Learning Socially Assistive Robot Behaviors for Personalized Human-Robot Interaction

by

Christina Moro

A thesis submitted in conformity with the requirements for the degree of Master of Applied Science
Department of Mechanical and Industrial Engineering
University of Toronto

© Copyright by Christina Moro 2018
Learning Socially Assistive Robot Behaviors for Personalized Human-Robot Interaction

Christina Moro
Master of Applied Science
Department of Mechanical and Industrial Engineering
University of Toronto
2018

Abstract

Caregivers play a crucial role in assisting seniors having difficulty accomplishing activities of daily living (ADLs) due to physical or cognitive limitations. A global decline in the caregiver-to-senior ratio is making it increasingly more difficult to care for these seniors. Socially assistive robots are promising alternative technologies for supporting seniors in living independently. However, limited research has gone into developing a learning-based method for designing assistive robot behaviors. This thesis aims to: (1) identify the key features necessary for assistive robots supporting seniors with cognitive impairments in completing ADLs; and (2) develop a novel behavior-learning architecture to teach robots how to display assistive behaviors using expert demonstrations and personalize these learned behaviors to the senior’s cognition using reinforcement learning to increase task performance. Experiments with a socially assistive robot validated the robot’s ability to learn and personalize new behaviors to a user’s cognition from expert demonstration using the proposed architecture.
Acknowledgments

I would like to thank my supervisor, Professor Goldie Nejat, for her mentorship during my Master’s degree, for constantly making herself available to answer questions and review my writing at all hours of the day or night, and for providing me with numerous opportunities for personal and professional growth. I also thank my co-supervisor, Prof. Alex Mihailidis, for his continuous support, for helping me brainstorm areas of research, and for always picking up the phone regardless of where he was traveling. Thank you to my committee members for taking time out of their days to review this thesis and provide feedback on my research.

Thank you to Ruojing Song, Carrie Yang, and Shayne Lin for their valuable contributions to this project. Thanks to all my lab mates for creating an awesome work environment and for helping me learn something new every day. Thanks to Essam for supporting me in putting this thesis together, and to Alex Hong for always making himself available to go climbing when we needed to de-stress. And lastly, a special thank you goes out to my mom and dad, who have both dedicated so much of their time, energy, and support throughout these past two years. This Master’s degree has been a rollercoaster of a ride; couldn’t have done it without all these incredible people cheering me on!
# Table of Contents

Acknowledgments ........................................................................................................ iii

Table of Contents ......................................................................................................... iv

List of Tables ............................................................................................................... vii

List of Figures ............................................................................................................. viii

Chapter 1 Introduction ................................................................................................. 1
  1.1 Motivation ............................................................................................................. 1
  1.2 Socially Assistive Robots for Eldercare ............................................................... 2
    1.2.1 Social Features to promote Human-Robot Interactions ................................. 2
    1.2.2 Development of Socially Assistive Robot Behaviors ................................. 3
  1.3 Thesis Objectives ............................................................................................... 5
  1.4 Proposed Methodology ..................................................................................... 5
    1.4.1 Literature Review ......................................................................................... 5
    1.4.2 Investigating the Influence of Communication Modes in Senior-Robot Interaction .................................................................................. 5
    1.4.3 An Architecture for Learning Socially Assistive Robot Behaviors ............. 6
    1.4.4 Behavior Learning Experiments .................................................................... 6
    1.4.5 Conclusion .................................................................................................... 6

Chapter 2 Literature Review ....................................................................................... 7
  2.1 Socially Assistive Robots .................................................................................. 7
    2.1.1 Preferences for Socially Assistive Robots .................................................... 7
  2.2 Socially Assistive Robot Behaviors .................................................................. 9
    2.2.1 Teaching Robot Behaviors through Demonstration ...................................... 10
    2.2.2 Teaching Robot Behaviors through Reinforcement Learning ...................... 11
    2.2.3 Learning Behaviors Using Both LfD and RL ................................................ 13

Chapter 3 Investigating the Influence of Communication Modes in Senior-Robot Interaction ...................................................................................................................... 15
3.1 User Study Goals ........................................................................................................... 15
  3.1.1 Study Hypotheses ................................................................................................. 15
3.2 User Study Design ....................................................................................................... 16
  3.2.1 Participants ........................................................................................................ 16
  3.2.2 The Interactive Platforms .................................................................................. 17
  3.2.3 The Tea-making Activity .................................................................................... 19
  3.2.4 Procedure .......................................................................................................... 20
  3.2.5 Evaluation Metrics ............................................................................................. 21
3.3 Results and Discussion ................................................................................................. 24
  3.3.1 Engagement levels during interaction ................................................................. 26
  3.3.2 Perceived social intelligence across platforms ................................................... 27
  3.3.3 Perceived usefulness across platforms ................................................................ 28
  3.3.4 Technology preferences ...................................................................................... 29
3.4 Chapter Summary ........................................................................................................ 30

Chapter 4 An Architecture for Learning Socially Assistive Robot Behaviors .................... 31
  4.1 Proposed Robot Behavior Learning Architecture Design ........................................ 31
    4.1.1 Learning Behaviors from Demonstration ......................................................... 32
    4.1.2 Personalizing Behaviors Using Reinforcement Learning ................................. 35
  4.2 Chapter Summary ...................................................................................................... 39

Chapter 5 Behavior Learning Experiments ........................................................................ 40
  5.1 Implementation on the Casper Robot ....................................................................... 40
    5.1.1 Gesture Imitation ............................................................................................ 40
    5.1.2 Speech Imitation ............................................................................................. 43
  5.2 Demonstration Study Setup ....................................................................................... 43
    5.2.1 Participants ...................................................................................................... 43
    5.2.2 Tea-Making Activity ....................................................................................... 43
List of Tables

Table 1 Pre-study questionnaire results showing experience with computers and robots, and initial robot preferences ................................................................. 17

Table 2 Measured variables obtained from video data ......................................................... 21

Table 3 Semi-structured interview questions for each platform ......................................... 24

Table 4 Summary of Friedman test results for video data variables .................................... 25

Table 5 The assistive tea-making behaviors demonstrated by our participants .................. 45

Table 6 Questions asked at the end of the demonstration session ....................................... 47

Table 7 Mean true positive rate (TPR), false negative rate (FNR), and false positive rate (FPR) ................................................................................................................. 49

Table 8 Labeling distribution for speech assertiveness and movement activity .................... 51

Table 9 Reward distribution based on the robot state, user functioning state, and user activity state .............................................................................................................. 52

Table A1 Descriptive statistics of measured variables from the video data for interactions with Casper (C), Ed (E), and the Tablet (T). The table provides the mean of the number of instances when a participant displayed the respective measured variable with each platform along with the standard deviation, as defined in section 3.2.5 Evaluation Metrics ....................................................... 69
List of Figures

Figure 1 The Ed robot .................................................................................................................. 18
Figure 2 The Casper robot ......................................................................................................... 18
Figure 3 The tablet display ....................................................................................................... 18
Figure 4 Excerpts from the six COACH videos showing each step in the tea-making task: (a) Turn the water on; (b) Pour water into the kettle; (c) Turn the water off; (d) Turn the kettle on; (e) Put the tea bag in the cup; and (f) Pour hot water into the tea cup .................................................. 19
Figure 5 Study setup showing interactions with: (a) the Casper robot; (b) the Ed robot; and (c) the tablet display in the kitchen environment with the tea-making items ........................................................................ 20
Figure 6 Socially Assistive Robot Behavior Learning Architecture ........................................... 32
Figure 7 Example CART decision tree with four behaviors ...................................................... 35
Figure 8 The Casper robot and its kinematic model. Notation representations are: “t” for the torso, “h” for the head, “rs” for the right shoulder, “ls” for the left shoulder, “re” for the right elbow, and “le” for the left elbow .................................................................................................................................. 41
Figure 9 Visual representation of the demonstrator’s arm joint vectors mapped onto the robot and the corresponding calculated robot joint rotations ........................................................................ 42
Figure 10 Study setup displaying the Casper robot, sensors, and kitchen utilities ................ 44
Figure 11 Demonstration learning sessions with the teacher performing a demonstration (left) and the Casper robot replicating the demonstrated behavior (right). The behaviors demonstrated are (a) pointing to the sugar in the cabinet; (b) conversing socially with the user, e.g. asking about their day and health; (c) pointing to the box of tea on the counter .................................................. 46
Figure 12 The decision tree generated by the CART algorithm showing the splitting nodes as rectangles with their splitting attribute and node impurity, and the rounded nodes representing leaves showing the predicted class ........................................................................................................... 48
Figure 13 Cumulative reward for six user cognitive models with their own behavior label preferences.
Chapter 1
Introduction

1.1 Motivation

The global population is aging at a remarkable rate, with potentially serious implications on the quality of care available to seniors. In 2015, the United Nations reported that 12% of the world’s population was aged 60 years or more [1]. Globally, this percentage is expected to grow to nearly 25% by 2050 due to decreasing fertility rates and increased life expectancy. Along with population aging comes a significant decline in the caregiver-to-senior ratio [2]. Formal and informal caregivers provide critical assistance to aging seniors who often encounter challenges related to physical and cognitive decline, making it difficult for them to accomplish activities of daily living (ADL) independently. Examples of ADL include preparing meals, taking medication, and undertaking household chores [3]. Given the declining caregiver-to-senior ratio, the development of assistive technologies to support seniors with ADL is of primary importance.

Assistive technologies that enhance daily care can take on a variety of forms. For example, seniors suffering from cognitive decline may use signs and notices placed around the home, medication reminders, orientation devices, and alarms [4], which are sometimes used in conjunction with cell phones [5]. Wearable devices are also used by some older adults for fall detection and daily health monitoring [6]. Most recently, the use of smart home technologies to monitor the home environment settings as well as a person’s health has become more prevalent [4], [6], [7]. However, these technologies generally have limited intervention capabilities, and require the senior to be pro-active in their use. For example, the use of a tablet requires a senior to actively search for it or remember its location, and then open the respective application they would like to use. Seniors with cognitive impairments may not always be capable of using such devices. Given these limitations, the potential use of autonomous assistive robots as alternative technologies for aiding seniors with ADL has become an emerging area of research [8]–[16]. Assistive robots are promising technologies as they are capable of autonomously providing physical [9], social [17], and cognitive interventions [16], as well as adapt their interactions and assistance levels to the needs and preferences of their user.
1.2 Socially Assistive Robots for Eldercare

Robots have successfully been used to support seniors with a wide variety of ADL [15], [17]–[21]. Socially assistive robots, which use social interactions rather than physical interactions to provide assistance [22], have been shown to provide particularly positive contributions to seniors’ overall wellbeing by improving engagement and performance in ADL [21], [23]–[26]. They can provide support for physical activities, for example motivating seniors to do physical exercise [9] or stroke patients in rehabilitation exercises [27], cognitive support, for example encouraging cognitively impairment seniors with meal-eating [28] and assisting with making a cup of tea [16], and social support, for example by facilitating Bingo games [25] and encouraging social interactions between residents [17] in long-term care facilities.

While seniors and caregivers alike have shown positive responses to these robots, not all robots are created equal. Studies show that the quality of the interaction can be significantly affected by a robot’s physical appearance and behaviors, as appearance shapes a user’s expectations for the robot’s capabilities [29]. It is therefore important to understand what physical robot features are conducive to increasing engagement and compliance with assistive robots, and what behaviors are expected of them.

1.2.1 Social Features to promote Human-Robot Interactions

Current assistive robots for seniors have varying morphologies, even when assisting with similar tasks. These robots vary in size, level of anthropomorphism, mobility, and forms of communication. For example, platforms used for telepresence between seniors, family members and healthcare professionals range from tablets placed on mobile platforms such as Giraff [30] and VGo [31], to human-like robots such as Kompai [32], Tangy [33], and the NAO [34] robots. Other examples of robots designed to assist seniors include the child-like robot Bandit, equipped with a physically emotive face, which has been used to assist seniors in playing music games [19], and the Care-O-Bot, a technomorphic robot capable of fetching and carrying objects, and triggering emergency interventions [35].

While previous research has shown that a robot’s appearance can directly affect interaction quality and acceptance of the robot [36], it remains unclear which dynamic social features, such
as facial expressions and gestures, are needed for such robots to provide effective and engaging assistance to seniors. This is due in large part to the varying designs of existing robots which provide similar assistance.

Focus groups and video-based comparison studies have been used to assess preferences in embodiment and robot behaviors [37]–[39]. However, participants in such studies did not interact directly with the robots, making it difficult to assess how robot features affect a user’s behavioral responses and the overall human-robot interaction (HRI). A handful of interaction studies have compared the effects of embodiment on interaction by asking participants to interact with a physical robot, a remotely located robot, or a virtual robot on a screen [40]–[44]. These studies have found that a physically collocated robot significantly increases engagement, perceived enjoyment, and acceptance of the technology as compared to a virtual or remotely-located robot. Other studies have investigated the effects of morphology by comparing interactions with a physical robot having different faces [45], and a robot compared to a tablet [46]. These studies have found that using a robot for healthcare applications provides a more engaging and generally preferred interaction. Though these studies validate the need for a physically embodied robot, they provided little information about how a robot’s dynamic social features affect the interaction.

