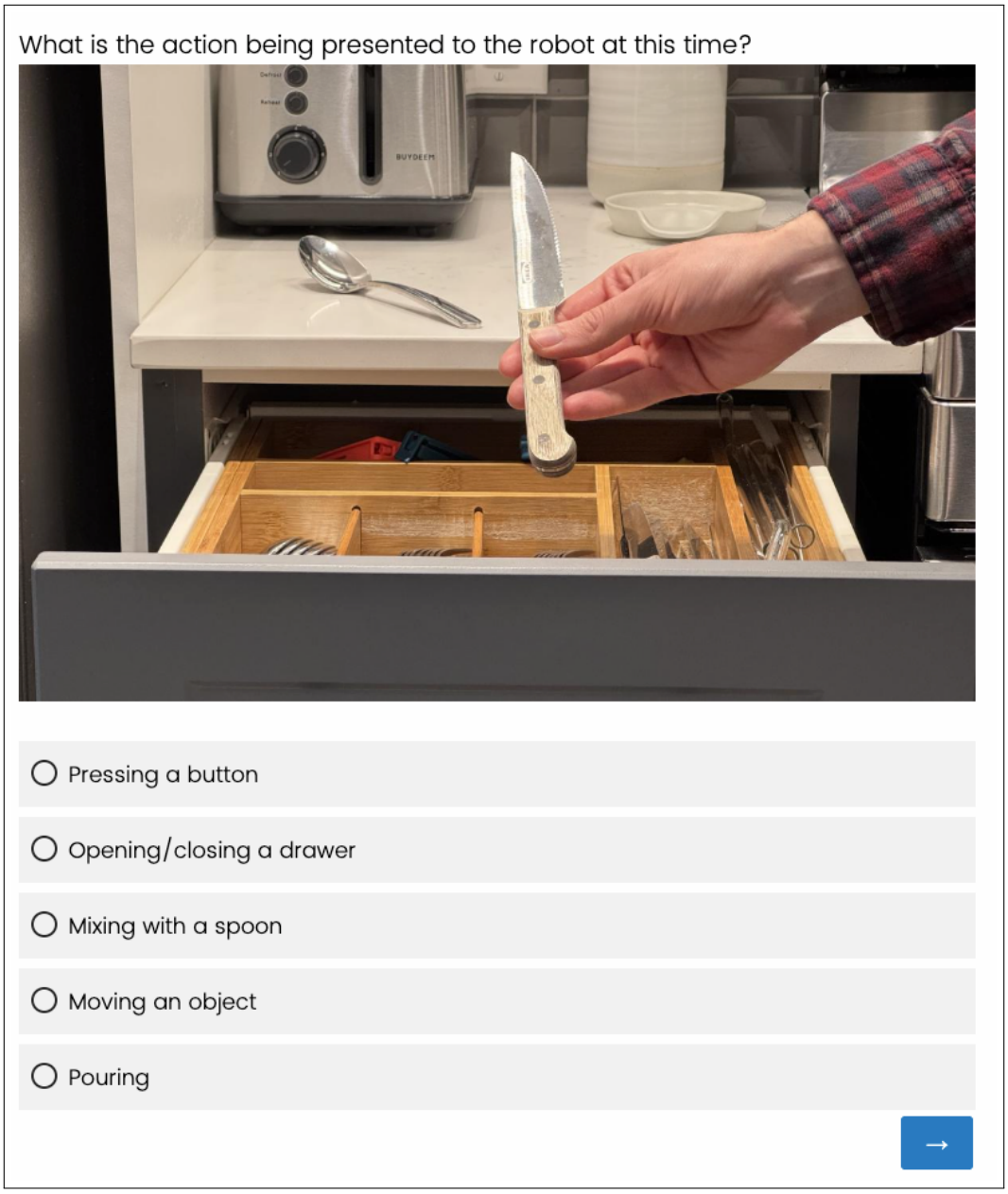


Figure C.3: [Cont.] The GHA about the scenario of a future advanced robot learning how to declutter a kitchen space - Page 3

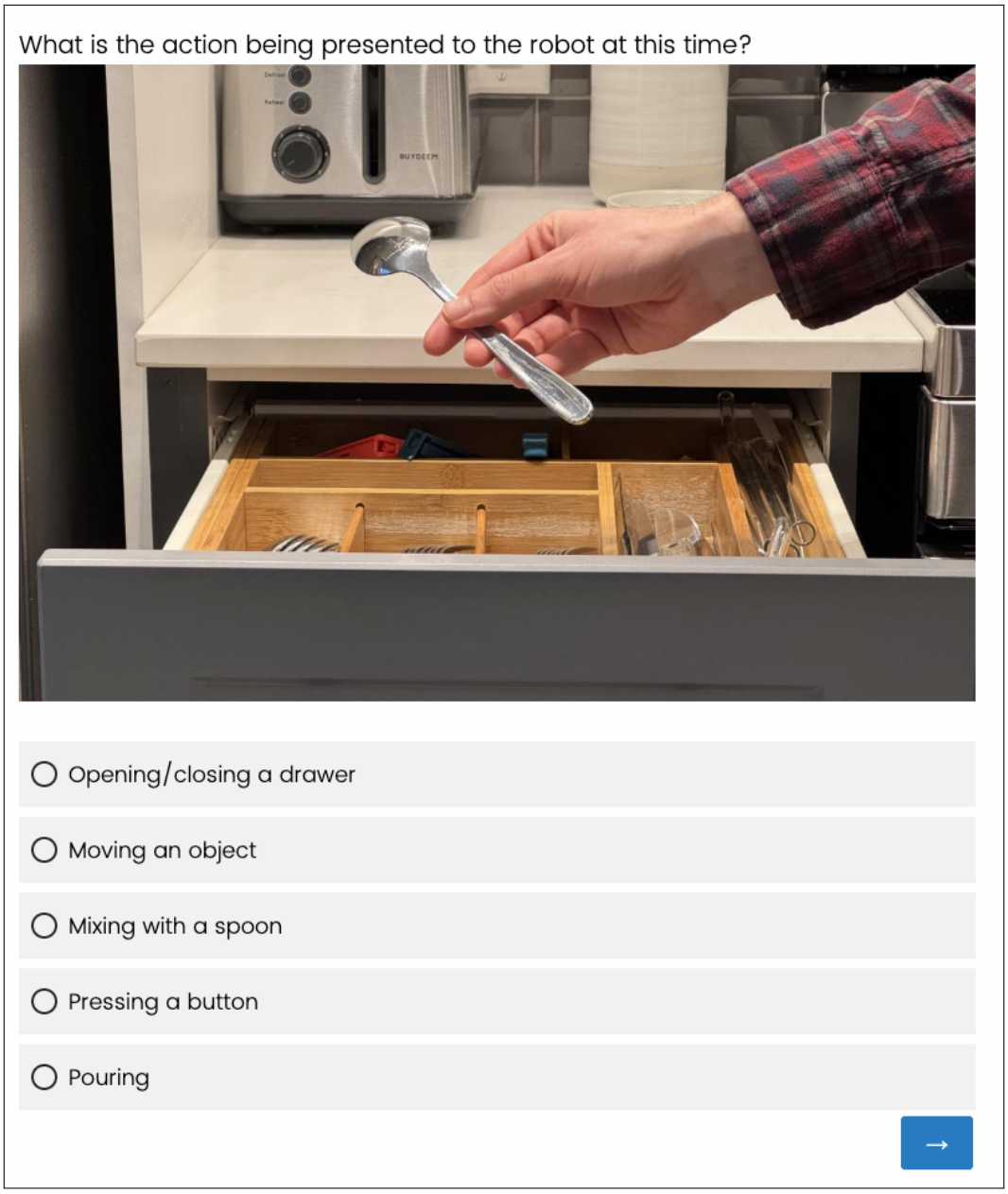


Figure C.3: [Cont.] The GHA about the scenario of a future advanced robot learning how to declutter a kitchen space - Page 4

