An Emotion and Memory Model for Social Robots: A Long-term Interaction

by

Muneeb Imtiaz Ahmad

A dissertation submitted in partial fulfillment
of the requirements for the
degree of Doctor of Philosophy (Human-Robot Interaction)
at the MARCS Institute
in Western Sydney University
2018

Doctoral Committee:

Dr Omar Mubin - Senior Lecture, School of Computing, Western Sydney University
Dr Joanne Orlando - Senior Lecture, School of Education, Western Sydney University

Special thanks to all the beautiful children who participated in my studies during the past few years.

keywords: Adaptive Social Robot, Children Robot Interaction, Educational Robots, Social Engagement, Personalisation

STATEMENT OF ORIGINALITY

It is to hereby declare that the work presented in this thesis has not been previously submitted towards the fulfillment of any degree to any educational institute. To the best of Knowledge and belief, this thesis does not contain material written by another person except where due references are made.

Muneeb Imtiaz Ahmad

This work is dedicated to my father Dr. Imtiaz Ahmad and my mother Naila Imtiaz without whom i would not have been able to fulfill my dreams.

ACKNOWLEDGEMENTS

It's a commencement of a long journey when you ride on to the road of your PhD study. Consequently, It was no different for me. As I begin to write the acknowledgement section of my dissertation, I continue to recall all the wonderful people who made it possible for me to finish my PhD.

I would first thank my supervisor Dr Omar Mubin for his continuous and generous support throughout my PhD Candidature. Thank you for accepting my ideas and enabling me to implement them. Dr Mubin has played an instrumental role throughout my candidature and helped me understand the true value of academic life. I cannot recollect the number of occurrences he has helped me in both personal and professional capacities beginning from my first day in Australia until today. In particular, thank you for providing an opportunity to lecture on the Human-Computer Interaction and Advanced Topics in User-System Interaction. I need to confess that it has helped me immensely to broaden my knowledge and clear my understanding of the Subjects. Additionally, it also helped me gain most valuable teaching experience. Moreover, thank you for also encouraging me to take part in a variety of research projects outside of my PhD research. It aided me in improving and enhancing my research skills and abilities.

I would also like to thank Dr Suleman Shahid for motivating me to continue my career in Academia. It was him who convinced me to leave Germany and move to Australia to continue my PhD. It would not be incorrect to highlight that he has been a constant source of advice and encouragement throughout my academic life.

Thank you for your continuous support beginning from providing guidance during my bachelor, master as well as PhD thesis in Pakistan, Germany and Australia. I guess we must have created a record in terms of working together while residing in entirely different parts of the world. Once again, Thank you, Suleman (Bhai), for everything. In the future, I wish to collaborate with you from another destination of the world.

I would also like to thank Dr Joanne Orlando for helping me and providing me guidance during my PhD. In particular, many thanks for helping me understand the field of education and educational technologies. I highly appreciate your guidance with conducting the qualitative data analysis during my candidature.

A bundle of thanks goes to the girls (Mona, Sarah, Valeria, and Nhung) at the MARCS Institute. I have really enjoyed our late evening conversations on a number of things, especially on the Friday Evenings. Special thanks for organising my surprise birthday and my CoC completion celebrations at the MARCS Institute. Many thanks to Yassine Frej, my roommate at the MARCS Institute for helping me with a number of things. To say the least, I highly appreciate and acknowledge your help in finding a nice place for myself and my wife. I would most certainly cherish the time spent with everyone at the MARCS Institute.

Special Thanks to Michele Balogh-Caristo from the St Margaret Mary's School for helping me and also convincing the school authorities to allow me to conduct my studies at the school. A bundle of thanks to all the teachers, children who participated in my studies. All of you are my rock stars. Your willingness to provide feedback helped me complete my thesis.

I cannot stop thanking my beautiful sisters, Ghazia Ahmad and Mehreen Zeeshan and my brother in Law Zeeshan Junaid due to countless reasons. In particular, all of you have given me a peace of mind because all of you are there with our parents to take care of them. You guys have also fulfilled my part of responsibilities towards them. Thank you guys for enabling me to follow my dreams without worrying about my lovely parents.