In [47], it was noted that the more proximate the relationship between a robot and user, the more social a robot should be. Given that such robots are expected to maintain long-term, proximate relationships with seniors, it is important to understand how a robot’s features and corresponding behavior enable it to promote engagement, establish trust, incentivize compliance, and lead to successful activity completion. Developing appropriate social behaviors for assistive robots using its dynamic social features remains an ongoing area of research.

1.2.2 Development of Socially Assistive Robot Behaviors

For robots to effectively function in human-centric environments, they must be equipped with the physical and social behaviors necessary for interacting with humans. This is especially true of socially assistive robots that use social interactions to provide assistance to vulnerable populations, such as those with cognitive impairment, as the quality of a robot’s behaviors,
particularly its speech and gestures, directly influences the effectiveness of the assistance provided [22].

Behaviors for socially assistive robots have traditionally been designed using one of three methods: 1) manually hand-crafting combinations of speech, gestures, and other communication modes necessary to display a behavior [48]–[51]; 2) teaching a robot multi-modal behaviors through learning from demonstration (LfD) [52], [53]; or 3) autonomously learning multi-modal behaviors via reinforcement learning (RL) [54], [55]. Manually designing robot behaviors involves tedious annotation, with limited potential for expanding the robot’s skillset once deployed in an environment, i.e. it is unable to improve its behaviors over numerous interactions with its user. LfD and RL allow robots to learn behaviors without having to pre-program them, however, they may require large numbers of interactions with demonstrators and their intended users for training, which may not always be available. Indeed, it is not safe for users with cognitive impairment to receive assistance from a technology still being trained. In addition to learning general behaviors, socially assistive robots may also have to adapt their behaviors to their specific users, as behavior personalization can positively affect its acceptance [27], [48], and increase the use of the robot over time [48].

Only a handful of work has focused on personalizing robot assistive behaviors to user profiles [27], [56]–[58]. This research has mainly focused on personalizing behaviors to either a specific user group, for example extroverted vs introverted users [27], or measuring one specific user state during an activity regardless of the individual user, such as stress level during a memory game [56]. Previous work has not investigated the personalization of robot behaviors to a single user’s cognitive model, which is important for a robot to effectively assist users with cognitively impairment. This thesis aims to develop an effective method for teaching a robot personalized assistive behaviors for providing user-centric care. A novel robot learning architecture for learning personalized socially assistive behaviors to assist with activities of daily living is presented.
1.3 Thesis Objectives

The objective of this thesis is to develop an architecture for assistive robots to learn and personalize socially assistive behaviors for HRI applications involving seniors with cognitive impairments. To do so, the effectiveness of various robot social features is validated through direct user studies with seniors having cognitive impairment. Two robots with different morphologies and a tablet are used in the investigation, and the preferred morphology is identified. A functional relationship between the preferred social features and social behaviors is established. Finally, an architecture for learning socially assistive robot behaviors using the identified social features is developed. This architecture uses expert demonstrations in LfD to teach a robot how to display the desired behaviors using appropriate combinations of social features, and when to display the behaviors based on activity and user states. A Q-learning RL algorithm with an Upper Confidence Bound exploration strategy is then used to personalize robot behaviors to a user’s cognitive model. The main contribution of this work is a unique architecture combining LfD and RL to learn social feature combinations, state-behavior pairings, and behavior personalizing for HRI applications.

1.4 Proposed Methodology

A methodology for identifying key social communication features and developing the behavior-learning architecture is presented in Chapters 2 to 4.

1.4.1 Literature Review

A literature review on socially assistive robot features and behaviors is conducted. Preferences and impacts of social robot features and morphologies are investigated, along with methodologies for designing assistive robot behaviors.

1.4.2 Investigating the Influence of Communication Modes in Senior-Robot Interaction

A user study evaluating seniors with cognitive impairments’ preferred robot social features is presented. The seniors prepare a cup of tea with assistance from two different robots and a tablet. The results of the study are discussed in conjunction with results from the literature.
1.4.3 An Architecture for Learning Socially Assistive Robot Behaviors

The procedure for developing a unique architecture for learning and personalizing robot behaviors is introduced. First, the behaviors are learned through expert demonstrations. Next, these behaviors are labeled according to their communication mode types. Finally, Q-learning is used to learn personalized labeled behaviors based on the user’s cognitive model.

1.4.4 Behavior Learning Experiments

Experiments are presented involving expert demonstrators teaching a robot appropriate behaviors and state-behavior pairs. The demonstrations involve assisting a senior with cognitive impairments in making a cup of tea. A Q-learning algorithm is trained in simulation to teach the robot a policy for personalizing the learned behaviors to a user’s cognition.

1.4.5 Conclusion

Lastly, the main contributions of this work with regards to identifying key robot social features and developing a novel architecture are presented, as well as future areas of research.
Chapter 2
Literature Review

This chapter provides a review on the literature investigating senior’s preferences for assistive robot features (section 2.1) and methods for learning desired robot behaviors (section 2.2).

2.1 Socially Assistive Robots

Socially assistive robots use social, rather than physical, interactions to provide care and assistance to vulnerable populations [22]. These robots come in a variety of forms and morphologies depending on the assistance provided, ranging from a seal-like robot Paro [17] to the human-like robot Brian 2.1 [15]. The following robot morphology classifications are defined:

1) Animal-like: a robot having animal-like features and embodiment, such as four legs, a tail, fur, etc.;
2) Machine-like: a robot having little to no physical human-like features;
3) Character-like: a robot having cartoon- or creature-like features which do not resemble a human or animal;
4) Human-like: a robot having physical human-like features such as an upper body, arms, head, while navigating an environment using either a mobile base or biped motion; and
5) Android: a robot that looks and acts like a human (i.e. having artificial skin, hair).

Senior preferences for the different morphologies are discussed in the following sections.

2.1.1 Preferences for Socially Assistive Robots

Senior’s preferences for socially assistive robot features are discussed based on results from focus groups, single robot user studies, and multi-robot user studies.

2.1.1.1 Results from Focus Groups

Focus groups have shown that preferences for robot design and functionality depend on user needs [37]–[39], [59]. In [59], a live demo of the human-like RobuLAB robot and videos of ten other socially assistive robots were presented to a group of 25 participants composed of healthy seniors, seniors living with mild cognitive impairment and family caregivers of persons with dementia. Healthy seniors preferred machine-like robots, caregivers preferred mechanical-like robots, and seniors living with mild cognitive impairment preferred animal-like robots. Similarly in [39], a group of 15 seniors who were shown 26 photos and video clips of robots with varying morphologies (e.g. Aibo, Asimo, Care-o-bot3, Kobian, Mamoru, Nao, Paro) indicated a
preference for robots that are small and have some character- or animal-like traits, while human-like and android robots were the least preferred. Their preferences were mainly due to ethical concerns rather than esthetics, as participants were concerned that human-like and android robots would substitute real human interaction.

These results provide convincing evidence for designing a robot’s form and functionality to meet user needs. However, the results should be validated by observing seniors directly interacting with assistive robots.

2.1.1.2 Results from Single-Robot User Studies

User studies involving seniors directly interacting with an assistive robot have demonstrated that these robots provide positive contributions to a senior’s overall wellbeing by improving engagement and performance in ADL [17], [20], [21], [23]–[26], [60]. One of the main reasons for these positive effects comes from a robot’s social attributes, suggested through its physical form and perceived social intelligence. For example, residents in a care facility interacting with the human-like Brian 2.1 robot during both a meal-eating activity and a memory game liked the emotional responses the robot displayed via facial expressions and vocal intonation, as well as its life-like appearance and the companionship provided [15], [20]. Their interactions with the Brian 2.1 robot contributed to high levels of social engagement and compliance during the respective ADL. A HRI study with the Tangy robot facilitating Bingo games with seniors in a long-term care facility showed that participants specifically appreciated the robot’s gestures and social characteristics, which included telling jokes, doing a celebration dance when a player won Bingo, and playing music [25]. An ethnographic study in a Japanese residence showed that the human-like Robovie2 platform successfully integrated into everyday lives of residents thanks to its social characteristics, including greeting the residents by name and casually conversing [21].

These studies were conducted with robots each having different features and corresponding levels of sociability based on their physical forms. All of the robots were met with enthusiasm and high levels of engagement during their respective interactions. However, it is not clear how these robots compare to each other with respect to effectiveness, quality of interaction, and acceptance by seniors. This can only be evaluated through multi-robot user studies.
2.1.1.3 Results from Multi-Robot User Studies

A handful of HRI studies with multiple robots having different morphologies have been conducted to investigate how a robot’s physical design affects user behavior during HRI [45], [46], [61]. For example, the physical embodiment of a mobile tele-presence robot, resembling a tablet supported at human-height level on a mobile base, has been shown to promote greater trust in its users than a hand-held tablet [61]. Another HRI comparison study required a group of adults recruited from the general population to interact with both a character-like iRobiQ robot and a tablet during three activities: a healthcare interview, basic physical exercises, and a relaxation exercise [46]. Compared to the tablet, the robot received higher ratings for trust, enjoyment, and desire for future interaction as measured from questionnaire responses. The robot also promoted more dialogue and increased positive affect in participants during the interaction. In [45], seniors received medicine from three robots in a waiting room consisting of the same PeopleBot platform with different heads, i.e. a tablet, a machine-like face having two eyes and a mouth, or a human-like face having eyes, a nose, and a mouth. The human-like face and voice led to higher levels of valence in participants, as measured by questionnaire results, though no differences were found in the levels of arousal between the machine-like and human-like robots. The authors attribute this result to a lack of dynamic features in the robots: the robot faces were static during the entire interaction.

While these studies seem to suggest than more evolved robots tend to be preferred over static tablets, they have not specifically examined the impact of a robot’s dynamic behavior on the interaction. Namely, these studies have compared robots designed with static features, such as a static face [45] or a static body [46], without considering the influence of dynamic social communication modes such as facial expressions, gestures, body language, and vocal intonation on the quality of interaction with seniors. It is important to understand how these dynamic social communication modes are perceived, as they form the basis for developing robot behaviors.

2.2 Socially Assistive Robot Behaviors

For robots to effectively function in human-centric environments, they must be equipped with the physical and social behaviors necessary for interacting with humans. These behaviors are
displayed through the aforementioned social communication modes, i.e. speech, gestures if available, facial expressions, and mobility.

Manually defining the combinations of social communication features to display behaviors involves sorting through a large dataset (e.g. videos) of human-human interaction, labeling the desired interaction features [51], [62], and manually recreating generalized, representative features on the robot in question [50], [51]. This process is tedious and time-consuming, but most importantly it does not allow a robot to learn or adapt new behaviors over time or across different environments. Therefore, learning-based approaches can be considered. This section focuses on the use of learning-from-demonstration and reinforcement learning algorithms for teaching robots new behaviors.

2.2.1 Teaching Robot Behaviors through Demonstration

While the majority of work on learning from demonstration (LfD) has focused on teaching robots physical tasks [63]–[66], a handful of researchers have considered teaching robots social interactions from demonstration [52], [53]. For example, in [52], videos of 16 demonstrators narrating paper-making to another person were recorded and manually coded for four classes of gestures (deictic, iconic, metaphoric, and beat) and four classes of gaze directions (reference, recipient, narrator’s own gesture, and other). Speech was also coded into lexical affiliates for each gesture class. The coded demonstrations were used as input to a dynamic Bayesian network (DBN) that learned the most probable gesture and gaze direction given a segment of speech. The authors noted large variances in the way demonstrators displayed each behavior, making it difficult to identify a single appropriate gesture and gaze direction with high probability. Therefore, to replicate the behaviors onto a human-like Wakamaru robot, the researchers identified the most common gestures made and created a similar gesture for the robot by manually moving its arms and tracking the gesture trajectory. Given a segment of speech, the robot learned which gesture and gaze direction to display using the learned probabilities from the DBN. In [53], a Robovie II humanoid robot was taught how to interact with customers in a camera store using data from 178 interactions between a seller and customer. Both the seller and customer’s speech, motion, and spatial formations were autonomously clustered into joint behavior states using dynamic hierarchical clustering. The clusters were used in a variable-length Markov model to predict possible behaviors for the robot.
In these two studies, the robots learned a single way to display a behavior, selected based on it having the highest probability against all other demonstrations. However, there may be more than one equivalent way of displaying the same behavior, as both studies mention variance in the way behaviors were displayed. The demonstrations may also have been sub-optimal, as there may be a more optimal way of displaying the behavior which was not provided in the demonstrations. Alternatively, reinforcement learning can allow a robot to learn to display optimal behaviors through interactions with users. RL can additionally be used to learn a user-specific policy, which has been shown to play an important role in the effectiveness of an activity in HRI [19], [56], [67].