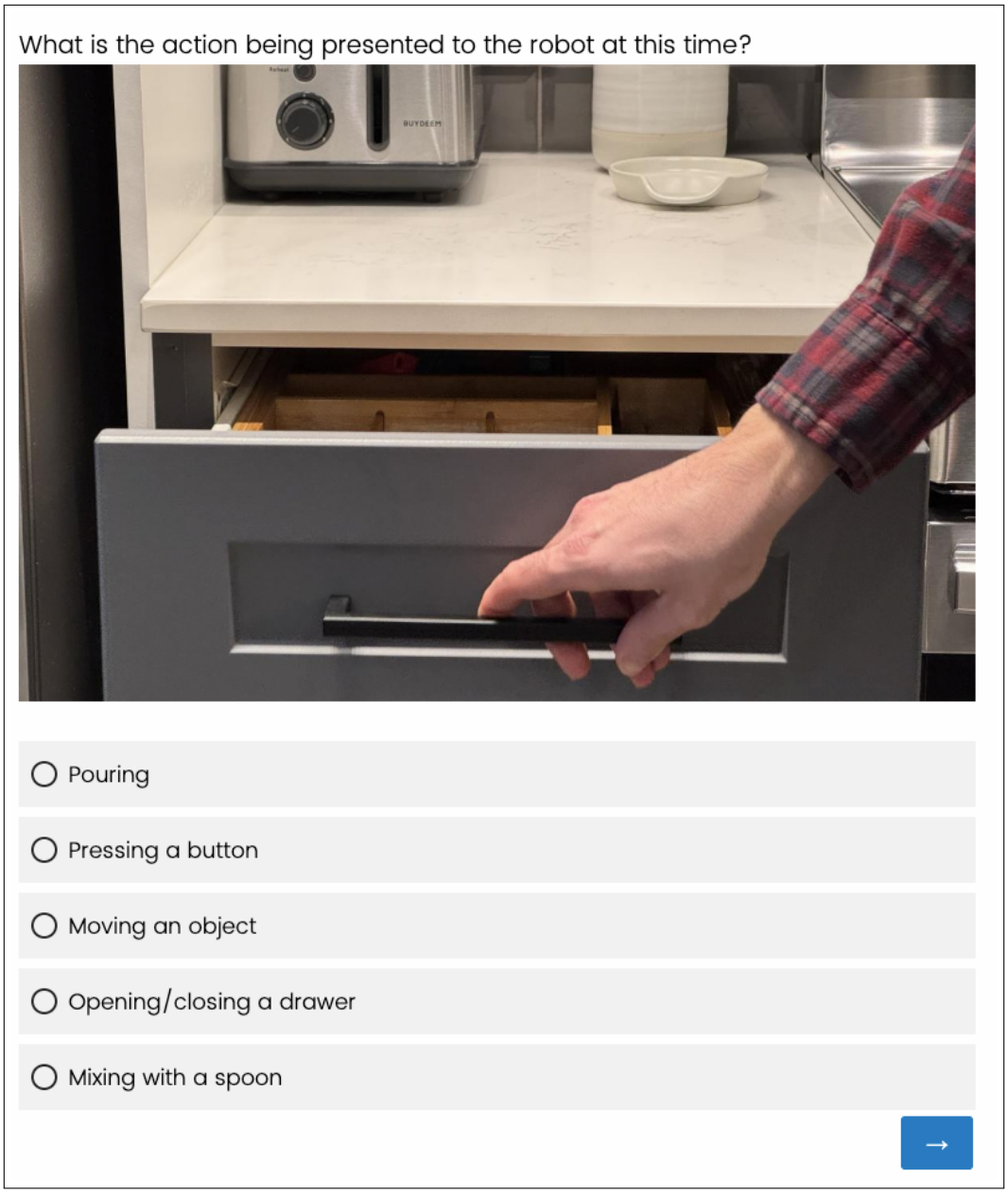


Figure C.3: [Cont.] The GHA about the scenario of a future advanced robot learning how to declutter a kitchen space - Page 5

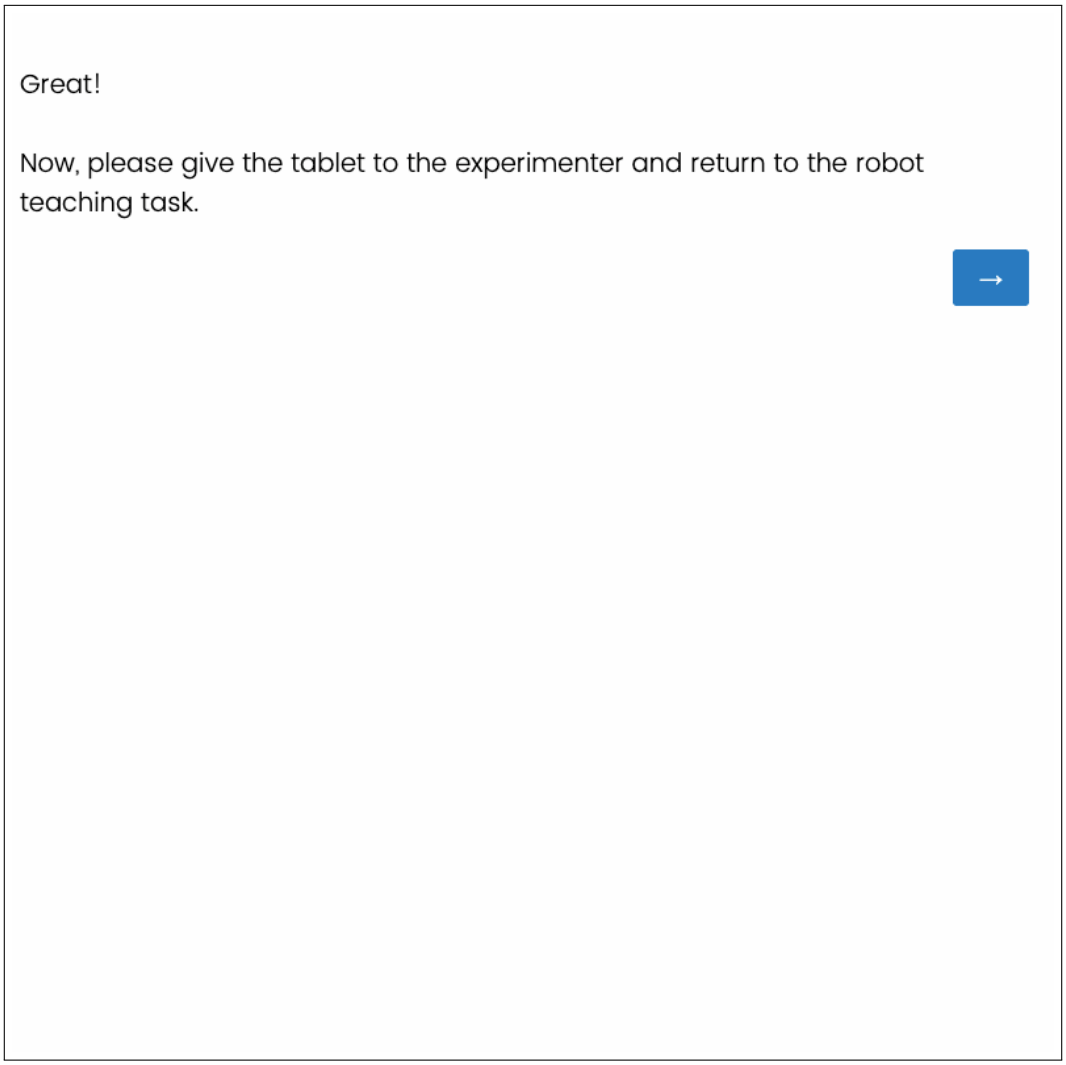


Figure C.3: [Cont.] The GHA about the scenario of a future advanced robot learning how to declutter a kitchen space - Ending page