This thesis acknowledgement cannot be complete without thanking my beautiful wife, Sunbul Muneeb Ahmad. Special thank goes to my wife for her consistent support during my PhD candidature. Most importantly, I appreciate you for letting me enter the house once it was over 7 PM every day and also later feeding me with delicious dinner. Thank you Sunbul for your love, help and support and with our little baby boy on the way, you have filled my life with countless joy and happiness.

In the end, Shukar Alhamdulillah (Thanks to Allah Jee) for giving me the ability to achieve this accord and recognition in my life.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

LIST OF APPENDICES

Appendix

LIST OF ABBREVIATIONS

HCI - Human-Computer Interaction
cHCI - Children-Computer Interaction
HRI - Human-Robot Interaction
cHRI - Children-Robot Interaction
ADSOR - Adaptive Social Robot
AUI - Adaptive User Interface
WoZ - Wizard of Oz
SAR - Socially Assistive Robots
SR - Social Robotics
AVA - Adaptive Virtual Agent
H - Hypothesis
RQ - Research Question
RL - Reinforcement Learning
GIFT - Generalized Intelligent Framework for Tutoring
API - Application Programmable Interface

ABSTRACT

An Emotion and Memory Model for Social Robots: A Long-term Interaction

by

Muneeb Imtiaz Ahmad

Supervisor(s): Omar Mubin and Joanne Orlando

In this thesis, we investigate the role of emotions and memory in social robotic companions. In particular, our aim is to study the effect of an emotion and memory model towards sustaining engagement and promoting learning in a long-term interaction. Our Emotion and Memory model was based on how humans create memory under various emotional events/states. The model enabled the robot to create a memory account of user's emotional events during a long-term child-robot interaction. The robot later adapted its behaviour through employing the developed memory in the following interactions with the users. The model also had an autonomous decision-making mechanism based on reinforcement learning to select behaviour according to the user preference measured through user's engagement and learning during the task.

The model was implemented on the NAO robot in two different educational setups. Firstly, to promote user's vocabulary learning and secondly, to inform how to calculate area and perimeter of regular and irregular shapes. We also conducted multiple long-term evaluations of our model with children at the primary schools to verify its impact on their social engagement and learning. Our results showed that the behaviour

generated based on our model was able to sustain social engagement. Additionally, it also helped children to improve their learning. Overall, the results highlighted the benefits of incorporating memory during child-Robot Interaction for extended periods of time. It promoted personalisation and reflected towards creating a child-robot social relationship in a long-term interaction.

CHAPTER I

Introduction

The use of social robots is becoming more widespread in our everyday lives and our encounters with them in different domain areas (education, industry, telecommunication) are increasing.

A social robot can be defined as an embodied semi autonomous entity capable of performing social interaction after perceiving information from the given environment (Nwana, 1996; Outtagarts, 2009). A social robot may have some of the following capabilities: understanding and displaying emotions, communication with high-level dialogue, learn/adapt according to user responses, establishing a social relationship, react according to different social situations and have varying social characteristics and roles (Dautenhahn, 2014).

It is a common finding in the field of Human-Robot Interaction (HRI) that robots possessing social skills result in establishing human-robot social relationships. Several studies performed in the past showed that robots with social capabilities, including displaying gestures, emotions, and turn-taking were preferred over the robot without having such socio-emotional capabilities (Heerink et al., 2007, 2010). For example, in one study, it has been shown that that humans developed the social association with Roomba, an autonomous vacuum cleaner robot, by calling it with a name or sometimes talking to it (Forlizzi, 2007). In addition, many research studies conducted with robots that possessed human-like adaptation capabilities have also reported in a range of positive findings of the long-term effective integration of robots in various social settings such as education, and health (M. Ahmad, Mubin, & Orlando, 2017b). Furthermore, the use of emotions is an integral part of a social interaction and robots are supposed to make the user believe that they are capable of caring about its environment (Bates et al., 1994) as well as involving the user in engaging interactions (Bartneck, 2003).

1.1 Motivation

The applications of the social robots possessing afore-mentioned capabilities across various social domains are growing. It has been observed that in the recent past, most research with social robots has been performed in the following areas: social robots capable of helping aged-adults at homes (Matarić, 2014), creating robots as social companions (Beer et al., 2017), and utilisation of robots in education (Mubin, Stevens, Shahid, Al Mahmud, & Dong, 2013; M. Ahmad, Mubin, & Orlando, 2017b).