2.2.2 Teaching Robot Behaviors through Reinforcement Learning

RL algorithms have been used by robots to learn behaviors from direct interactions with humans. For example, in [55] the Furhat robot, a table-top robotic head, used Q-learning to determine the optimal combination of communication modes (speech, head gestures, gaze direction, and facial expressions) required to direct a person’s attention in a memory game. The participant’s gaze and speech were used as state inputs along with the game state. Combinations of communication modes were selected according to an $\epsilon$-greedy exploration strategy. Costs were assigned to each of the communication modes, while a positive reward was assigned if the user redirected their attention to the game. The robot’s goal was to minimize the overall cost while increasing the number of positive rewards. In [54], the Pepper robot used Q-learning to learn how to gain a person’s attention in a public space using combinations of speech, gestures, and gaze direction. The robot was placed in a public environment for 14 days where it interacted with passersby. Then a deep Q-network was trained offline by sampling from the interaction data. The robot successfully learned which combination of communication modes had the maximum likelihood of getting a person’s attention after approximately 14,000 interactions. While these Q-learning approaches have been used to learn general behaviors for all users, RL techniques have also been used to learn personalized robot behaviors.

2.2.2.1 Robot Behavior Personalization using RL

In [58], Q-learning was used by the human-like ARIO robot to identify the optimal combination of its gestures (e.g., head shake, arm wave), speech (e.g., call the person’s name, make a sound), and navigation (e.g., move to user, move in user’s field of view) to obtain a user’s attention.
while they read. A Hidden Markov Model was trained to identify the user’s state as input based on face direction, body direction, and speech. After 26 interactions, the robot was able to develop user-specific policies based on the way users preferred to be interrupted while reading, for instance one user preferred having his name called out continuously whereas another user preferred having his name called out once followed by a wave. In [56], a MAXQ hierarchical reinforcement learning (HRL) approach was used by the human-like Brian 2.0 robot to learn assistive behaviors for a cognitive training game. Learning occurred in two stages: 1) first, the robot learned appropriate assistive behaviors for the different game states through offline learning, and 2) it then used online learning during interaction with a user to learn to personalize its behaviors to the user’s stress levels by measuring heart rate. These studies adapted the robot’s behavior according to specific activity parameters, which are difficult to generalize to other types of activities.

Other research has focused on adapting the robot’s behavior to general user types. For example, policy gradient reinforcement learning was used to adapt the child-like Bandit robot’s behaviors to extroverted or introverted user personalities during a stroke rehabilitation activity [27]. The robot’s speech, speed, proximity to the user, and voice gender were manually designed to represent extroverted and introverted personalities. During interactions with users, the robot learned to select extroverted or introverted versions of each parameter to yield the highest compliance rates while developing a model for personality-matching.

Rather than personalize behaviors to general user types, focus groups conducted with seniors suggest that socially assistive robots capable of adapting their behaviors to a user’s level of impairment could lead to increased technology acceptance, higher intentions for use, and a more overall positive attitude toward these robots [48]. As previously mentioned, our research focuses on the personalization of robot assistive behaviors to a user based on their level of cognition to provide person-centered assistance. The use of RL for both behavior learning and behavior personalization would require a very large number of interactions with users to develop an optimal policy: in [54] 14,000 interactions were required to learn a policy, and in [27] and [55] the policies never fully converged due to a limited number of interactions. As socially assistive robots interact with vulnerable populations, it is not feasible to train a robot using such a high number of iterations through direct user interactions. To minimize the number of iterations required for behavior learning, LfD can be used to quickly learn an initial policy for selecting
assistive robot behaviors to display based on the user’s state, while RL can be used to optimize the learned policy.

### 2.2.3 Learning Behaviors Using Both LfD and RL

The use of both LfD and RL techniques have either focused on learning agent actions during simulated games such as Atari [68], robot navigation [69], or robot behaviors to facilitate group activities [57]. In the first two cases, the objective was to reduce computation time in traditional RL algorithms by using sub-optimal human demonstrations of the activity to correctly shape the policy, and then finding the optimal policy using RL. For example, in [68], human demonstrations were used with deep Q-learning to teach a computer to play several Atari games. Initially, the algorithm randomly samples from a set of human demonstration data which include the state, action taken, reward received, and next state at every time step. The demonstration samples are placed in a batch, which is used to iteratively update the neural network and shape the weights. The game is played by the RL using the current policy, and the state, actions, rewards, and next states found using RL slowly replace the demonstration samples in the batch. This algorithm was tested on a series of Atari games, and outperformed other Deep Q-Networks and the highest performing human demonstrations. In [69], an algorithm called Approximate Policy Iteration with Demonstrations (APID) was developed and tested in both a car driving simulation and a real robot navigation task. APID is initially given a sample of human demonstrations relevant to the activity being learned, which include the state, action taken, reward received, and next state at every time step, and may be sub-optimal or limited in quantity. The demonstrations are initially used to shape the value function in Approximate Policy Iteration by imposing a set of linear constraints during policy evaluation, after which a policy improvement step is applied. The algorithm outperformed both Least-Squares Policy Iteration and supervised learning in both experiments, irrespective of the quantity of demonstrations provided.

With respect to HRI applications, in [57] LfD was used to teach a robot assistive behaviors for a stimulating group activity and RL was then used to personalize the robot’s speech content using persuasion strategies and Thompson Sampling. Non-expert demonstrators used a GUI to define the sequence of assistive behaviors using known behavior types. Only three activity demonstrations were needed for the robot to learn the appropriate behavior sequence given the
user activity state and user assistance request state. The robot learned which of its four persuasion strategies (i.e. praise, suggestion, scarcity, and neutral) was most likely to achieve compliance from users playing a Bingo game.

To the best of the authors’ knowledge, no robot behavior-learning architecture has been developed to learn communication mode combinations necessary to display an assistive behavior, behavior sequences, and user personalization. To address this challenge, this thesis proposes to leverage the strengths of both LfD and RL. Namely, LfD will be used to teach the robot how and when to display assistive behaviors provided by expert demonstrations. The expert demonstrations teach the robot combinations of speech and gestures required to display a behavior, and map the behaviors to robot and user states. As previous research using LfD has shown, there can exist a large variance in how behaviors are displayed by different demonstrators [52]. Rather than use general behaviors, the variance in speech and gestures in the demonstrated behaviors is used to personalize the robot’s behaviors to an individual user’s cognition. Behavior personalization is done using an RL algorithm that learns which demonstrated behavior, labeled according to its speech and gesture types, is most likely to transition the user into a desirable state (i.e. focused and completing the correct step).
Chapter 3
Investigating the Influence of Communication Modes in Senior-Robot Interaction

User studies were performed with seniors having cognitive impairment to identify their preferences for social robot communication modes. The seniors’ preferences were evaluated after preparing a cup of tea with assistance from a human-like robot, a non-human-like robot, and a tablet.

3.1 User Study Goals

This HRI comparison study investigated how the use of dynamic robot features and social communication modes, i.e. facial expressions and gestures, affect the interaction between seniors and robots. Three platforms (two robots and a tablet) of increasing levels of sociability, afforded by the dynamics of their physical design, were considered: a human-like robot with facial expressions, arm gestures, and head motions; a character-like robot with a virtual face expressing either a smile or talking mouth motion; and a tablet with no facial constructs. All platforms were capable of displaying the same media content on a screen. The study examined how the varying degrees of sociability afforded by the three platforms’ respective social features influenced the participants’ engagement, affect, perceived social intelligence of the robot, collaborative behavior, trust, and compliance during an assisted ADL.

3.1.1 Study Hypotheses

Four hypotheses were investigated based on the three different platforms tested:

1. A human-like robot with dynamic social features will increase user engagement in an assistive activity when compared to a character-like robot with an animated virtual face; which will in turn increase engagement compared to a static tablet.
2. A human-like robot with dynamic social features will be perceived as more socially intelligent when compared to a character-like robot with an animated virtual face, which in turn will be perceived as more socially intelligent than a static tablet.
3. A human-like robot with dynamic social features will be perceived as more useful than a character-like robot with an animated virtual face, which will in turn be perceived as more useful than a static tablet.

4. A human-like robot with dynamic social features will be preferred overall as an assistive technology over a character-like robot with an animated virtual face, which will in turn be preferred over a static tablet.

3.2 User Study Design

The comparative study design for evaluating seniors’ perceptions of human-like robots, character-like robots, and tablets is presented in the following sections. We obtained ethics approval prior to the commencement of the study.

3.2.1 Participants

Participants were recruited from a local retirement home in Toronto. An initial information session was conducted to explain the study to the residents. The inclusion criteria for the participants were to: 1) be 60 years of age or older; 2) have mild cognitive impairment to healthy functioning, with a MOCA cut-off score of 19 out of 30 [70]; 3) have difficulty completing tasks in the kitchen as described by their caregivers; 4) have no other health problems that would otherwise affect their ability to perform the tea-making task; 5) be fluent in English; and 6) have normal levels of hearing. In total, six participants, all female, participated in the study ranging in ages from 82 to 96 ($\mu$=89.5, $\sigma$=5.32). The mean MOCA score for all participants was 25.8 with a standard deviation $\sigma$=2.48.

Participants’ experience with computers ranged from beginner to intermediate, whereas they had no experience or beginner experience with robots, Table 1. Computer experience was classified as either: 1) No experience, 2) Beginner (e.g. email, simple programs), 3) Intermediate (e.g. internet browsing, chat), and 4) Experienced (editing documents, using complex programs). Experience with robots was classified as either: 1) No experience, 2) Beginner (have seen robots at museums, science centers, or on TV), 3) Intermediate (have worked with/used industrial robots), and 4) Advanced (have worked on developing either robot hardware or software).
The participants were also asked to select their preferred robot(s) from a set of images with the following options: a) Paro [71]; b) iCat [72]; c) Pomi [73]; d) Mamoru-Kun [74] e) Eve [75]; f) RobuLAB [76]; g) Kampai [32]; h) Pearl [60]; i) NAO [34]; j) Nexi [77]; k) Kobian [78]; l) Telenoid [79]. This information was gathered to determine preferences and expectations of robots prior to the participants interacting with any platforms. Overall, participants favored human-like robots, though two participants also liked the character-like robot Pomi and one participant also chose the RobuLAB robot. Four participants selected the NAO humanoid robot.

Table 1 Pre-study questionnaire results showing experience with computers and robots, and initial robot preferences

<table>
<thead>
<tr>
<th>Participant</th>
<th>Experience with computers</th>
<th>Experience with robots</th>
<th>Preferred robot(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participant 1</td>
<td>Beginner</td>
<td>No experience</td>
<td>NAO</td>
</tr>
<tr>
<td>Participant 2</td>
<td>Intermediate</td>
<td>Beginner</td>
<td>RobuLAB, Pearl</td>
</tr>
<tr>
<td>Participant 3</td>
<td>No experience</td>
<td>No experience</td>
<td>Pomi, NAO, Kobian</td>
</tr>
<tr>
<td>Participant 4</td>
<td>No experience</td>
<td>No experience</td>
<td>NAO, Pomi</td>
</tr>
<tr>
<td>Participant 5</td>
<td>No experience</td>
<td>No experience</td>
<td>NAO</td>
</tr>
<tr>
<td>Participant 6</td>
<td>Intermediate</td>
<td>Beginner</td>
<td>Kampai</td>
</tr>
</tbody>
</table>

3.2.2 The Interactive Platforms

Three interactive platforms were used to assist the senior participants in preparing a cup of tea: the human-like Casper robot, the character-like Ed robot, and a tablet display.

The Casper Robot

The Casper robot [80], [81], Figure 1, is a human-like robot developed in the ASBLab with an expressive face, two arms, and a torso mounted on an omnidirectional base. It can display five facial expressions using LEDs for its eyebrows and mouth: happy, sad, surprised, angry, and neutral. The robot uses the Amazon Polly Text-to-Speech API [82] and a speaker on its base for communication through speech. The robot’s neck has 2-degrees of freedom (DOF) which allow for nodding and shaking its head. Casper’s 3-DOF arms are used to display gestures. An ASUS RGB-D sensor on the robot’s torso is used for person detection and tracking, while a Hokuyo laser rangefinder is used for environment mapping, robot localization and navigation. Casper also has a 10” touch screen tablet mounted on its chest for displaying multimedia such as videos, images, and text.
The Ed Robot

The Ed robot [83], Figure 2, developed in the Intelligent Assistive Technology and Systems (IATSL) Lab consists of a long torso mounted on an iRobot Create platform. On top of the torso is an LCD screen that displays both the robot’s virtual face, which either smiles or performs speaking motions during audio prompts, and the tea-making prompting videos. Ed uses Cepstral’s US English voice, William [84], as its text-to-speech service. Ed has two speakers on the front of its body, and two mounted RGB cameras: one on its head and one on its base. The top camera uses a fisheye lens to provide a larger field of view of a user and environment, and the base camera is used to tele-operate the robot through an environment.

The Tablet

A 10” GeChic 1002 touch-screen display connected to a mini-PC was placed on a supporting stand on the
counter, Figure 3. A connected speaker played audio prompts, while the display showed videos and images.

3.2.3 The Tea-making Activity

Tea-making was chosen as the assistive ADL as cognitively impaired seniors have mentioned requiring assistance from a caregiver to accomplish this activity [85], this activity has been used in previous studies with the Ed robot and seniors with dementia [16].