C.2 Online Interface for Collecting Questionnaire Data



Human-robot teaching study

Research Team:

Principal Investigator: Professor Kerstin Dautenhahn, Canada 150 Research Chair in Intelligent Robotics, Departments of Electrical and Computer Engineering/Systems Design Engineering, University of Waterloo, Canada
Co-Investigator: Professor Christopher Nehaniv, Departments of Systems Design Engineering/Electrical and Computer Engineering, University of Waterloo, Canada
Co-Investigator: Professor Moojan Ghafurian, Department of Systems Design Engineering, University of Waterloo, Canada
Student Investigator: Pouyra Alasghari, Department of Electrical and Computer Engineering, University of Waterloo, Canada

To help you make an informed decision regarding your participation, this page will explain what the study is about, the possible risks and benefits, and your rights as a research participant. If you do not understand something, please feel free to ask one of the investigators before consenting to the study.

Purposes and objectives of the study: You are invited to participate in a research study about social learning in robotics. With the rapid advancements occurring in robotic technologies, social robots will play a prominent role in the future of our world. These robots should include sophisticated social capabilities to be accepted as our natural interaction partners. This study sets out to understand human-robot teaching more deeply, with a goal to improve robots' learning abilities in the future.
This experiment will be a part of a PhD dissertation that will be submitted for the degree compilation of student investigator, Pouyra Alasghari.

I. Your responsibilities as a participant

Participation in the study will consist of 1 session scheduled according to your convenience, which will approximately take 30-45 minutes. This session will be conducted in the Social and Intelligent Robotic Research Laboratory, E7, 5th floor. Before your session, we will ask you to give formal consent, complete a demographic information form asking about your gender, age, highest degree, as well as your previous experience with robots. In the session, you will be asked to teach a few simple tasks to our robot.

Procedure: You will be asked to stand in front of a humanoid robot called iCub and show it how to perform 3 simple household tasks, 5 times each. These tasks are presented in an abstract form and are placed on a table between you and the robot. The robot will give you the instructions on when to start, but you need to first place task objects in front of the robot and say 'ready'. Once the robot confirms, you can start teaching and you need to say 'done' or 'finished' to end teaching. After teaching each task, the robot will perform what it has learnt from you. Please be careful to avoid touching the robot before and after teaching periods.
Please listen to the robot carefully and follow its instructions. At the end of each session, we will ask you to complete a questionnaire that asks about your experience in the teaching session.

Participation criteria: The study will involve up to 100 participants and is accessible to University of Waterloo students. To be eligible to participate, you should not have any special knowledge or background in humanoid robotics or Human-Robot Interaction. This means that you should not be a robotics or Human-Robot Interaction researcher. That is why we will ask about your department, current program, and/or field of study in the demographic information form. People who have implanted medical devices should not participate in the study as the proper functionality of the devices could be impaired during the experiment. You will see a "Screening Questionnaire" next for more details and confirm those criteria. In case you are not eligible, your screening data will be deleted.

II. Your rights as a participant

Your participation in this study is voluntary. You may decide to leave the study at any time by communicating this to the researcher. You may decline to answer any question(s) in the questionnaire you prefer not to answer by leaving them blank. You can request your data be deleted after the session by contacting us.

Remuneration: In appreciation for your time, you will receive the amount of CAD $15 loaded to your WatCard by approximately two weeks after your session. If you withdraw earlier, you will receive prorated amount: a base amount of CAD $7.5 plus a fraction of CAD $7.5 depending on the amount of time you spend on the study prior to withdrawal. For example, if you withdrew halfway through the study you will receive the amount of CAD $(7.5+3.75) = CAD $11.25 on your WatCard.

Benefits of the study: Participation in this study may not provide any personal benefit to you. However, the study will benefit the academic community and the research team. We plan to publish the results of this study, and researchers in the field of social robotics can potentially benefit from that. We hope the outcomes of this experiment pave the way for designing more intelligent robotic systems.

Risks associated with participation: There could be a potential risk of the robot's moving parts colliding physically with your body if you come very close to the robot while it is moving. To avoid those risks, you should sit as instructed by the researchers conducting the study and wait for their prompts to move. During the experiment, the robot will be stopped in case of any detected potential collision between the robot and you, any unexpected faults in the behaviour of the robot, or if you stand up. Please let us know if you feel uncomfortable during the experiment. The researchers will be ready to take action in that situation.

Confidentiality of identities and data: You will be participating in a study that uses Qualtrics. Qualtrics has implemented technical, administrative, and physical safeguards to protect the information from loss, misuse, and unauthorized access, disclosure, alteration, or destruction. However, no internet transmission is ever fully secure or error free. Your names will be replaced with a numeric study code, and will not appear in any report, presentation or publication resulting from this research. The connection between your identity and research data that is confidential. The study data will be kept for a minimum of 7 years, and only the researchers associated with this project will have access to it. Participant data will be stored on a secure lab server, accessed by the researchers associated with this project for the purposes of this study. Video footage will not be shared outside the research team without your consent. Given that we want to explore the facial expressions of the participants, and therefore, videos need to be identifiable during the storage, videos will be encrypted. We will delete the videos as soon as they are no longer needed. Still, if you give permission for the videos to be shared publicly, all identifying features will be blurred. The data captured will be securely stored on a secure password-protected lab server with access only for the researchers. The lab servers are secure and created by the University of Waterloo. Immediately after the session, the data files will be uploaded to that server and will be removed from the local computer. You can withdraw from the study at any time by informing the researcher. However, your data cannot be withdrawn after submission of papers for publication, anticipated by May 2025. If you decide to withdraw in any part, we will delete all the files and data corresponding to you. Your identity will remain confidential even in that case. The data collected in this study could be used with your permission in future research (i.e., related research, follow-up studies, etc.) by the listed researchers. The anonymized data collected through this study (i.e., robot sensory data, videos with blurred/removed participant faces, and questionnaire responses) may be shared with your permission in an online repository or archive (e.g., GitHub) for use by other researchers. You may withdraw your permission for any of the above items or withdraw your study participation without penalty by advising the researchers. It should be noted that other researchers utilizing this anonymized published data may not be subject to the Tri-Council Policy Statement on research ethics and they may use this anonymized published data in unspecified research.