One rapidly growing category of application is found in learning or educational settings. Robots have been previously utilised as tools to facilitate basic programming skills and computer engineering concepts learning (Williams, 2003; Kay & Moss, 2012), gymnastics (Görer et al., 2013), music (Gifford et al., 2011) and other relevant educational fields (Tanaka & Kimura, 2009; Fagin & Merkle, 2002). Most recently, researchers are interested in implementing an intelligent or autonomous robot takes different facilitation roles to engage with children or adults. Many researchers have conducted studies with robots possessing different social capabilities in the educational settings or learning environment where users indulged with the robots in various interactions over time (Kanda et al., 2004; Komatsubara et al., 2014). One of the common findings of these studies were the loss of interest towards interacting with the robot due to the loss of novelty effect. The conjectured reasons for this loss of interest were due to robot's repetitive behaviour and non-flexible design.

Mubin et al. (2013) has identified four different research challenges in the domain of educational robots. One of the challenges is building an adaptive, reactive and proactive pedagogical social robot that is capable of effectively adapting itself based on the user actions in a working (educational) scenario. The need for designing such a social robot has been highlighted by various researchers across different social domains (Matarić, 2014; Dautenhahn, 2014; Wuttke, 2014). Additionally, the significance of designing an interface that can adapt based on the user actions can be witnessed in the field of Human-Computer Interaction (HCI) (WU, 2010; J. Liu, Wong, & Hui, 2003). It has been conjectured that the design of such a social robot may be able to mitigate the problem of the loss in user interest due to the loss of the novelty effect (Kanda et al., 2004).

The design of such an adaptation mechanism for the social robot can be based on various categories of user actions during a Human-Robot interaction. It can be based on the user's emotion, personality, or on the history of previous user actions in a certain situation. Leite et al. (2013) reported a survey on the social robots for long-term interactions and highlighted the use of adaptation based on memory and emotions. They conjectured that such adaptations will lead to personalisation and this will be helpful to address the problem of the loss of interest during HRI for an extended period of time and will also promote engagement and learning during educational interactions. The survey also reflected on the limited research on the use of social robots that can be employed in education domain for repeated interactions in schools. Similarly, They highlighted the need of incrementally implementing novel behaviours in social robots. In general, the mechanism for a machine (a robot, agent, computing device) to adapt based on different user characteristics and then apply it in a real-setting is one of the major challenges in the field of HRI and HCI (Tapus et al., 2007b).

1.2 Research Goals

In this thesis, we propose to implement and evaluate a mechanism for a social robot to perform various adaptations based on the user's memory and emotions. We intend to investigate the effect of these adaptations on the user's social engagement and learning performance during long-term learning or educational HRI. We expect that robots adapting based on the user's emotions and memory will result in creating a social relationship with humans and this will result in sustaining social engagement during long-term interactions. Additionally, we also reflect on the positive effect of the level of student's engagement on their learning as widely reported in the education literature (Carini et al., 2006). We therefore believe that a robot performing aforementioned adaptations will result in promoting learning of the individuals.

Our aim is to implement a computational model for social robots that enables them to perform adaptation based on the users memory and emotions. We later want to study the effect of such a model towards sustaining user social engagement and promoting learning during a long-term interaction. To achieve our aims, we expect to find answers to the following research questions (RQs):

- **RQ1 -** What are the views of teachers and children on the effect of a social robot adapting its behaviour based on their emotions, memory or personality during a educational interaction? [Chapter III]

- **RQ2 -** Should social robot's behaviour adaptation based on user emotions or memory result in sustaining user's (children) social engagement in a long-term interaction? [Chapter IV]

- **RQ3 -** Does users (children) prefer a social robot adapting its behaviour based on user emotions over a robot adapting its behaviour on the user memory in terms of social engagement during long-term interactions? [Chapter IV]

- **RQ4 -** Does a robot creating memories of user's emotional situations effect their social engagement and learning during long-term interaction? [Chapter V]

- **RQ5 -** How does the positive, negative and neutral behaviour or role of the social robot impacts the learning performance during a long-term interaction? [Chapter V]

- **RQ6 -** Does the process of the selection of the robot's behaviour effect the social engagement and learning during long-term interaction? [Chapter VI]

- **RQ7 -** Whether the proposed model/selection/robot behaviour can apply across different learning context? [Chapter VII]

To address the research questions, we began the study that is the focus of this thesis, by conducting user-studies with teachers and children to understand their perspective on the social robot performing various adaptations. We also conducted a long-term study with children to observe the effect of the robot adapting its behaviour based on the user's emotions and memory of the previous interaction on children's engagement. Based on our findings, we devised a model for a social robot that enabled the social robot to create the memory of different emotional events during a long-term interaction. We later used this memory to create an adaptive dialogue for the robot.