All interactive platforms were equipped with the same videos and audio prompts, obtained from the COACH system [86]. The COACH prompting system is a software designed to assist individuals with cognitive impairments in completing activities of daily living using audio and visual cues. The system decomposes an ADL, such as tea-making, into individual step-by-step instructions each linked to a video prompt. The video prompt also contains audio explaining each step (e.g. “Try pouring some hot water into the tea cup”). When a person finishes a given step, the next video prompt is played, until the activity is done. For the tea-making activity, a total of six videos were played, beginning with a prompt to turn the faucet on, and ending with a prompt to pour hot water into the cup, as shown in Figure 4.

Figure 4 Excerpts from the six COACH videos showing each step in the tea-making task: (a) Turn the water on; (b) Pour water into the kettle; (c) Turn the water off; (d) Turn the kettle on; (e) Put the tea bag in the cup; and (f) Pour hot water into the tea cup.
3.2.4 Procedure

The tea-making activity took place in a public kitchen at the retirement home. The study setup with each platform is shown in Figure 5. A sink, electric kettle, mugs and tea were provided. The platforms were positioned opposite the sink. Two cameras were placed in the environment to record the interaction sessions. One camera was placed behind the interactive platform and another camera in front of the counter on which the tea-making items were placed.

Participants individually interacted with each of the platforms within the kitchen environment. They were each asked to make a cup of tea with assistance from one of the three platforms. Each participant interacted with a specific platform twice within a 10-day interval, for a total of 6 sessions. Sessions were spread out by at least one day.

Sessions with the Ed robot, Casper robot, and tablet followed a randomly assigned sequence. During the first session with each platform, a member of our research team brought each participant to the kitchen, where the respective platform’s functions and dynamic features were explained, and any questions were answered.

The researcher then left the kitchen area and began tele-operating the platforms using a Wizard-of-Oz approach. The robots/tablet provided a brief introduction about their own abilities, demonstrated their features, and initiated the tea-making activity using the COACH system. Audiovisual prompts were provided for each step in the tea-making task, or when a participant encountered difficulty in completing the step. While the water in the kettle was put for boiling, the interactive platforms asked each participant three social questions: (1) “How are you doing
today?; (2) “How is the weather outside?”; and (3) “What do you have planned this afternoon?”. Once the water had boiled, prompts from the COACH system were resumed until completion of the activity, at which point the robot/tablet thanked the participant for making tea with them and ended the activity.

3.2.5 Evaluation Metrics

Data from videos, and interviews were analyzed to evaluate the quality of interaction between the seniors and each of the platforms, as well as their feature preferences.

3.2.5.1 Video data

Videos of each session were recorded to identify participant behavioral cues pertaining to the following variables: 1) engagement, 2) trust, 3) affect, 4) perceived social intelligence, 5) collaborative behavior, and 6) compliance. These independent variables have been previously used in human-robot interaction studies to evaluate the quality of interactions with robots [15], [16], [25], [47], [87]. Table 2 summarizes the user behaviors associated with each measured variable.

Table 2 Measured variables obtained from video data

<table>
<thead>
<tr>
<th>Measured variable</th>
<th>Corresponding Behaviors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Engagement</td>
<td>Visual focus of attention (gaze direction and length of time)</td>
</tr>
<tr>
<td>Trust</td>
<td>Asks or turns toward the platform for help/explanations</td>
</tr>
<tr>
<td>Positive affect</td>
<td>Smiles, laughs</td>
</tr>
<tr>
<td>Negative affect</td>
<td>Grimaces, frowns, shows restlessness</td>
</tr>
<tr>
<td>Perceived social intelligence</td>
<td>Social cues toward the platform (asking the platform: “Do you like tea?”; thanking it, using its name)</td>
</tr>
<tr>
<td></td>
<td>Backchannels (e.g. head nods, linguistic cues such as “uh-huh”, “okay”)</td>
</tr>
<tr>
<td>Collaborative behavior</td>
<td>Status updates (e.g. “The water is boiling”, or “I put the teabag in the cup”)</td>
</tr>
<tr>
<td></td>
<td>Questions toward the platform (e.g. asking: “What’s next?” or “What do I do now?”)</td>
</tr>
<tr>
<td>Compliance</td>
<td>Ratio of prompts complied with over total number of prompts given</td>
</tr>
</tbody>
</table>
**Engagement**

High levels of engagement have been shown to increase positive emotions, thus improving general quality of life and willingness to perform ADL [88]. We measure engagement towards a platform during the assistive activity using visual focus of attention [28], [89], defined herein as the percentage of time a participant’s eye gaze was towards the platform during the social interaction portions of the activity.

**Trust**

Trust is a fundamental factor in the success of human-robot collaborations, as it directly affects a person’s willingness to share information, accept the platform’s suggestions, and rely on it in moments of uncertainty [90], [91]. Users who trust robotic technology are more inclined to ask it for assistance and accept its suggestions [90]. Trust was measured as the number of instances when a participant actively asked the platform for assistance and followed its suggestion.

**Affect**

Ensuring positive affect is important to promoting collaborative behavior in HRI [92], while contributing to increased trust between two agents [91]. An individual’s affect level is displayed via his/her verbal and non-verbal cues [93]. During the tea-making activity, the number of instances when a participant smiled or laughed were recorded as signs of positive affect, whereas instances where a participant frowned, grimaced, or showed signs of restlessness through excessive movements or hand-wringing were identified as negative affect.

**Perceived Social Intelligence**

Perceived social intelligence is a measure of how much human-like intelligence a technology is believed to have as displayed through multi-modal forms of communication [94]. An assistive technology lacking skills relevant to the social context, such as appropriate speech and spatial distance from the user, may go unused and fail to be perceived as a useful assistant [47]. Perceived social intelligence was determined by the display of four behaviors: (1) the number of times a participant asks the platform a social question, e.g. asking the robot/tablet “Do you like tea?”; (2) the number of times a participant thanked the platform or responded to it politely; (3) the number of times a participant used the robot’s name to greet it or ask it a question; and (4)
the number of times a participant used backchannels during dialogue, defined by head nods and shakes, and utterances of “yes”, “uh-huh”, and “okay”. Backchannels are brief verbal or gestural cues that act as social reinforcers indicating agreement and understanding in dialog [89], [95].

**Collaborative Behavior**

Research in HRI has shown that collaboration with a technology indicates willingness to consider it as an interactive partner and valuable team member [90]. During tea-making, proactive collaboration was defined as the provision of status updates and activity-based questions [16]. Status updates were defined as instances when a participant provided the platform with information about the activity status or an action he/she was completing, for example: “The water is boiling” or “I put the teabag in the cup”. Activity-based questions were defined as instances where a participant asked the robot/tablet about next steps or for permission to act, for example: “What’s next?” or “Can I put the teabag in?”

**Compliance**

Compliance with an assistive technology is important to ensure a senior follows the correct steps, recovers from mistakes, and successfully completes the activity [87]. Compliance in this study was measured as a ratio of the number of prompts followed to the total prompts provided [28].

### 3.2.5.2 Semi-structured interviews

After the second interaction session with a platform, a semi-structured interview was conducted with the participants in order to understand their preferences, behaviors, and perceptions of the three platforms [96]. The specific questions asked in the interviews are presented in Table 2. They were chosen to better understand preferences for the platforms’ features and to have participants reflect on their experience and interactions with the platforms. Special emphasis was placed on topics related to a robot’s dynamic social features (gestures, facial expressions), voice, behavior (prompts, conversation, display) and practical use (in a home setting). Participants were encouraged to speak freely, and elaborate on answers given.
Table 3 Semi-structured interview questions for each platform

<table>
<thead>
<tr>
<th>#</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>What are your overall impressions of the robot/tablet?</td>
</tr>
<tr>
<td>2</td>
<td>What did you think about using the robot/tablet to make tea?</td>
</tr>
<tr>
<td>3</td>
<td>What did you think of the way prompts were given?</td>
</tr>
<tr>
<td>4</td>
<td>Would you like to have the robot/tablet in your home?</td>
</tr>
<tr>
<td></td>
<td>What would you use it for?</td>
</tr>
<tr>
<td>5</td>
<td>If I left the robot/tablet here for everyone to use, who would use it?</td>
</tr>
<tr>
<td></td>
<td>Would you use it?</td>
</tr>
<tr>
<td></td>
<td>How?</td>
</tr>
<tr>
<td>6</td>
<td>What did you think of the robot/tablet’s appearance?</td>
</tr>
<tr>
<td></td>
<td>Did anything strike you?</td>
</tr>
<tr>
<td>7</td>
<td>What did you think of the robot/tablet’s voice?</td>
</tr>
<tr>
<td>8</td>
<td>What did you think of the robot/tablet making conversation while water</td>
</tr>
<tr>
<td></td>
<td>boiled?</td>
</tr>
<tr>
<td>9</td>
<td>How did you find the robot’s face? (N/A for the tablet)</td>
</tr>
<tr>
<td>10</td>
<td>What do you think of the robot/tablet’s size?</td>
</tr>
<tr>
<td>11</td>
<td>Is there anything you would like to improve on the robot/tablet?</td>
</tr>
<tr>
<td>12</td>
<td>Did the robot/tablet meet your expectations?</td>
</tr>
<tr>
<td>13</td>
<td>Which platform did you prefer the most? The least? Why?</td>
</tr>
<tr>
<td></td>
<td>(asked after all sessions completed)</td>
</tr>
</tbody>
</table>

3.3 Results and Discussion

A non-parametric Friedman Two-way test [97] was used to evaluate the participant data for each of the measured variables across all three platforms. The data for the first and second sessions were averaged for each participant, to eliminate biases for novelty effects, daily variations in cognitive state, and random environmental variations such as visitors or staff entering the room during a session. The detailed results for each participant for each platform with respect to the measured variables are presented in Appendix A. A summary of the Friedman test results is provided in Table 4, calculated with $df = 2$. A post hoc Wilcoxon signed-rank test with a Bonferroni correction was used for pairwise comparisons. A significance level of $\alpha_{\text{revised}} = 0.033$ was computed to avoid Type 1 error across multiple comparisons. The Bonferroni correction was calculated given an original $\alpha = 0.10$ divided by three platform tests.
Table 4 Summary of Friedman test results for video data variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>$\chi^2(2)$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Engagement</td>
<td>7.00</td>
<td>0.0302</td>
</tr>
<tr>
<td>Trust</td>
<td>1.58</td>
<td>0.4538</td>
</tr>
<tr>
<td>Positive affect</td>
<td>7.58</td>
<td>0.0226</td>
</tr>
<tr>
<td>Negative affect</td>
<td>1.08</td>
<td>0.5800</td>
</tr>
<tr>
<td>Perceived social intelligence</td>
<td>5.33</td>
<td>0.0696</td>
</tr>
<tr>
<td>Collaborative behavior</td>
<td>3.25</td>
<td>0.2000</td>
</tr>
<tr>
<td>Compliance</td>
<td>1.00</td>
<td>0.6100</td>
</tr>
</tbody>
</table>

The mean number of instances when participants displayed each of the measured variables when interacting with each platform, along with the respective standard deviations, are presented in Appendix A. In summary, the results of the Friedman test showed statistical significance for the variables Engagement, Positive Affect, and Perceived Social Intelligence, as seen in Table 4. A post-hoc Wilcoxon signed-rank test revealed statistically significant differences between the Casper robot and tablet for Engagement, Positive Affect, and Perceived Social Intelligence, and between Casper and Ed for Engagement only. None of the measured variables showed statistically significant differences between Ed and the tablet.

The semi-structured interviews were subjected to a thematic analysis to determine recurring themes. The frequency of similar responses to each question was evaluated. Responses with a recurrence of 50% or more across participants (i.e. when at least 50% of participants mentioned a similar response) for at least one platform were grouped into themes.

The results of these analyses are discussed in the following sections.
3.3.1 Engagement levels during interaction

Hypothesis 1: A human-like robot with dynamic social features will increase user engagement in an assistive activity when compared to a character-like robot with an animated virtual face; which will in turn increase engagement compared to a static tablet.

The first hypothesis was partially validated. Results showed that participants were significantly more engaged with Casper than with the other two platforms. Questionnaire and interview data show that engagement levels improved with increasing levels of dynamic features. For example, when asked to select their preferred features, all participants selected Casper’s face and 67% chose Casper’s facial expressions, whereas 67% of participants selected Ed’s face and only 33% selected Ed’s facial expressions. It was also found that participants directly responded to these facial expressions, for example, they laughed and smiled when the robot changed its facial expressions to surprised or happy. Participants were less engaged with Ed’s digital face, and not at all with the tablet.