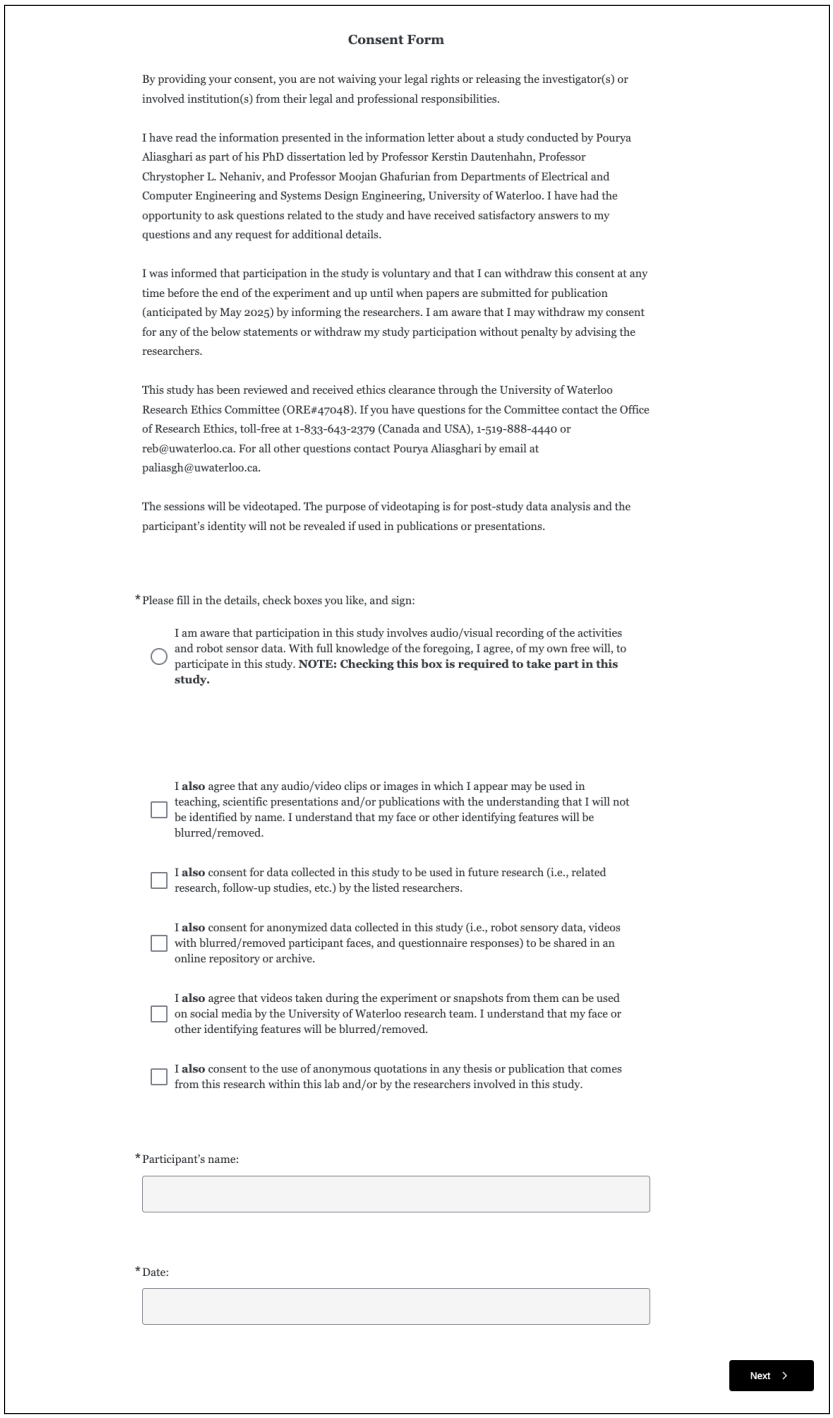
III. Questions, comments, or concerns

Ethics clearance: This study has been reviewed and received ethics clearance through a University of Waterloo Research Ethics Board (REB #47048). If you have questions for the Board, contact the Office of Research Ethics toll-free at 1-833-643-2379 (Canada and USA), 1-519-888-4440 or reb@uwaterloo.ca

Contact Information: If you have any questions regarding this study, or would like additional information about your participation, please contact Pouyra Alasghari by email at palasgh@uwaterloo.ca.

Continue to consent form

Figure C.4: The pre-session online module - Information page. Zoom in for readable text.



Consent Form

By providing your consent, you are not waiving your legal rights or releasing the investigator(s) or involved institution(s) from their legal and professional responsibilities.

I have read the information presented in the information letter about a study conducted by Pourya Aliasghari as part of his PhD dissertation led by Professor Kerstin Dautenhahn, Professor Chrystopher L. Nehaniv, and Professor Moojan Ghafurian from Departments of Electrical and Computer Engineering and Systems Design Engineering, University of Waterloo. I have had the opportunity to ask questions related to the study and have received satisfactory answers to my questions and any request for additional details.

I was informed that participation in the study is voluntary and that I can withdraw this consent at any time before the end of the experiment and up until when papers are submitted for publication (anticipated by May 2025) by informing the researchers. I am aware that I may withdraw my consent for any of the below statements or withdraw my study participation without penalty by advising the researchers.

This study has been reviewed and received ethics clearance through the University of Waterloo Research Ethics Committee (ORE#47048). If you have questions for the Committee contact the Office of Research Ethics, toll-free at 1-833-643-2379 (Canada and USA), 1-519-888-4440 or reb@uwaterloo.ca. For all other questions contact Pourya Aliasghari by email at paliasgh@uwaterloo.ca.

The sessions will be videotaped. The purpose of videotaping is for post-study data analysis and the participant's identity will not be revealed if used in publications or presentations.

* Please fill in the details, check boxes you like, and sign:

- ○ I am aware that participation in this study involves audio/visual recording of the activities and robot sensor data. With full knowledge of the foregoing, I agree, of my own free will, to participate in this study. **NOTE: Checking this box is required to take part in this study.**
- ☐ I **also** agree that any audio/video clips or images in which I appear may be used in teaching, scientific presentations and/or publications with the understanding that I will not be identified by name. I understand that my face or other identifying features will be blurred/removed.
- ☐ I **also** consent for data collected in this study to be used in future research (i.e., related research, follow-up studies, etc.) by the listed researchers.
- ☐ I **also** consent for anonymized data collected in this study (i.e., robot sensory data, videos with blurred/removed participant faces, and questionnaire responses) to be shared in an online repository or archive.
- ☐ I **also** agree that videos taken during the experiment or snapshots from them can be used on social media by the University of Waterloo research team. I understand that my face or other identifying features will be blurred/removed.
- ☐ I **also** consent to the use of anonymous quotations in any thesis or publication that comes from this research within this lab and/or by the researchers involved in this study.

* Participant's name:

* Date:

Next >

Figure C.4: [Cont.] The pre-session online module - Consent from. Zoom in for readable text.