The use of memory to inform adaptations has been under-studied in the field of HRI and social robotics. In addition, the effects of such an adaptive social robot on the learning outcomes of individuals during an educational setting in a long-term interaction has also been under-studied (Leite et al., 2013). Moreover, it has been also conjectured that the future of the social HRI lies in the past (Baxter, 2016). Therefore, in this thesis, we attempted to create a mechanism for the robot to create the memory of the user's emotional events and later used it to inform user-specific memory based adaptation. We evaluated the model during a long-term interaction to validate its effect on the social engagement and learning of the individuals. The robot was evaluated in schools with children aged between 9-10 years. To evaluate the model, we first implemented the model on the social robot in two different educational scenario where it was enabled to play the updated version of the Snakes and Ladders game and complete a mathematics challenge with children. The snakes and ladders game was modified to promote children's vocabulary learning whereas the mathematics challenge was designed to learn to calculate area and perimeter of different shapes in a playful manner.

1.3 Expected Contributions

In this thesis, we expect to make the following contributions in the area of HRI and social robotics through addressing the aforementioned research questions:

Contribution 1: Researchers have emphasized the implementation of the mechanism for the robots based on memory and emotions however, there is limited understanding of the comparison on how these different adaptations by the social robot may effect the social engagement of the children in a long-term interaction scenario (M. Ahmad, Mubin, & Orlando, 2017b). Our first *contribution* addressing RQ2 and RQ3 lies in studying the effect of a Social Robot capable of adapting its behaviour based on the memory and emotions of the user on maintaining children long-term engagement. To the best of our knowledge, research that focuses on understanding the impact of different adaptations on the social engagement during long-term cHRI is not available and is needed.

Contribution 2: The need for a computational model for the social robot that enables it to perform adaptations based on the user emotions and memory is highlighted in the literature (Beer et al., 2017). Our second *contribution* addressing RQ4 lies in the creation of a computational model for the social robot. This model directed the robot to create a memory of the information stored by users under different emotional states and later also enabled it to generate dialogue through combining both verbal and nonverbal behaviours based on the memory during a long-term interaction. We also evaluated our model to study the effects of this dialogue towards sustaining the children-robot social engagement and promoting vocabulary learning during a long-term interaction.

Contribution 3: Our computational model enables the robot to provide feedback (positive, negative, and neutral) through combining the memory of emotional events during an educational scenario. We also understand the significance of providing feedback on learning outcome during an education scenario as it can have positive effect on student's learning (Butler & Winne, 1995). In addition, we also find different claims with respect to the effect of emotions on human memory. A body of research shows that emotions enhance human memory in tone, while another claims that emotions enhance central information at the cost of peripheral details (Levine & Pizarro, 2004). Moreover, the impact of robot's feedback during an educational long-term interaction has been under-studied. Considering these claims on the relationship between emotions and memory and the relationship between memory and learning, we believe, it is intriguing to analyse about the impact of robot's emotional feedback on the human's memory. Consequently, our third *contribution* addressing RQ5 lies in studying the effect of the positive (emphatic, encouraging), negative (critical, competitive) and neutral emotional feedback of the robot on children's vocabulary learning during long-term interactions.