When referring to Casper, one participant said: “the expressions ...draw in your attention.” Comparatively, when asked what she thought of Ed, one senior stated she preferred Casper because “it’s [Casper’s] better to talk to. It’s more like someone, like they’re looking at you and talking.” To her, the robot had the same level of presence as a person would. Another senior pointed out that “he’s [Ed’s] got a little smile on his face. But I think work could be done on it. See something a little more human, not just smile or sad.” Engagement levels for the tablet were especially low: one participant never looked at the tablet during the tea-making activity. One senior commented “it’s [the tablet] not as interesting as using a robot that looks like a person... You need to keep the person’s attention.” Lack of engagement in assistive tablets is not without precedent: two thirds of participants never looked at the tablet in a study involving seniors with dementia playing a music game with a human-like robot and a digital representation of the robot on a tablet [41]. Designing robots to have an expressive, human-like face seems to be very important to maintaining engagement and interest during interactions with assistive technologies.

While few participants rated Casper’s gestures and arm movements as a preferred feature, video data shows that participants naturally responded to the robot’s movements during the interaction. In particular, participants often mimicked the robot’s actions. For example, Figure 5(a) shows the participant imitating Casper’s greeting wave motion.
3.3.2 Perceived social intelligence across platforms

Hypothesis 2: A human-like robot with dynamic social features will be perceived as more socially intelligent when compared to a character-like robot with an animated virtual face, which in turn will be perceived as more socially intelligent than a static tablet.

Video data results suggest that dynamic anthropomorphic features increase a technology’s perceived social intelligence. Though only the Casper-tablet comparison yielded statistically significant differences in perceived social intelligence, overall the data show an increasing trend of perceived social intelligence starting from the tablet to Ed, and then to Casper, as suggested by the second hypothesis.

Communication patterns show that participants thought Casper had the highest level of social intelligence. Though participants were equally likely to call both robots by their respective names, they repeatedly thanked Casper and asked it personal questions such as “Would you also like some tea?”. The biggest difference across platforms was found in the number of backchannels provided toward each platform (i.e. head nods, “uh-huh” utterances). Casper received almost twice as many backchannel cues as Ed, and almost ten times more backchannels cues than the tablet. These results may be due to participants viewing the platforms as increasingly more human-like, not just in appearance but also in behavior. Previous research has shown that technologies capable of showing sociable behavior are typically preferred and prompt social responses from individuals [8], [44], [50], [98]–[100].

In terms of social features, one participant stated “that Casper’s eyes “really make the face. Very life-like.” Its facial expressions made one participant feel the robot was similar to a person, i.e. “It’s more like someone, like it’s looking at you and talking. That’s a desirable thing.” One participant also emphasized that the robot’s dynamic movements made it feel more human-like: “It had a bit of movement. More of a person than Ed. I want it to feel more like a person.” Another participant mentioned Casper’s arms have “a friendly manner.” These comments suggest that the participants felt Casper was like a real person, and were likely to socialize with the robot as they would another human.

Participants also liked conversing with the Ed robot – it was an important feature for its acceptance. One participant said “I like talking to him. The conversation aspect is important.
Especially if he has any thoughts of his own!” However, Ed was seen as much more of a robot than a person. Although participants wanted to converse with Ed, because it was not viewed as human-like, they may have been less social towards it during dialog. Conversely, the tablet prompted little to no social responses from the participants. As a static platform with no dynamic features or corresponding behaviors, the lack of perceived social intelligence is expected.

3.3.3 Perceived usefulness across platforms

_Hypothesis 3: A human-like robot with dynamic social features will be perceived as more useful than a character-like robot with an animated virtual face, which will in turn be perceived as more useful than a static tablet._

Our third hypothesis was not directly validated. No statistical differences were found in perceived usefulness among the platforms. However, interviews reveal that Casper was perceived to be useful for social activities and ADL, whereas Ed and the tablet were perceived to be useful for ADL only. Casper was the only platform associated with social activities such as accompanying seniors on walks and playing games. Conversely, the tablet was not associated with any interactive activities.

Though the robots had no arms capable of physical interactions, participants believed they could pick up and fetch objects, highlighting the important of these tasks to this user group. The desire for robots to perform physically interactive tasks is common with elderly populations [26]. A robot with arms should likely also be able to complete physical-contact tasks. In the event that a user assumes a robot is capable of more than it actually is, they risk over-trusting the technology, or becoming disappointed and under-using it [101].

Participant comments also show that the tablet was regarded as a tool, while Casper was regarded as a companion, and Ed an intermediate between the two. For example, one participant stated that she would “use it [tablet] only as a tool, not to make a conversation.” On the other hand, one participant said Casper “would be a companion.” All participants wanted to converse with Casper, while only half of participants wanted to converse with Ed or the tablet.

Lastly, the platform features also influenced how participants thought of using the robots/tablet in their own homes. When asked where they would use and store the platforms at home, five out of six participants were inclined to keep Casper in social or private areas (living room, bedroom,
and bathroom), whereas only one participant suggested keeping Ed in a social area (living room) near the entrance. The tablet did not need a specific storage location as it can be placed in any drawer. Interestingly, although the activity presented to participants in kitchen, none of the suggested assistive activities mentioned by the participants included kitchen-based activities. One participant said that she “really wouldn’t see it [Casper] in the kitchen because it wouldn’t have the same level of communication. By having it in the living room, it implies conversations and friends, or companionship.” On the other hand, the same participant preferred to keep Ed in the kitchen rather than the living room, saying: “I wouldn’t want... a guest to just stumble on it. I want to dress it up a little, and after introduce the robot... I just don’t want them to think ‘Wow she’s gone way down now – so much so that she needs a robot to help her.’”

3.3.4 Technology preferences

_Hypothesis 4: A human-like robot with dynamic social features will be preferred overall as an assistive technology over a character-like robot with an animated virtual face, which will in turn be preferred over a static tablet._

Our fourth hypothesis was partially validated through affect levels measured from video data, supported by participants’ verbal responses during the semi-structured interview. Showing more positive affect is a reflection of greater technology acceptance [48]. Measures of positive affect from video data were statistically higher for participants interacting with the human-like robot Casper compared to the tablet. Though no statistical significance was found for other platform comparisons, average data suggests an increasing trend in positive behavior towards the platforms from the tablet, followed by Ed, and finally Casper. These findings were confirmed through the interviews. One participant explicitly mentioned that the activity was “more fun with Casper.”

In the interviews, Casper was chosen as the preferred platform overall. The main reasons for participants preferring Casper were due to its human-like appearance and expressive face. One participant justified her preference by saying “because there is more emotion shown in Casper. Ed has more of a computer expression.” One participant said: “I found myself looking into his [Casper’s] eyes...” The tablet was preferred over the Ed robot, even though participants showed higher levels of engagement, positive affect, and perceived social intelligence when interacting with Ed compared to the tablet. The participant preferring the tablet emphasized its convenient
size: “it wouldn’t take up any space, and I can bring it with me.” However, those who liked the tablet least stated that it was “ordinary” and “sterile, just directions, like a poster, a sign, or a video in the subway.”

3.4 Chapter Summary

Based on the user studies conducted, it was found that the human-like robot Casper was the overall preferred assistive platform due to its human-like appearance. The robot’s speech, eyes, and facial expressions were found to be the most important features. Results from video data also suggest that the robot’s use of gestures could lead to increased levels of perceived social intelligence. Participants made many assumptions about the platforms’ abilities and behaviors. Most strikingly, the importance of conversing with the platforms was greatly emphasized. The more human-like the technology, the more they treated it as a companion with whom they could converse and interact with in social activities. Indeed, the tablet was regarded as a tool, while Casper was regarded as a companion, and Ed an intermediate between the two. Effective use of multimodal forms of communication, both verbal and non-verbal, is therefore important to the success of the technology. Developing robots that display assistive behaviors using multimodal communication modes that meet their perceived level of social intelligence is a challenge, as human communication follows complex rules that can be difficult to manually define. This thesis proposes to equip a robot with such a behavior skillset by having it learn human-like assistive behaviors directly from expert demonstrators and interactions with its user. A unique architecture combining \textit{LfD} and \textit{RL} to teach a robot socially assisting behaviors is presented in the next chapter.
Chapter 4
An Architecture for Learning Socially Assistive Robot Behaviors

Once the preferred robot social features were identified, a behavior learning architecture was developed for learning behaviors from combinations of these social features. The architecture uniquely uses a combination of LfD and RL to learn and personalize the behaviors to the user’s cognition.

4.1 Proposed Robot Behavior Learning Architecture Design

The proposed robot behavior learning architecture is presented in Figure 6. In the Learning Behaviors from Demonstration module, the robot uses expert demonstrations to first learn the combination of speech and gestures that make up specific assistive behaviors, and second, the environment state-assistive behavior pairs identifying when the robot should display each assistive behavior based on both the robot activity state and user state. The set of all environment state-assistive behavior pairs are stored in a behavior repository. These demonstrated behaviors are labeled according to their speech content and levels of movement activity in the Personalized Behavior Deliberation module. In this module, Q-learning, coupled with an Upper Confidence Bound strategy for exploring behavior selection, is used to teach the robot which of the labeled behaviors transition the user to desirable states based on the User Cognitive Model. A policy for selecting the appropriate labeled behaviors based on the user cognition is learned, and the selected behavior is displayed on the robot by imitating the speech and gestures initially learned from demonstrations.
Learning Behaviors from Demonstration

To teach the robot new behaviors, expert demonstrators physically perform a demonstration of each behavior required for a given assistive activity in front of the robot. The demonstrator’s speech is recorded through a microphone, while a depth sensor is used to track and record their gestures. These demonstrated assistive behaviors are stored in a behavior repository, along with the respective environment states. The robot learns to display behaviors by imitating the combinations of speech and gestures from the behaviors stored in the behavior repository. Supervised learning, in the form of a Classification and Regression (CART) decision tree classifier, is used to learn the environment state-assistive behavior pairs, i.e. a policy to select behaviors during an assistive activity. The robot learns how to display assistive behaviors for the assistive activity by imitating the demonstrated speech and gesture combinations, and learns when to display assistive behaviors based on the environment states.

Figure 6 Socially Assistive Robot Behavior Learning Architecture

4.1.1 Learning Behaviors from Demonstration

To teach the robot new behaviors, expert demonstrators physically perform a demonstration of each behavior required for a given assistive activity in front of the robot. The demonstrator’s speech is recorded through a microphone, while a depth sensor is used to track and record their gestures. These demonstrated assistive behaviors are stored in a behavior repository, along with the respective environment states. The robot learns to display behaviors by imitating the combinations of speech and gestures from the behaviors stored in the behavior repository. Supervised learning, in the form of a Classification and Regression (CART) decision tree classifier, is used to learn the environment state-assistive behavior pairs, i.e. a policy to select behaviors during an assistive activity. The robot learns how to display assistive behaviors for the assistive activity by imitating the demonstrated speech and gesture combinations, and learns when to display assistive behaviors based on the environment states.
4.1.1.1 User States

The user model represents two states: user functioning state, $s_{fn}$, and user activity state, $s_{ac}$, such that $s_u = \{s_{fn}, s_{ac}\}$. The user functioning state represents one of the mental functioning states known to be displayed by seniors with cognitive impairment while performing ADLs [102]–[104] and includes: focused, distracted, having a memory lapse, showing misjudgement, or being apathetic. The user activity state is defined as one of the possible actions performed by seniors with cognitive impairment during ADLs [102]–[104] and includes: successfully completing a step, being idle, repeating a step, conducting a step incorrectly, or declining to continue the activity. The desired user states are $s_u = \{focused, successfully completing a step\}$. The cognitive and activity states defined herein are the symptoms identified in persons with dementia that inhibit them from completing ADL independently [102]–[104].

4.1.1.2 Activity Model

The activity is represented as a set of $M$ sequential robot activity states, $s_r = \{s^1_r, s^2_r, ..., s^M_r\}$. The robot activity states represent individual steps in the assistive activity, which can include for example initiating the activity and instructing a particular activity step. Transitions from one robot activity state to another are a function of both the robot activity state and the user state, i.e. $s'_r = f(s_r, s_u)$. The robot will transition to the next activity step if the user is in a desirable state, otherwise it may remain in the same state or choose to skip a step.

4.1.1.3 Robot Model

The robot is equipped with a set of $N$ behaviors it needs to learn, $B = \{b^1, b^2, ..., b^N\}$. Each behavior $i$ is composed of a set of $n$ communication mode combinations, each containing a combination of speech and gestures displayed by an expert demonstrator, i.e. $b^i = \{cm^i_1, cm^i_2, ..., cm^i_n\}$. $cm^i_j$ is a function of the robot’s arm joint angles ($\theta$) and speech ($sp$), i.e. $cm^i_j = f(\theta, sp)$.

4.1.1.4 Environment State-Assistive Behavior Mapping

The Behavior Learning sub-module uses supervised learning, in the form of a CART decision tree [105], to learn the environment state-assistive behavior mapping for the given assistive
activity. A CART decision tree was used as it provides accurate results even with a small number of demonstrations [106], and can easily handle outliers due to different interpretations or variations across multiple demonstrators without overfitting [107].

In our *Learning from Demonstration* module, the demonstrated behaviors $b^i$ are the targets (i.e. classes) and the environment states, $s = \{s_r, s_u\}$, are the features. CART samples from the pairs of demonstrated behaviors and environment states stored in the *behavior repository*, i.e. $\{s^i, b^i\}$, to learn the behavior classifications.