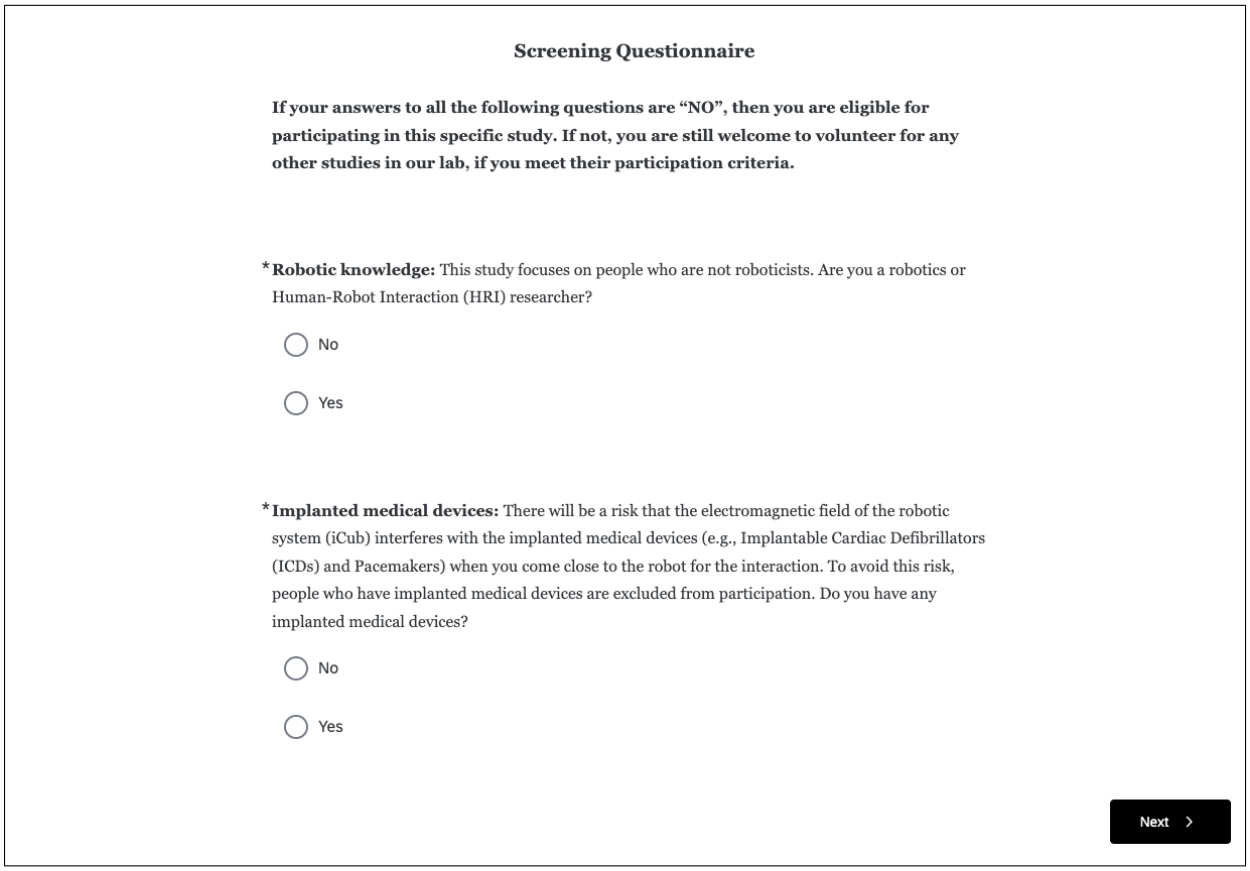


Figure C.4: [Cont.] The pre-session online module - Screening questionnaire. Zoom in for readable text.

Demographic Information Form

Please provide a response for the following questions:

Age:

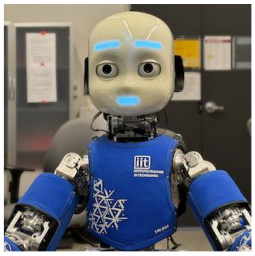
Gender:

Department, current program and/or field of study:

Have you ever programmed any robot?

Have you ever interacted with a human-like robot?

Have you ever interacted with an iCub robot (pictured below) in particular?



Have you ever participated in a study and taught something to a robot?

Done >

Figure C.4: [Cont.] The pre-session online module - Demographic information form. Zoom in for readable text.

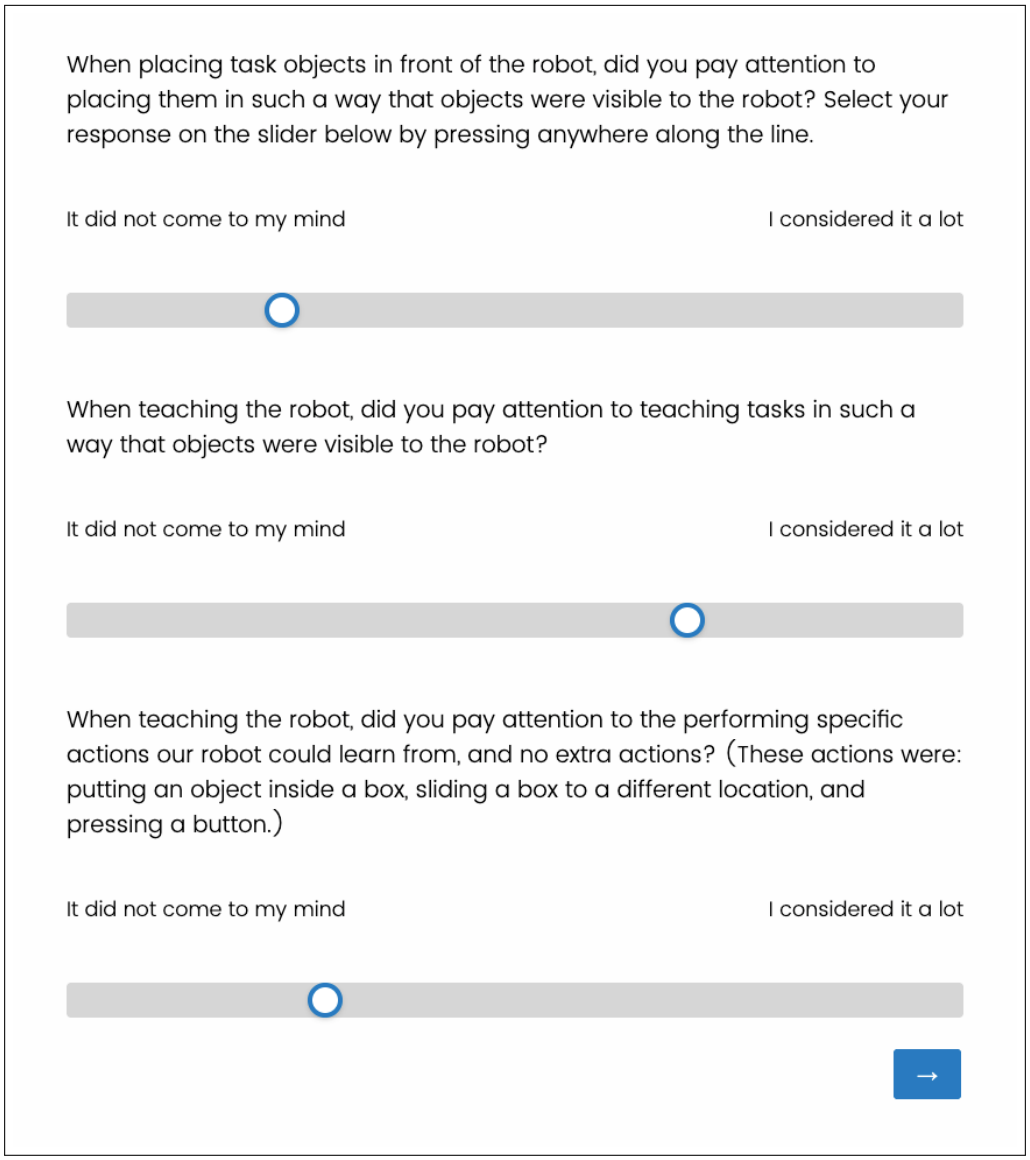


Figure C.5: The ending questionnaire module - Q1-3, as noted in Table 5.2. Random responses are selected as an example.