Contribution 4: Our findings highlighted the need for robotic tutoring systems where the robot takes a balanced yet positive role and appreciate children's learning efforts. In addition, we also found some evidence of the benefits of using negative feedback such as displaying sadness during feedback results in improved learning (Butler & Winne, 1995; M. Ahmad, Mubin, Shahid, & Orlando, 2017). Based on these findings, we used a reinforcement learning mechanism for the social robot to select an appropriate role (positive, negative and neutral) for each child based on their level of social engagement and learning performance during the interaction as a part of our model's behaviour selection mechanism. To study the effect of this learning mechanism towards maintaining children's social engagement and promoting children's learning performance, we conducted two long-term evaluation studies towards promoting language learning and mathematics learning. In both studies, we compared our findings with the scenario where no learning mechanism was implemented on the robot. The purpose of conducting two different long-term studies was to highlight the scalability of our model across different education scenarios. Therefore, the final *contribution* of our work addressing RQ6 and RQ7 lies in understanding that if the process of the selection of the robot's behaviour effects the social engagement and learning during long-term interaction. Another aspect of our final *contribution* lies in evaluating our model in the educational scenario that was based off a real life curriculum in a real setting.

1.4 Outline

In the following chapters, we define adaptivity and report on the systematic review on the adaptivity in the Human-Robot interaction [Chapter II]. We also report our findings on the teachers and children's perspective on different adaptations portrayed by the social robots in the domain of education [Chapter III]. In [Chapter IV], we compare our findings on the effect of emotion and memory adaptations by the robot on children's level of social engagement during a long-term interaction [Chapter IV]. Based on the findings, we created our emotion and memory model for the social robot that enables the robot to create memory of the past user's emotional events. We also report the results of the effect of our model towards sustaining social engagement and promoting children vocabulary learning during a long-term interaction [Chapter V]. Based on the initial positive findings of our model, we refine the desicion making process of our model. We present a reinforcement learning mechanism for the social robot to select one of the three feedback behaviours (positive, negative and neutral) based on the learning outcome and social engagement of the user. We also present the evaluation results of our refined model during a long-term interaction [Chapter VI]. In [Chapter VII], we also present results of another exploratory long-term evaluation of our Emotion and Memory model during a mathematics learning task in the wild to further show the value of our model. Lastly, we conclude our results, their implications in the field of social HRI and social robotics [Chapter VIII]. The outline can also be seen in figure 1.1.