An example CART decision tree is shown in Figure 7. The root node contains all the environment state-assistive behavior pairs stored in the *behavior repository*. The environment state-assistive behavior pairs contained in each node will be referred to as the node sample. At the root node, a splitting feature is selected to classify behaviors, which consists of either the robot activity state, user functioning state, or user activity state. The samples that satisfy the splitting feature are moved down the left branch, otherwise they are moved down the right branch into two new nodes, shown in Figure 7. Each node is characterized by its impurity $H$, which measures the homogeneity of behaviors contained in a node’s sample [105]:

$$H = \frac{z_l}{Z} G(X_l) + \frac{z_r}{Z} G(X_r),$$  \hspace{1cm} (1)

where $z_l$ and $z_r$ are the number of environment state-assistive behavior pairs moved down the left and right branches, and $Z$ is the total node sample size. The node impurity is based on the Gini Index, $G(X)$, which measures the sample impurity. Both the node and sample impurities vary between $[0.0, 1.0]$: if the sample, $X$, only holds one behavior (i.e. is pure), then $G(X) = 0.0$; if a sample contains two equally partitioned behaviors, then $G(X) = 0.5$. The Gini Index for each behavior $b^i$ in node $m$ is [105]:

$$G(X_m) = 1 - \sum_{b^i} p_{mb^i}^2,$$  \hspace{1cm} (2)

where $p_{mb^i}$ is the proportion of samples containing behavior $b^i$ in node $m$:

$$p_{mb^i} = \frac{1}{Z} \sum_b I(b^i).$$  \hspace{1cm} (3)
A leaf node, or terminal node, is reached when \( H = 0.0 \), i.e. when there is only one behavior contained in the node sample, indicating a behavior has been fully classified, Figure 7. In the end, we have at least one leaf node for each possible behavior.

![Example CART decision tree with four behaviors](image)

**Figure 7 Example CART decision tree with four behaviors**

The developed CART decision tree provides a list of binary rules for classifying behaviors. Given an environment state, the tree predicts the appropriate behavior for the robot to display. However, as each expert demonstrator demonstrated each behavior differently, the robot also needs to determine which behavior variation is appropriate for a given user. A user personalization model is obtained in the *Personalized Behavior Deliberation* sub-system.

### 4.1.2 Personalizing Behaviors Using Reinforcement Learning

Speech and gesture types are used to label behaviors, for example assertive, high-movement activity “motivation” or suggestive, low-movement activity “correction” behaviors. These labeled behaviors are used as actions in a Q-learning algorithm with Upper Confidence Bound action-selection to learn which labeled behaviors maximize the probability of achieving desirable user states.
4.1.2.1 Learned Behavior Labeling

Previous research has shown that behavior demonstrations can be displayed in a wide variety of ways depending on the demonstrator [108]. The learned behaviors were thus labeled according to speech content and movement activity levels.

*Speech labeling*

The sequence of speech used in each demonstration, $l_s$, is labeled as either *suggestive*, *assertive*, or *other*. These labels were chosen as users have shown personal preferences for robots that speak suggestively or assertively to them during an assistive activity [27]. An utterance is labeled *suggestive* if it contains propositional wording such as “can you”, “if you want”, “try”, “maybe”, whereas an utterance is labeled as *assertive* if it contains exclusively imperative wording such as “pull” or “fill”. Utterances are labeled as *other* if they are neither suggestive or assertive, for example, asking social questions not related to the activity.

*Gesture labeling*

Gestures are labeled according to the level of movement activity, $l_{ma}$. Movement activity refers to the amount of movements shown, i.e. how many gestures a person makes in a given time period [109]. Gestures enhance dialog [93], and are particularly important for directing attention and establishing the same context about an activity between two people [22]. The amount of gestures or movement made reflects the degree to which a demonstrator is directing the user’s attention and emphasizing aspects of the environment to explain a concept.

Movement activity levels are defined as high, medium and low and are measured by calculating the change of each arm joint angle across two sequential time frames, where a frame contains the 3D joint coordinates at a given timestep, over the course of the entire demonstration. The average change in all arm angles was taken over the entire behavior demonstration and used as the measure of movement activity for that behavior demonstration, i.e.:

$$\Delta \theta_j^r = \left\{ \left[ \theta_{p,a}^{r,t-1} - \theta_{p,a}^{r,t} \right] \right\} \forall t, p, a,$$

$$l_{ma} = \sum_j \frac{\Delta \theta_j^r}{j},$$
where $\theta_j^r$ represents the robot arm joint angle, $t$ is the time frame, $p$ is the joint position in an arm, $a$ is the selected arm, $j$ is the selected arm joint, and $J$ is the sum of all arm joints.

### 4.1.2.2 User Cognitive Model

The user cognitive model represents the cognitive processes governing the user’s functioning and activity state transitions. The user functioning transition probabilities, $T_{fnc} = P(s'_{fnc}|s_{fnc}, s_r, b_l^i)$, depend on the current user functioning state, robot activity state, and labeled behavior displayed by the robot, $b_l^i = \{b_l^i, l_s, l_{ma}\}$. The user activity state transition probabilities, $T_{ac} = P(s'_{ac}|s'_{fnc}, s_{ac}, s_r)$, in turn depend on the new user functioning state, previous user activity state, and robot activity state. The user model regulates the state transition probabilities in the Personalized Behavior Learning module, where the Q-learning algorithm learns which labeled behavior consistently transitions the user to a desirable functioning and activity state based on their individual transition probabilities.

### 4.1.2.3 Personalized Behavior Learning

In the Personalized Behavior Learning sub-module, the robot uses the speech and gesture labels of each behavior learned in the Learning Behaviors from Demonstration module to personalize its behaviors to its user’s cognition. Given a particular user cognitive model, Q-learning is used to learn the value of selecting each labeled behavior in an environment state. Q-learning was chosen as it is a model-free strategy which does not require learning the exact state transition probabilities [110], which can be very complex for modeling a person’s cognitive processes. Rather, only the value of selecting behaviors in a given state is required.

Q-learning uses a Markov Decision Process (MDP) formulation [110]. Our MDP consists of a tuple $(S, B_s, R, T, \gamma, \alpha)$, where $S$ is the set of environment states, $s = \{s_r, s_u\} \forall s \in S$, $B_s$ is the set of possible labeled behaviors at state $s$, i.e. $b_l^i = \{b_l^i, l_s, l_{ma}\} \forall b_l^i \in B_s$, $R(s, b_l^i)$ is the reward received for selecting labeled behavior $b_l^i$ while in state $s$, $T = P(s'|s, b_l^i)$ is the transition
probability function regulating the change in user functioning and activity state, as defined by \( T_{fnc} \) and \( T_{ac} \) in the user cognitive model, \( \gamma \) is discount factor, and \( \alpha \) is the learning rate.

The robot’s goal is to choose a labeled behavior that will maximize the probability of being in a desirable user functioning and activity state, and thus maximize its reward. The value of all environment state-labeled behavior pairs \((s, b^l_i)\) is given by \(Q(s, b^l_i)\). At every step in the activity, the environment state-labeled behavior values are updated according to the Bellman equation:

\[
Q(s, b^l_i) = \alpha \left( R(s, b^l_i) + \gamma \arg\max_{b^l_i'} Q(s', b^l_i') - Q(s, b^l_i) \right),
\]

(6)

The learned environment state-labeled behavior values, \((s, b^l_i)\), are used to develop a policy \(\pi\) through which the robot chooses the most appropriate (highest value) labeled behavior \(b^l_i\) at every state \(s\). To learn the \((s, b^l_i)\) values, the robot must explore all possible labeled behaviors while still selecting behaviors that yield high rewards. One strategy for doing so is to select all possible labeled behaviors at least once in each state, and subsequently select behaviors based on the probability of yielding a high reward. The Upper Confidence Bound 1 (UCB-1) algorithm uses such an exploration-exploitation strategy.

The Upper Confidence Bound 1 (UCB-1) algorithm [111] is used here to select the behavior with the highest probability of transitioning to the desired user state. Initially, UCB-1 selects all behaviors with equal probability. Over time, it learns with certainty which behavior has the highest probability of yielding the highest reward. UCB-1 is an optimistic policy that minimizes regret, i.e. the difference between the reward from the optimal behavior vs the reward received. UCB-1 has been proven to achieve near optimal regret, and faster convergence than traditional strategies such as decaying epsilon-greedy [111]. UCB-1 guarantees that all labeled behaviors will be selected at some point in the future, irrespective of how poorly it has performed in the past. No labeled behaviors are permanently ruled out, which can be a desirable trait if the user’s cognition changes over time.

Initially, the maximum reward \(r_{max}/1 - \gamma\) is attributed to each environment state-labeled behavior pair, where \(r_{max}\) is the maximum possible reward from \(R(s, b^l_i)\) at any time step. Each
labeled behavior is selected once, providing an initial approximation for the empirical mean reward $\hat{\mu}_i$ of each labeled behavior $b_t^i$. Let $n_i$ be the number of times labeled behavior $b_t^i$ has been observed, and $t$ be the total number of time steps. At each time step, the labeled behavior $b_t^i$ that maximizes $\hat{\mu}_i + \sqrt{2\log(t)/n_i}$ is selected and its empirical mean updated by the reward observed, i.e.

$$b_t^i = \arg \max_{b_t^i} (\hat{\mu}_i + \sqrt{2\log(t)/n_i}).$$

(7)

The result of the proposed Q-learning with UCB-1 exploration algorithm is a policy determining which labeled behavior to display at each state.

4.2 Chapter Summary

In this chapter, a novel architecture for learning combinations of social communication features, environment state-assistive behaviors pairs, and behavior personalization was presented. The architecture uses expert demonstrations to teach the robot speech and gesture combinations necessary to display a set of socially assistive behaviors. A CART decision tree is used to map the behaviors to robot activity, user functioning, and user activity states using the environment state-assistive behavior pairs stored in the behavior repository. The learned behaviors are labeled according to their speech content and levels of movement activity. The labeled behaviors are then used as actions in a Q-learning with Upper Confidence Bound behavior-selection to identify appropriate behavior labels to transition the user to a desirable state based on their cognitive model. This method allows for personalizing the robot’s behavior selection to the user’s cognition, which is important for providing assistance to seniors with cognitive impairments.
Chapter 5
Behavior Learning Experiments

The developed learning architecture was tested both in experiments and in simulation. As the Casper robot was shown to be the preferred robot in our user study, it is used in this experiment. An LfD experiment was designed where participants demonstrated assistive behaviors to the Casper robot. The robot learned the required social feature combinations through imitation. The environment state-assistive behavior pairs were learned using a decision tree and labeled according to their speech and gesture types. The labeled behaviors were used as actions in the proposed Q-learning algorithm to personalize the behavior selection to a simulated user’s cognitive model.

5.1 Implementation on the Casper Robot

To evaluate the effectiveness of the proposed architecture, behavior demonstrations were provided by graduate students in the healthcare sciences. The robot imitated their gestures and speech using external sensors to track the participants’ joint angles and audio data. This section discusses the robot imitation abilities.

5.1.1 Gesture Imitation

The Casper robot only has 3 DOF in each of its arms compared to a human’s 7 DOF, as shown in Figure 8. The robot uses arm joint positions to approximate a demonstrator’s gestures. The Asus depth sensor and the ROS OpenNI tracker package [112] are used for skeleton tracking of the demonstrator. OpenNI uses both RGB and depth images to identify the demonstrator and track their joint positions in 3D space at a rate of 30 Hz.
Vectors representing the demonstrator’s wrist ($\overrightarrow{W_d}$), elbow ($\overrightarrow{E_d}$), and shoulder ($\overrightarrow{S_d}$) positions are evaluated with respect to their torso’s reference frame. To determine the robot’s required shoulder rotation, a vector, $\overrightarrow{SE_d}$, aligned to the demonstrator’s upper arm is defined as the distance between the demonstrator’s elbow and shoulder positions. The $z$-component is set to zero as Casper’s shoulder has no rotation about the $z$-axis (see Figure 9). Similarly, to determine the robot’s elbow rotations, a vector between the demonstrator’s wrist and elbow, $\overrightarrow{EW_d}$, is defined as the distance between the demonstrator’s wrist origin and elbow origin vectors. These vectors are represented as:

$$\overrightarrow{SE_d} = \left( S_x^d - E_x^d, S_y^d - E_y^d, 0 \right) \quad (8)$$

$$\overrightarrow{EW_d} = \left( E_x^d - W_x^d, E_y^d - W_y^d, E_z^d - W_z^d \right) \quad (9)$$
The vectors are illustrated in Figure 9, mapped onto the robot. The dot products between unit vectors $\hat{y}_0$ and $\hat{z}_0$, defined according to the robot’s fixed reference frame, and vectors $\overrightarrow{SE^d}$ and $\overrightarrow{EW^d}$ are used to compute the required rotation angles, see Figure 9. The mapping to Casper’s shoulder ($\theta^r_s$) and elbow angles ($\theta^r_{ep}$ for elbow pitch, and $\theta^r_{ey}$ for elbow yaw) are determined by:

$$\theta^r_s = \cos^{-1} \left( \frac{\overrightarrow{SE^d} \cdot -\hat{z}_0}{\|\overrightarrow{SE^d}\|} \right),$$

(10)

$$\theta^r_{ep} = \cos^{-1} \left( \frac{\overrightarrow{EW^d} \cdot \overrightarrow{SE^d}}{\|\overrightarrow{EW^d}\| \|\overrightarrow{SE^d}\|} \right),$$

(11)

$$\theta^r_{ey} = \cos^{-1} \left( \frac{\overrightarrow{EW^d} \cdot \hat{y}_0}{\|\overrightarrow{EW^d}\|} \right) - \pi/2.$$  

(12)

Figure 9 Visual representation of the demonstrator’s arm joint vectors mapped onto the robot and the corresponding calculated robot joint rotations.
5.1.2 Speech Imitation

Speech recognition is performed using the IBM Watson Speech-to-Text API [113], recorded using a Acoustic Magic Voice Tracker II microphone array. Speech by the demonstrator is segmented by at least 1 second of silence into a string of text (i.e. utterance) used as input to Casper’s text-to-speech module.