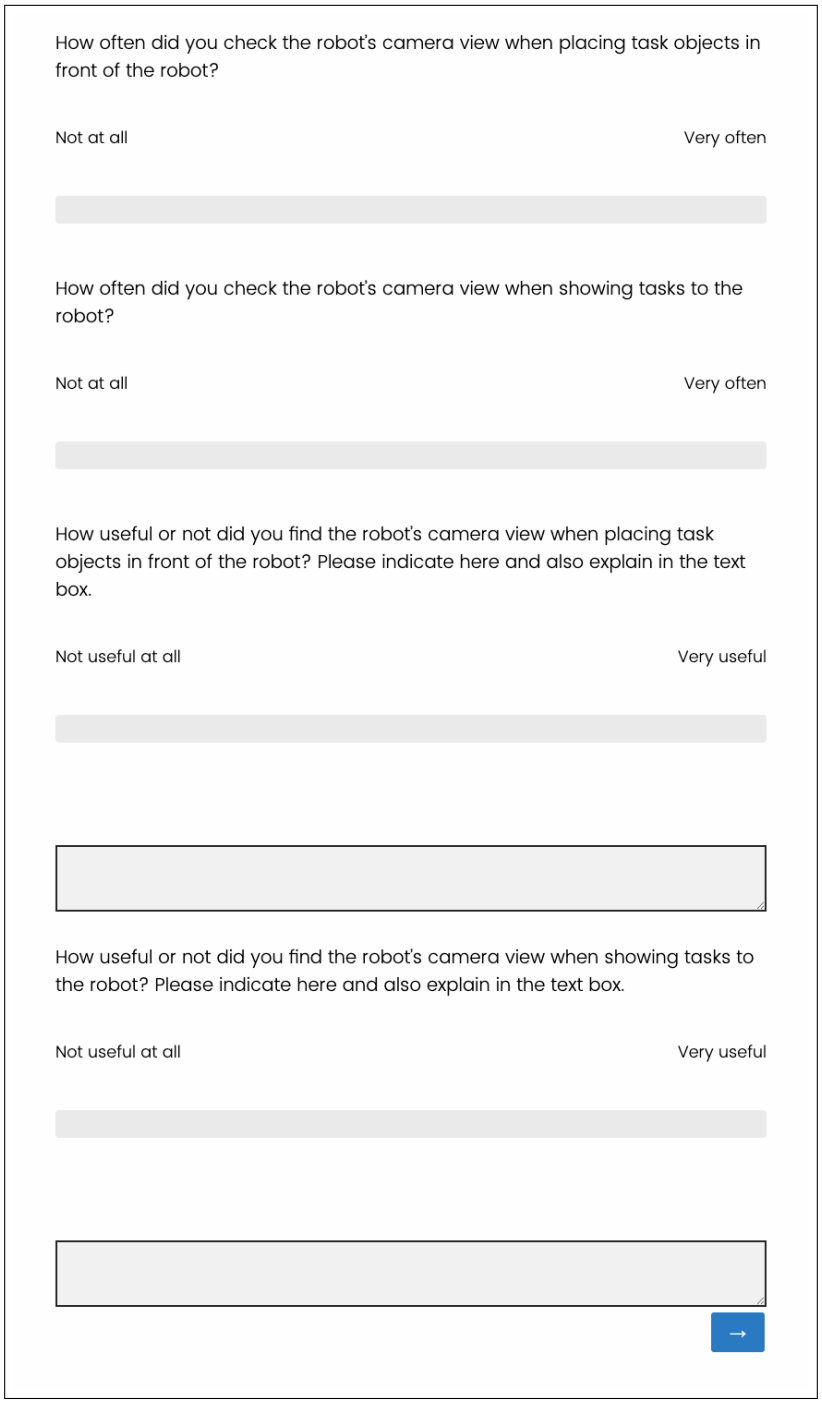


Figure C.5: [Cont.] The ending questionnaire module - Q4-7, as noted in Table 5.2. Only participants in RCV condition answered this part. Zoom in for readable text.

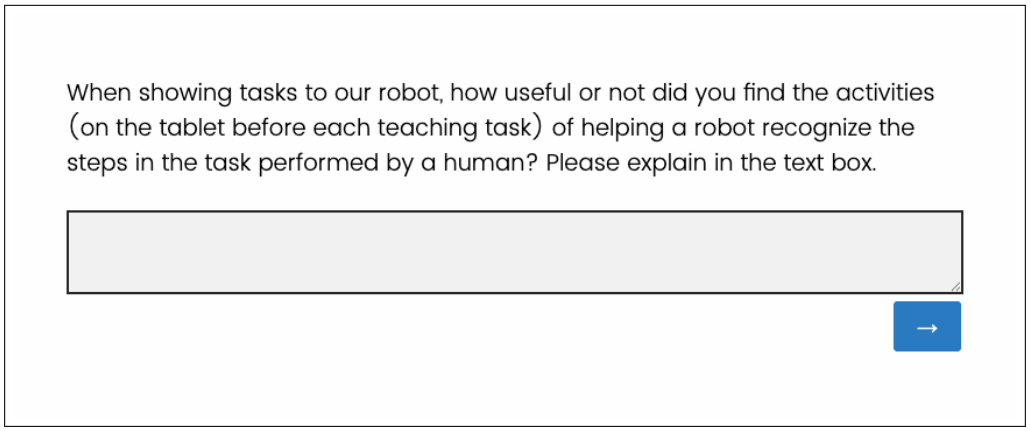


Figure C.5: [Cont.] The ending questionnaire module - Q8, as noted in Table 5.2. Only participants in GHA condition answered this part.

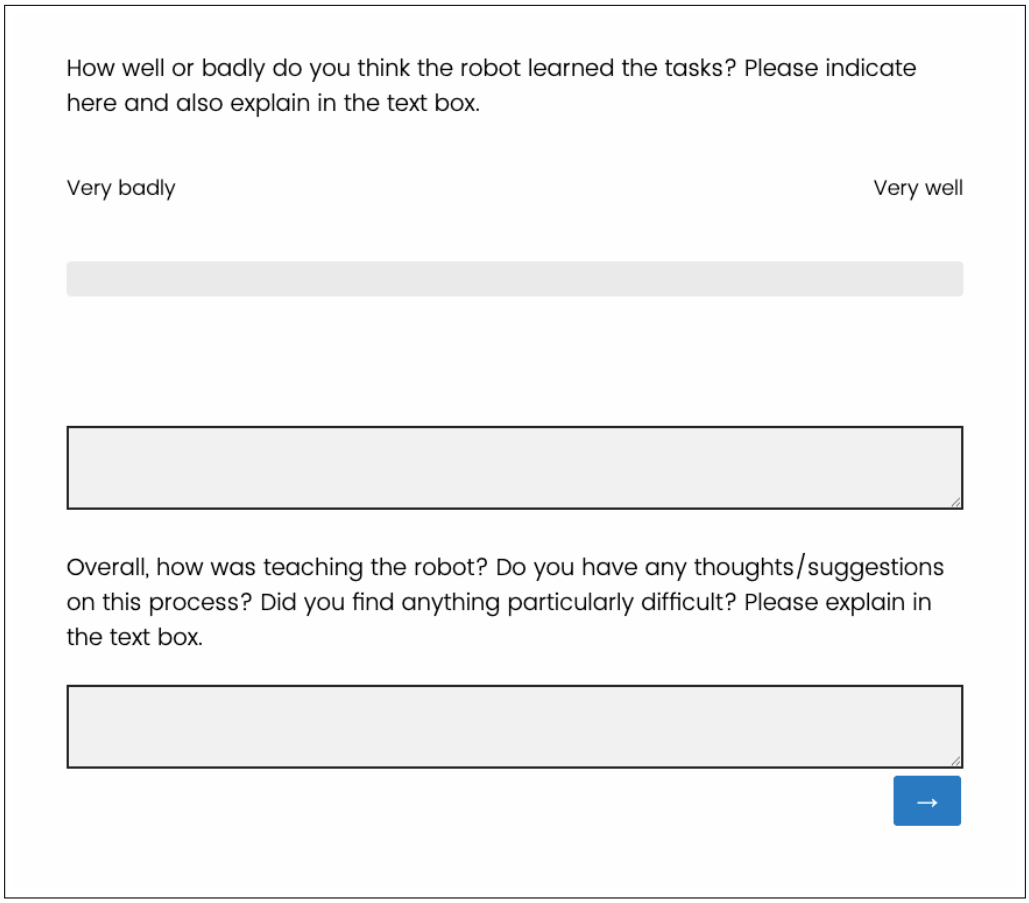


Figure C.5: [Cont.] The ending questionnaire module - Q9-10, as noted in Table 5.2