5.2 Demonstration Study Setup

For the robot to learn these behaviors, a study was implemented where teachers demonstrated assistive behaviors to the robot. The objective was to determine if the robot can effectively learn the activity from different teachers.

5.2.1 Participants

The recruited teachers were graduate students enrolled in allied healthcare programs at the University of Toronto. These individuals are professionally trained in assisting vulnerable populations using appropriate behaviors, and are expected to be the future teachers of these robots once deployed in real-world healthcare settings. None of the teachers had any prior robotics experience. A recruitment flyer was distributed to each respective department and sent to students through a mailing list. Ethics approval from the University of Toronto Research Ethics Board was obtained prior to commencement of the study. The inclusion criteria for the participants were to be a graduate student (Master’s or PhD) in the allied healthcare fields, and to have clinical training with vulnerable populations (including people with dementia). In total, 15 participants were recruited (Occupation Therapy = 9, Physical Therapy = 1; Clinical Psychology = 1; Biomedical Communications = 1; Speech-Language Pathology = 2; Kinesiology = 1).

5.2.2 Tea-Making Activity

Tea-making was chosen as the assistive activity because our potential users have mentioned needing assistance from a caregiver to perform this activity [85], and because this activity has previous been used with another robot called Ed to assist seniors with dementia in preparing a cup of tea [16]. The tea-making activity was composed of thirteen discretized steps.
First, a demonstration of the Casper robot’s capabilities was first provided to each participant. Participants were asked to demonstrate how they would assist a hypothetical senior with dementia, named Ms. Potts, in preparing a cup of tea. The experiment took place in a kitchen environment having a sink, fridge, an electric kettle, a box of teabags, a mug, spoons, sugar, and milk. The environment, robot placement, and kitchen utilities are shown in Figure 10. A table divided the participant and robot, which stood face-to-face during the demonstrations as shown in Figure 11. A member of our research team was hidden from view to monitor the experiment.

Audio data were recorded using a microphone array placed on the dividing table. The participants’ joint coordinates were recorded at their torso, shoulder, elbow, and wrist using an Xtion ASUS PRO depth sensor positioned behind the robot. RGB videos were also recorded from three separate cameras, situated behind the participant, perpendicular to the participant and robot, and behind the robot. To replicate the participants’ speech, the robot used the output from the Watson Speech-to-Text API in its text-to-speech module.

![Figure 10 Study setup displaying the Casper robot, sensors, and kitchen utilities](image-url)
5.2.3 Procedure

Before beginning the teaching session, the Casper robot performed a demonstration of its purpose and abilities (facial expressions, speech, and gestures). Afterwards, the participants were asked to act out each of the behaviors listed in Table 5. As Casper is a socially assistive and non-contact robot, the participants were instructed not to touch or grasp any of the objects. However, they were encouraged to point to them or make demonstrative gestures. During the demonstrations, Casper recognized the participants’ speech and gestures, and replicated them using its own platform, as shown in Figure 10.

Casper replicated the behavior after every demonstration. The participants were asked to validate the quality of the behavior or make corrections if necessary. If a correction was made, the previous demonstration was discarded.

<table>
<thead>
<tr>
<th>Table 5 The assistive tea-making behaviors demonstrated by our participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Invite the senior to make tea</td>
</tr>
<tr>
<td>Instruct the senior to turn the faucet on.</td>
</tr>
<tr>
<td>Instruct the senior to fill the kettle.</td>
</tr>
<tr>
<td>Instruct the senior to turn the kettle on.</td>
</tr>
<tr>
<td>Converse socially with the senior, e.g. asking about their day, their tea preferences, etc.</td>
</tr>
<tr>
<td>Ask if the senior would like sugar in their tea.</td>
</tr>
<tr>
<td>Instruct the senior to add sugar to their tea (if appropriate).</td>
</tr>
<tr>
<td>Ask if the senior would like milk in their tea.</td>
</tr>
<tr>
<td>Instruct the senior to add milk to their tea (if appropriate).</td>
</tr>
<tr>
<td>Instruct the senior to put a teabag in their cup.</td>
</tr>
<tr>
<td>Instruct the senior to add boiling water to the cup.</td>
</tr>
<tr>
<td>Instruct the senior to stir the contents of their cup.</td>
</tr>
<tr>
<td>Re-engage a senior who is distracted.</td>
</tr>
<tr>
<td>Motivate a senior to finish making tea if they want to end the activity before completion.</td>
</tr>
<tr>
<td>Motivate a senior who no longer wants to make tea and did not respond positively the first time.</td>
</tr>
<tr>
<td>Correct a senior who is putting a teabag in the kettle (doing a step incorrectly).</td>
</tr>
<tr>
<td>Correct a senior who is putting a second teabag in their cup (doing an incorrect step).</td>
</tr>
</tbody>
</table>
5.2.3.1 Questionnaire

At the end of the demonstration session, the participants were asked to complete the questionnaire in Table 6. The questions were designed to assess the teachers’ opinion of the
robot, potential use-cases based on their clinical experiences, and areas of improvement during the teaching process. The questions were left open-ended to avoid biasing their responses.

### Table 6 Questions asked at the end of the demonstration session

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>What was your overall experience like in teaching a robot to do this type of activity?</td>
</tr>
<tr>
<td>2.</td>
<td>Can you think of any other techniques you could use to teach Casper to assist with an activity?</td>
</tr>
<tr>
<td>3.</td>
<td>As a healthcare professional, how helpful would it be to have a robot take on one of these repetitive tasks or activities?</td>
</tr>
<tr>
<td>4.</td>
<td>Can you think of other activities the robot could assist with?</td>
</tr>
</tbody>
</table>

### 5.3 Results and Discussion

Experimental results and discussion for the LfD study, behavior classification, and RL simulation are presented.

#### 5.3.1 Learning Behaviors from Demonstration

The CART decision tree was evaluated using k-fold cross-validation. A 10-fold cross-validation was chosen to evaluate the quality of the model as it maximizes the dataset by using it both for training and testing, has low-variance, and avoids sample bias. In total, 233 interactions (demonstration samples) were used for training and validation. A tree was learned using all of the interactions except 10 which were left out as training data, according to the 10-fold cross-validation process. This process was repeated until all interactions had been used as training data. The cross-validation results showed a 93.0% identification rate for all training iterations. Since all demonstrators were given the same scenario, the environment state-assistive behaviors pairs, i.e. the features, were consistent across the demonstrators. Therefore, as the identification rate shows, it is likely that the classifier would achieve high prediction accuracy on a validation set composed of similar feature distributions. In the future, the decision tree could be tested on experimental data derived from observing a caregiver assisting a senior with cognitive impairments in making tea.
The final decision tree is presented in Figure 12. A rectangular node represents a branch split, including the respective splitting attribute (robot state, user functioning state, or user activity state) and node impurity. The rounded nodes are leaves and represent which of the nine behaviors (see Table 5 for a list of behaviors) was classified.

**Figure 12** The decision tree generated by the CART algorithm showing the splitting nodes as rectangles with their splitting attribute and node impurity, and the rounded nodes representing leaves showing the predicted class

A sensitivity and specificity analysis was conducted on the behaviors predicted by the decision tree during the k-fold validation. The true positive rates (TPR), false negative rates (FNR), and false positive rates (FPR) are listed in Table 7. For each behavior $i$, the true positives (TP) are the instances where behavior $i$ was correctly predicted, the true negatives (TN) are all the other correctly predicted behaviors $k$, the false negatives (FN) are the instances where behavior $i$ was incorrectly predicted to be another behavior $k$, and the false positives (FP) are the instances where another behavior $k$ was predicted to be behavior $i$. The TPR, FNR, and FPR were calculated as:
\[ TPR = \frac{\sum TP}{TP + FN} \]  
\[ FNR = 1 - TPR \]  
\[ FPR = \frac{\sum FP}{FP + TN} \]  

The decision tree was capable of predicting all behaviors with 100% accuracy except “Ask if the user wants sugar” and “Ask if the user wants milk”. The CART decision tree consistently classified the behavior “Ask if the user wants milk” as “Instruct the next step”, and classified the “Ask if the user wants sugar” behavior as “Instruct the next step” in five out of the 12 instances. The last leaf node contained both “Instruct the next step” and “Ask if the user wants milk” in its sample set. It also had a greater proportion of samples containing the “Instruct the next step” behavior, such that the tree has a higher probability of choosing this behavior during classification. Moreover, it is possible that the behavior “Ask if the user wants sugar” also had a splitting threshold close to that of the last leaf node, as they both originate from the same node, which could cause it to be classified as “Instruct the next step”.

Table 7 Mean true positive rate (TPR), false negative rate (FNR), and false positive rate (FPR)

<table>
<thead>
<tr>
<th>Behavior</th>
<th>TPR</th>
<th>FNR</th>
<th>FPR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Invite to tea</td>
<td>100%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Instruct next step</td>
<td>100%</td>
<td>0%</td>
<td>12.8%</td>
</tr>
<tr>
<td>Ask if the user wants sugar</td>
<td>41.7%</td>
<td>58.3%</td>
<td>0%</td>
</tr>
<tr>
<td>Ask if the user wants milk</td>
<td>0%</td>
<td>100%</td>
<td>0%</td>
</tr>
<tr>
<td>Make social conversation</td>
<td>100%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Motivate the user to make tea</td>
<td>100%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Re-engage a disengaged user</td>
<td>100%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Remind the user of a step</td>
<td>100%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Correct a user</td>
<td>100%</td>
<td>0%</td>
<td>0%</td>
</tr>
</tbody>
</table>
5.3.1.1 Questionnaire Results

When asked to evaluate the experience of teaching the robot in the first question, 93% of participants said it was easy to teach the robot and believed the robot would be helpful to them in assisting older adults. While 80% of participants were satisfied with the teaching process, five participants provided alternative suggestions such as having the robot observe two people making tea or using a hand-over-hand technique where the demonstrator physically moves the robot’s arms to demonstrate a gesture. In the last question, being open-ended, several answers were given by each participant. The ranking of the other activities that the robot could assist with were: meal preparation, personal toileting, dressing, housekeeping, providing reminders, doing laundry, and bathing.

Overall, the expert demonstrators were positive about the teaching method employed during LfD training. One recommendation provided to further improve our LfD training was to have a professional actor play the role of the senior with dementia during demonstrations, as demonstrators believed they would feel more comfortable interacting with another person. This type of interaction may also allow for analyzing more complex aspects of human interaction known to be important in HRI, for example facial expressions [114]–[116], gaze directions [50], [95], proximity between two users [117], [118], and the use of backchannels [89]. It should also be noted that while the use of facial expressions was found to be important for seniors with cognitive impairment in the social feature identification study (Chapter 3), the demonstrators’ facial expressions were not analyzed as participants remained neutral for the majority of the demonstration sessions. It is possible that the demonstrators’ facial expressions would be more dynamic if interacting with another person.

5.3.2 Behavior Labeling

Each behavior demonstration was labeled according to its speech assertiveness and movement activity as shown in Table 8. The speech label distribution was identified to be 44.21% assertive, 40.77% suggestive, and 15.02% other. The movement activity label distribution was 11.59% high movement, 48.07% medium movement, and 40.34% low movement. These distributions represent the diversity in providing assistance across all demonstrators.
Indeed, previous research using LfD found major variances in the way participants displayed the designated behaviors [52]. In their case, it became difficult to create generalized robot behaviors due to inconsistency in gestures across participants. The results from this experiment are aligned with their findings. However, rather than seek to create one general gesture from all demonstrations, the present work uses the variations in behavior demonstrations to determine a set of behaviors appropriate for the user’s cognitive requirements. For example, a user with more cognitive decline may require higher levels of movement activity in order to clearly link the spoken directions to interactions with objects in the environment. Pointing at a mug in a tea-making activity informs the senior of both which mug to use and its location in the environment. Another user with less cognitive decline may not require such gestures to understand instructions. The identified variations in displayed behaviors also highlight the need for learning behaviors from a range of demonstrators, since they may communicate information differently based on cultural norms [119] or professional experience. Users may perform better when behaviors are demonstrated by someone who understands a similar mental model.