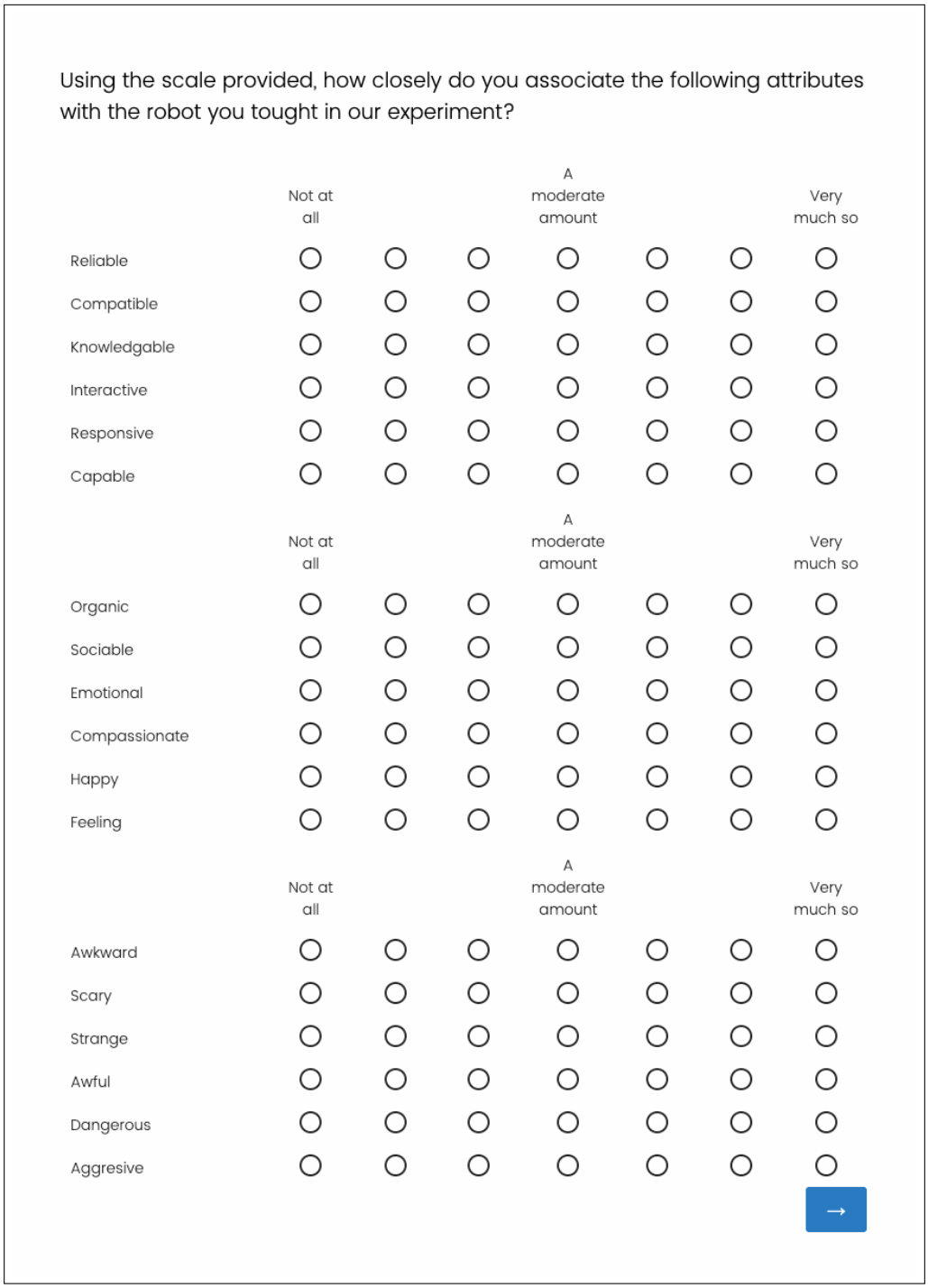


Figure C.5: [Cont.] The ending questionnaire module - RoSAS questionnaire [146]

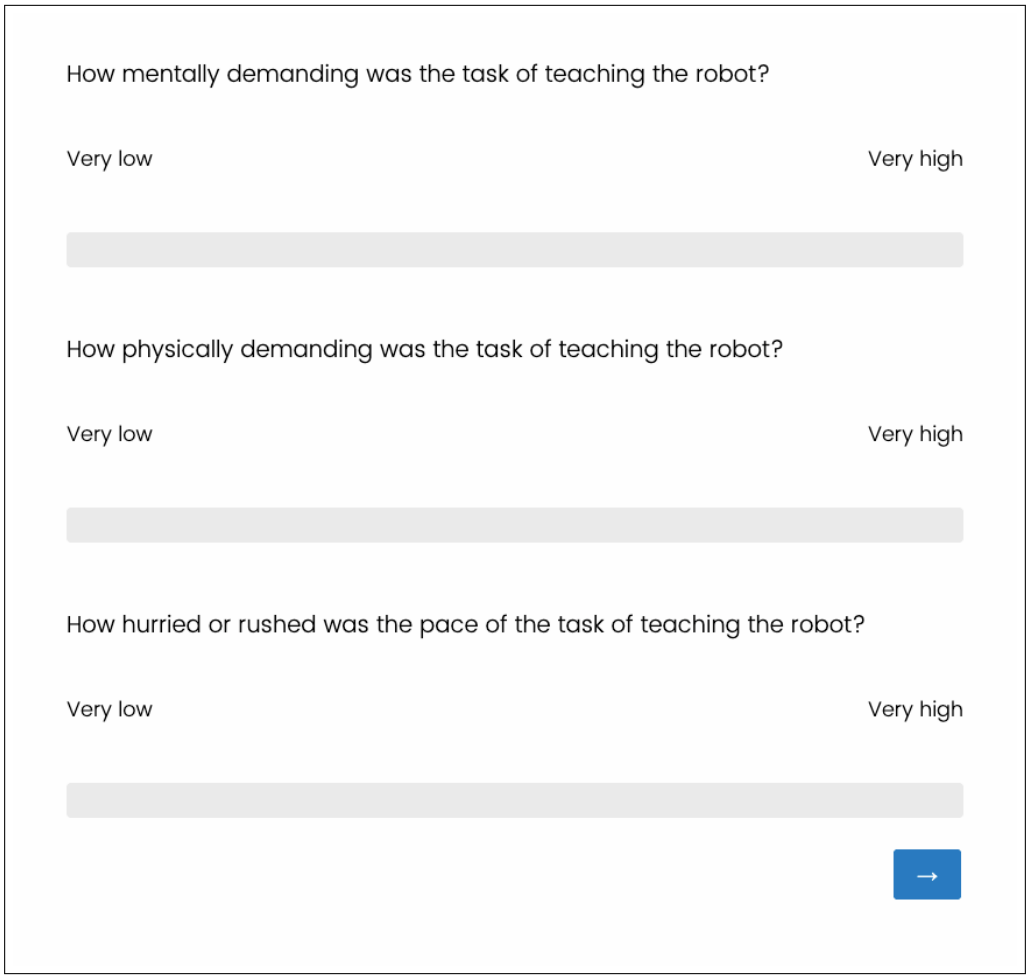


Figure C.5: [Cont.] The ending questionnaire module - NASA-TLX questionnaire [85], Part 1

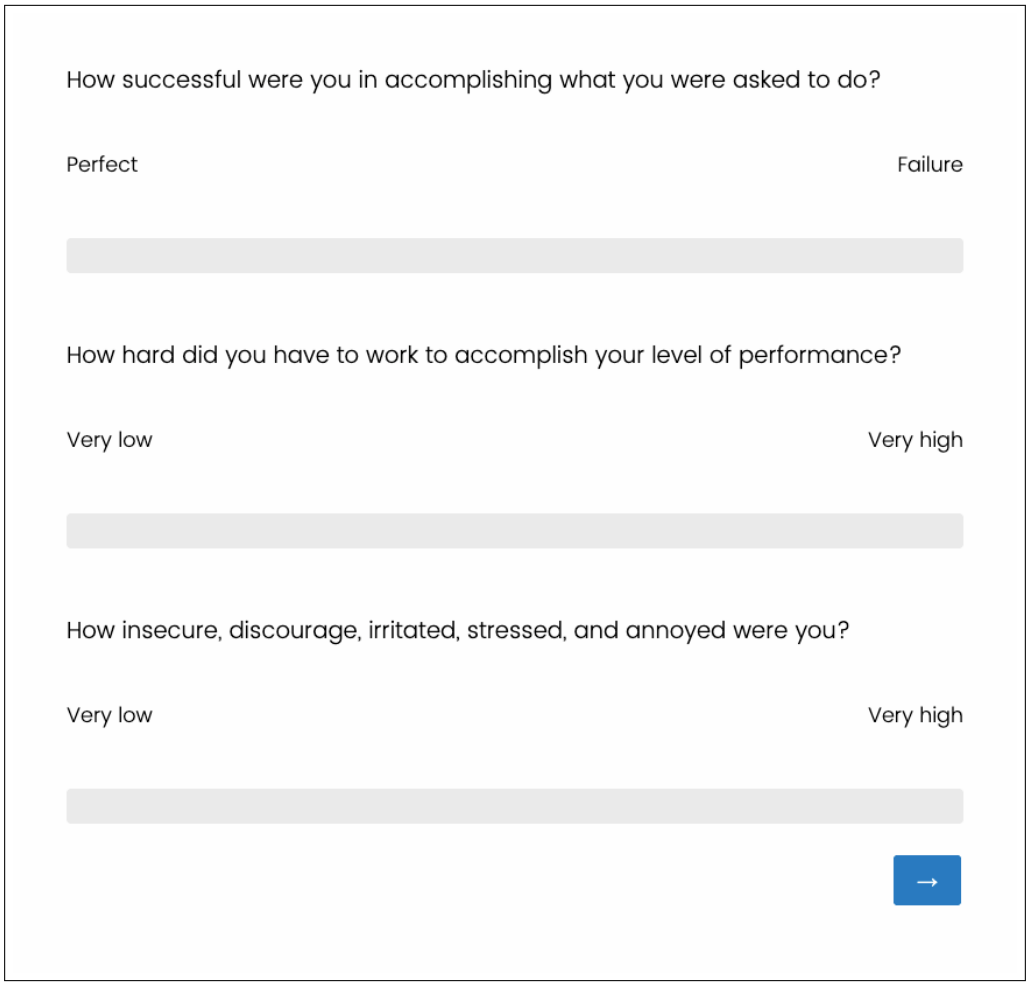


Figure C.5: [Cont.] The ending questionnaire module - NASA-TLX questionnaire [85], Part 2

C.3 Other Experimental Material and the Setup



Figure C.6: The flyer used for advertising the experiment, posted on a public board. Students interested in participating could either send an email directly to myself or scan the QR code to view additional information and book their session through Outlook calendar, selecting from available time slots.



Figure C.7: The setup in the experimental room, before the arrival of participants. On the left side, the setup for task learning of the iCub robot in *RCV* condition can be found, as was seen in Figure 5.2. Floor lamps were placed next to the robot to improve the visibility of the tasks in varying lighting conditions, especially due to running the experiment during winter. The robot was supported by a crane for safety reasons, to prevent it from falling in case an emergency stop was needed (fortunately, no such event occurred throughout the experiment). In the middle, the door connecting the experimental room to the control room is visible. On the right side, the abstract tasks and the two sets of objects involved in each task (for robot learning and task performance) are stored for easy switching during the session.

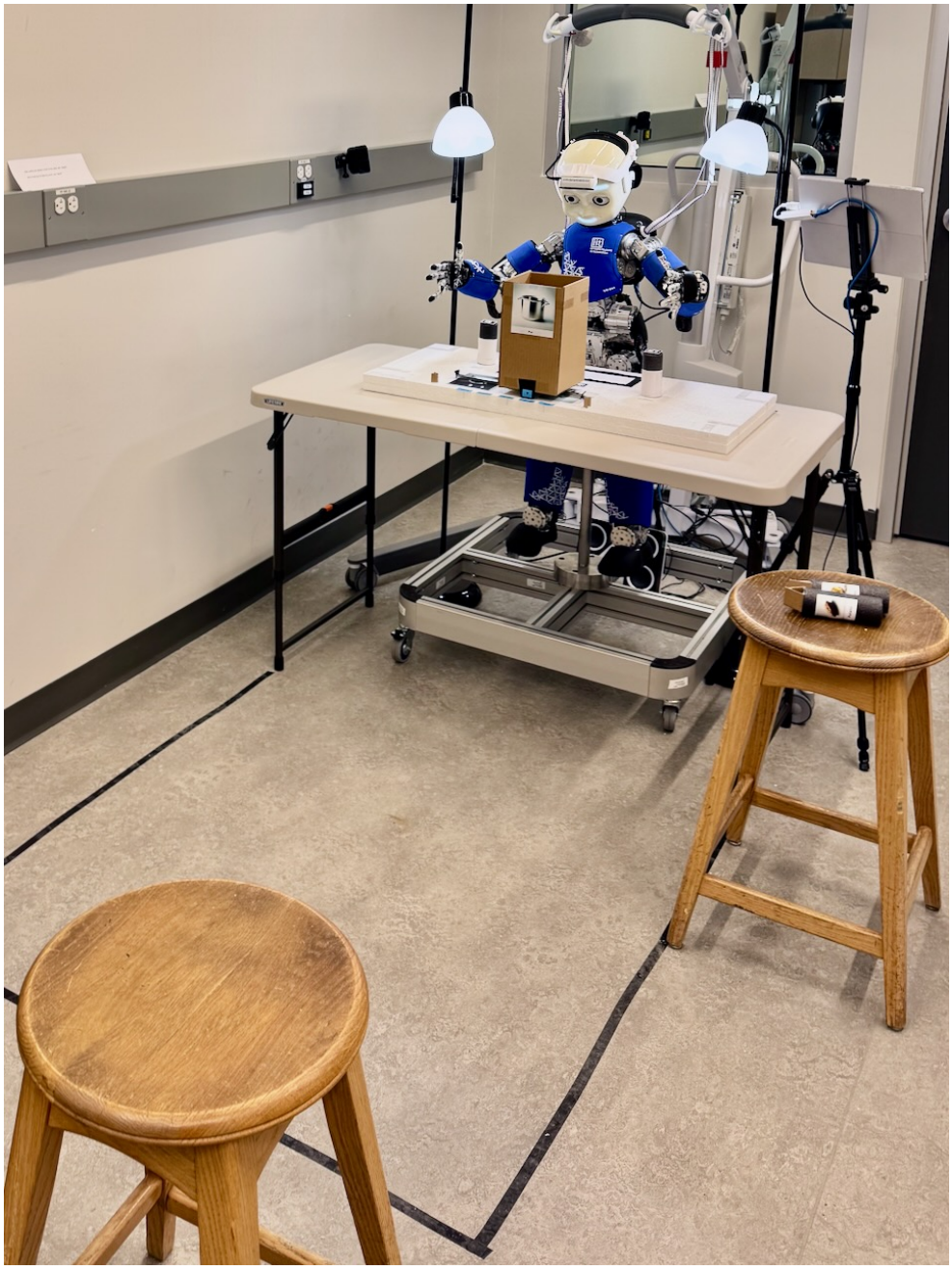


Figure C.8: Another view of the experimental setup for task learning of the iCub robot in *No-RCV* condition, as also shown in Figure 5.2. The stool at the bottom left of the photo was for participants to sit and watch the robot when it was performing the tasks. The area for participants to stand and teach the robot was marked with black tape on the floor.