### Table 8 Labeling distribution for speech assertiveness and movement activity

<table>
<thead>
<tr>
<th></th>
<th>Speech</th>
<th></th>
<th>Movement activity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Assertive</td>
<td>Suggestive</td>
<td>Other</td>
</tr>
<tr>
<td>Number of interactions</td>
<td>103</td>
<td>95</td>
<td>35</td>
</tr>
<tr>
<td>Percentage of interactions</td>
<td>44.21%</td>
<td>40.77%</td>
<td>15.02%</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### 5.3.3 Personalizing Behaviors

The environment states consisted of the robot activity state and the user state, $s = \{s_r, s_u, \}$, and the actions consisted of the labeled behaviors, $b_i^t = \{b^t, l_s, l_m\}$. In total, there were 25 possible states, based on all possible combinations of each environment state, and 9 possible behavior labels, based on the possible speech and gesture combinations of the behaviors, at each state. The rewards were distributed as described in Table 9. The discount factor $\gamma$ was 0.8, and the learning rate $\alpha$ was 0.3.
Table 9 Reward distribution based on the robot state, user functioning state, and user activity state

<table>
<thead>
<tr>
<th>Robot state</th>
<th>User functioning state</th>
<th>User activity state</th>
<th>Reward</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any</td>
<td>Any</td>
<td>Idle, Repeating a step, Conducting a step incorrectly, Declining to continue the activity</td>
<td>-0.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Instruct user to stir their tea Focused</td>
<td>1.0</td>
</tr>
<tr>
<td>Any</td>
<td>Focused</td>
<td>Idle, Completing a step correctly</td>
<td>0.0</td>
</tr>
</tbody>
</table>

**User Personalization**

The Q-learning algorithm was trained on all six combinations of different user cognitive models: User 1 preferred assertive, high-movement robot behaviors; User 2 preferred assertive, medium-movement robot behaviors; User 3 preferred assertive, low-movement robot behaviors; User 4 preferred suggestive, high-movement robot behaviors; User 5 preferred suggestive, medium-movement robot behaviors, and User 6 preferred suggestive, low-movement robot behaviors. Each user had a unique probabilistic cognitive model, each with their own behavior label preferences and cognitive difficulties. The different user profiles were used to investigate convergence of the maximum accumulated rewards across all possible behavior preferences in the cognitive models.

The labeled behaviors were used as input to the user cognitive model. Given a labeled behavior, the user’s next cognitive state was probabilistically chosen from the five possible cognitive states (focused, distracted, having a memory lapse, showing misjudgement, or being apathetic) according to their own user-specific cognitive state transition function. The user’s cognitive state also probabilistically determined their next activity state. Noise was added to the user cognitive and activity transition probabilities to model the unpredictability that can be caused by the onset of dementia [120]. This model guarantees that the user will enter an undesirable state at least once during the assistive activity. The robot’s goal is then to select the most appropriate labeled behavior to transition the user to a desirable state, i.e. “Focused” and “Successfully completing the correct step”.
The Q-learning algorithm was trained five times on each user profile. The accumulated rewards at each timestep in the five training sessions were averaged for each user profile. The overall cumulative reward per episode, where an episode refers to the entire sequence of making tea starting with inviting the user to make tea and ending by instructing them to stir their tea, is shown in Figure 13. As can be seen in the figure, the Q-learning algorithm converged to an policy with a maximum cumulative reward (upper bound) of 0.6 after approximately 85 to 125 iterations, depending on the user profile. These results show that the robot was successfully able to determine an appropriate policy for selecting labeled behaviors that maintain the user focused and completing the step correctly.

Upon convergence, the robot almost exclusively exploits the behaviors with highest Q-values. However, a small drop in User 1’s cumulative rewards, seen at episode 162 in Figure 13, shows that the algorithm continues to explore new behaviors over time, as expected with UCB-1. The algorithm also converged in half the number of episodes required by similar behavior personalization simulations [57].
5.4 Chapter Summary

In this chapter, experimental validation for the proposed learning architecture was conducted. Expert demonstrations of assistive behaviors were provided to the Casper robot, which learned the speech and gestures combinations for assistive behaviors. The environment state-assistive behavior pairs were learned from the demonstrations using a CART decision tree with 10-fold cross-validation. The demonstrated behaviors were classified according to their speech content and movement activity levels. The behavior demonstrations showed a wide variance in both speech and gestures labels, emphasizing the importance of accounting for differences in displayed behaviors. The labeled behaviors were used as actions in the Q-learning with Upper Confidence Bound exploration algorithm to personalize the behavior selection to the user’s cognition. The RL was tested in simulation for four different user profiles, each with their own corresponding user cognitive model. The algorithm converged to an upper bound of 0.6 for all users. These experiments validate the proposed architecture as a robust method for learning socially assistive robot behaviors.
Chapter 6
Conclusion

This chapter provides a summary of the main contributions of this thesis, followed by suggestions on future work, and a concluding statement.

6.1 Summary of Contributions

The main contributions of this thesis consist of: (1) identifying the key dynamic social features for assistive robots; (2) developing a novel architecture capable of teaching an assistive robot combinations of dynamic social features, behavior sequences, and personalization of behaviors to a user’s cognition.

6.1.1 Identification of key dynamic social features

This thesis work is the first to investigate the impact of a robot’s dynamic social communication features on interactions with seniors during an assistive activity. It is one of a few studies where participants directly interacted with multiple platforms of different physical forms. Though previous research using focus groups suggest that seniors are less likely to prefer human-like robot, based on direct interactions with various platforms in the presented study, it was found that a human-like robot was preferred over a character-like robot and a tablet. The Casper robot was preferred due to its human-like face, relatable facial expressions, and gestures. It was found that the robot’s role as a companion was important to this user group, in particular the ability to converse socially with the robot. The human-like robot showed higher engagement levels, was most likely to be used for both ADL and social activities, was regarded as a companion, and was more likely to be stored in social areas of the home. Only one senior preferred the tablet as it could be stored easily. The character-like Ed robot, capable of displaying a smiling and talking face, showed higher levels of engagement than the tablet. However, it was least preferred, suggesting that facial expressions and gestures provide significant value to HRI involving seniors and socially assistive robots.

6.1.2 Development of a novel behavior learning architecture

A novel architecture that uniquely uses \( LfD \) and \( RL \) to teach an assistive robot appropriate combinations of social communication modes, behavior sequences, and behavior personalization
was developed. Expert demonstrations are used to teach the robot combinations of social communication features (specifically speech and gestures) to display a set of behaviors. The environment state-assistive behaviors pairs are used to teach the robot when to display each behavior. As demonstrators each have different ways of displaying the same behavior, the demonstrated behaviors are labeled according to their speech content and movement activities. The labeled behaviors are then used as actions in a RL algorithm to personalize behaviors to a user’s cognitive model.

Whereas previous work on robot behavior personalization has focused on personalizing to groups of people, such as personality types, this work uniquely personalizes behavior selection to a single user’s specific cognition using RL to learn a control policy for selecting labeled behaviors that maintain the user focused and successfully completing the activity step. This mode of personalization is particularly important if the robot is to provide effective assistance to seniors with cognitive impairments. As learning the cognitive state transition probabilities for a user with cognitive impairments can be complex, due to unpredictable behaviors, a model-free Q-learning algorithm was used to learn the behavior personalization policy. UCB-1 was used for action selection as it: (1) has been shown to converge with near optimal regret; (2) converges faster than traditional exploration strategies such as decaying epsilon-greedy; (3) guarantees that all actions will be played in the future irrespective of their current probability of success. The latter condition is especially important as a user’s cognitive model may change over time, and new behaviors may become more effective. It is therefore important that no behaviors are completely ruled out. The model was tested and validated in simulation.

6.2 Future Work

6.2.1 Expanding the set of social communication features

In future work, it would be noteworthy to expand the set of imitated communication features from expert demonstrations. Specifically, it would be beneficial to include facial expressions, as the Casper robot’s facial expressions were important to the seniors in the user study. Facial expression analysis was not included in the presented demonstration study as participants predominantly remained neutral during the study sessions. However, this may change in future
activity demonstration sessions, for example if participants are asked to demonstrate a game-based activity or if the demonstration session involves another person. Gaze direction should also be included as previous research [50], [95] has shown gaze direction to play an important role in mediating dialog.

6.2.2 Incorporation of a senior with cognitive impairment in demonstrations

The expert participants used in the learning-from-demonstration study were asked for alternative proposals for teaching the robot assistive behaviors. Some participants mentioned that interacting with a professionally trained actor representing a senior could also be an alternative set-up for the experiments, as they would feel more comfortable interacting with another person.

6.2.3 User validation of the proposed architecture

It would be interesting to validate the robot’s learned behaviors with cognitively impaired users. A study evaluating user interaction between a senior and a robot randomly selecting behaviors from its behavior repository vs a senior and a robot personalizing its behavior selection to the user’s cognition would validate the personalization model proposed herein. It would also provide feedback on differences in behavior requirements from an assistive robot vs a human caregiver, if any.

6.3 Final Concluding Statement

In conclusion, this thesis investigated the impact of dynamic social communication features in HRI between seniors with cognitive impairment and socially assistive robots, and developed a novel architecture using a combination of LfD and RL to learn behaviors from experts and personalize them to a user. The human-like robot Casper was preferred over a character-like robot and a tablet due to its facial expressions, speech, and gestures. Once the preferred social communication features were determined, a learning-from-demonstration study was implemented to teach the Casper robot how to display the preferred communication modes effectively. The speech and gestures of expert demonstrators were recorded while they demonstrated behaviors necessary to assist a senior with cognitive impairment in making a cup of tea, which the robot imitated. Once the combinations were learned, the robot also learned
when to display the behaviors based on robot and user states using a CART decision tree. The behaviors were classified according to speech content and movement activity. A Q-learning algorithm with UCB behavior-selection was used to personalize behavior selection to a user’s cognitive model by learning which labeled behaviors maintain the user focused and completing an activity step correctly. The architecture was tested and validated in simulation. Future work will look at expanding the set of learned behavior features, modeling the dynamics between caregivers and seniors, and validating the architecture through long-term user studies.
References


Appendix A: Descriptive statistics from the multi-platform interaction study

**Table 10** Descriptive statistics of measured variables from the video data for interactions with Casper (C), Ed (E), and the Tablet (T). The table provides the mean of the number of instances when a participant displayed the respective measured variable with each platform along with the standard deviation, as defined in section 3.2.5 Evaluation Metrics.

<table>
<thead>
<tr>
<th></th>
<th>Engagement</th>
<th>Perceived social intelligence</th>
<th>Positive affect</th>
<th>Trust</th>
<th>Negative affect</th>
<th>Compliance</th>
<th>Collaborative behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>C</strong></td>
<td><strong>E</strong></td>
<td><strong>T</strong></td>
<td><strong>C</strong></td>
<td><strong>E</strong></td>
<td><strong>T</strong></td>
<td><strong>C</strong></td>
<td><strong>E</strong></td>
</tr>
<tr>
<td>Participant 1</td>
<td>0.95</td>
<td>0.74</td>
<td>2.00</td>
<td>2.00</td>
<td>2.00</td>
<td>11.0</td>
<td>6.00</td>
</tr>
<tr>
<td>Participant 2</td>
<td>1.00</td>
<td>1.00</td>
<td>21.5</td>
<td>9.50</td>
<td>2.50</td>
<td>7.00</td>
<td>3.00</td>
</tr>
<tr>
<td>Participant 3</td>
<td>0.92</td>
<td>0.80</td>
<td>6.50</td>
<td>0.50</td>
<td>2.00</td>
<td>10.0</td>
<td>4.00</td>
</tr>
<tr>
<td>Participant 4</td>
<td>0.95</td>
<td>0.93</td>
<td>1.00</td>
<td>1.50</td>
<td>0.00</td>
<td>7.00</td>
<td>9.00</td>
</tr>
<tr>
<td>Participant 5</td>
<td>0.99</td>
<td>0.80</td>
<td>7.00</td>
<td>6.00</td>
<td>0.50</td>
<td>6.00</td>
<td>5.00</td>
</tr>
<tr>
<td>Participant 6</td>
<td>0.92</td>
<td>0.70</td>
<td>10.0</td>
<td>7.00</td>
<td>0.50</td>
<td>10.0</td>
<td>8.00</td>
</tr>
<tr>
<td><strong>Total average</strong></td>
<td><strong>0.95</strong></td>
<td><strong>0.70</strong></td>
<td><strong>8.00</strong></td>
<td><strong>4.42</strong></td>
<td><strong>1.25</strong></td>
<td><strong>8.50</strong></td>
<td><strong>5.83</strong></td>
</tr>
<tr>
<td><strong>Standard dev.</strong></td>
<td><strong>0.03</strong></td>
<td><strong>0.36</strong></td>
<td><strong>7.41</strong></td>
<td><strong>3.60</strong></td>
<td><strong>1.04</strong></td>
<td><strong>2.07</strong></td>
<td><strong>2.32</strong></td>
</tr>
</tbody>
</table>