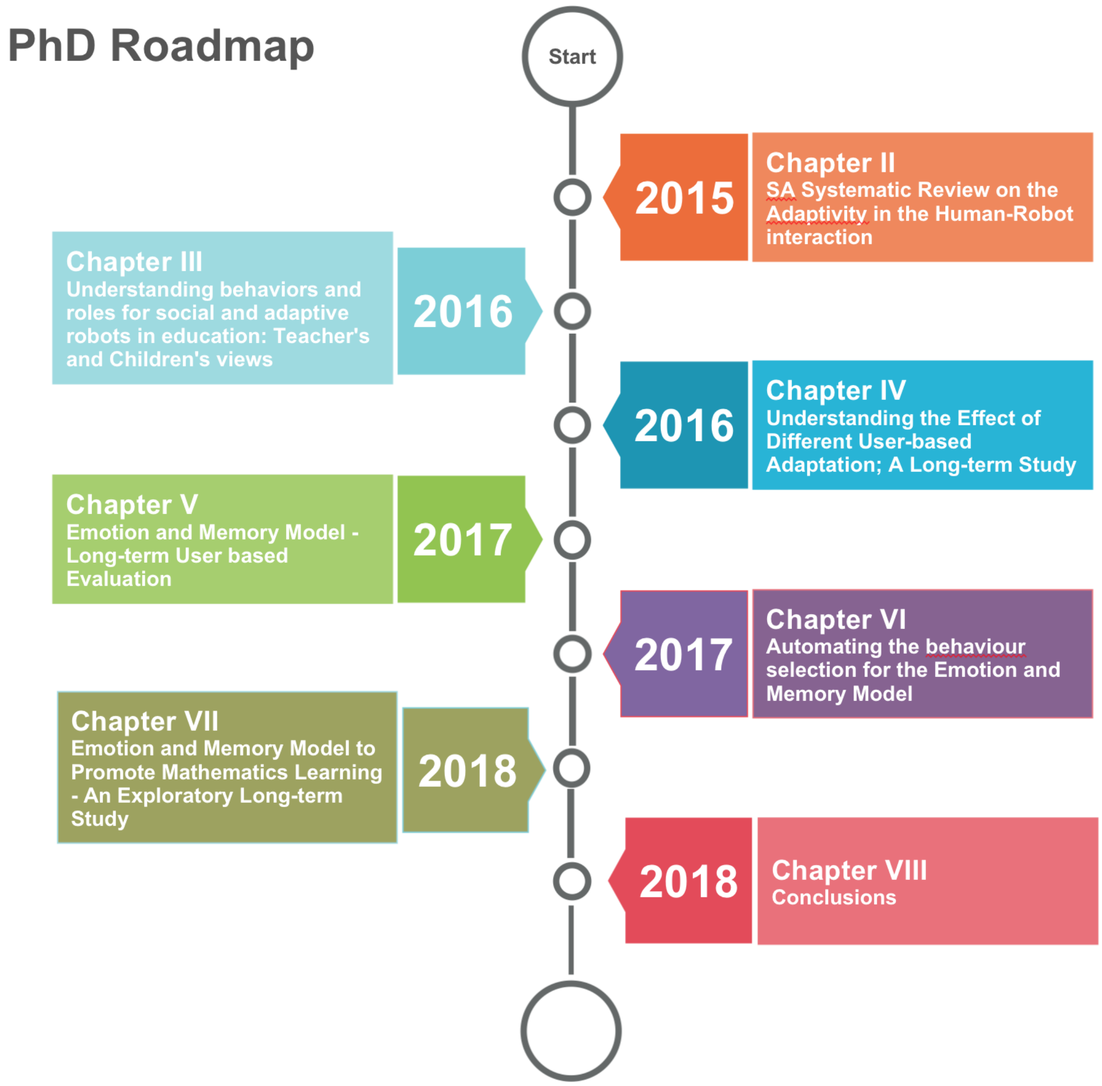

Figure 1.1: Thesis Road-map

CHAPTER II

A Systematic Review of Adaptivity in Human-Robot Interaction

This chapter [1] presents a survey on adaptive interactions displayed by different robots reported in the literature. We have limited the scope of our survey on autonomous or semi-autonomous social robots displaying various adaptive characteristics. Additionally, they have also been involved in various field trials and user-based studies conducted in both controlled and real-world settings. Furthermore, we dont consider industrial robots as part of our sampling strategy. Lastly, we also discuss literature on how has adaptivity been used in non-HRI contexts, in particular with Virtual agents.

2.1 Introduction

An Adaptive User Interface (AUI) can be defined as an agent that monitors user interactions, analyses the interactions to determine usage patterns, stores these patterns, and based on these patterns, presents a personalised interface to the user (WU, 2010).

[1] This Chapter has been published as a journal article - Ahmad, M., Mubin, O. and Orlando, J., 2017. A systematic review of adaptivity in human-robot interaction. Multimodal Technologies and Interaction, 1(3), p.14.

The need for implementing an AUI has also been widely emphasised in the Human-Computer Interaction (HCI) literature (Fischer, 2001). To understand the definition of an AUI in the context of social robotics, it is important to distinguish between an autonomous robot and an adaptive robot. A robot can be categorised as autonomous based on its level of automation such as manually controlled with human-intervention, artificially intelligent or fully autonomous as presented by (Endsley, 1999). A fully autonomous robot however can be defined as a robot that can perform the action without human intervention. An Adaptive Social Robots (AdSoR) is an autonomous or semi-autonomous robot, where speech may or may not be controlled by a human operator through a Wizard of Oz (WoZ), capable of making decisions through perceiving the user information from the given environment. The user information may include their profile, emotions, personality and past interactions (Beer et al., 2014).

Despite an immense amount of emphasis on the need of AdSoRs (Fong et al., 2003), it still remains one of the open issues to implement such robots due to various technical challenges such as mapping of user's emotions and user's personality and keeping track memory of previous interactions in a real-time environment (Tapus et al., 2007a). To overcome these challenges, research is currently utilising various techniques and tools to simulate and portray adaptivity in various social domains. These social domains includes education (Benitti, 2012), public places (Shiomi et al., 2013), domestic and work environments, health care and therapy (Robinson et al., 2014).

We also find a number of research studies that have been conducted on understanding how a user interacts with a robot, or the effect of a robot's social behaviour (Saerbeck et al., 2010), role (Bruce et al., 2002), anthropomorphism, animacy, likeability, perceived intelligence and perceived safety on user's perception Bartneck et al. (2009) in various settings. The applications of robots in various social domains are an evolving phenomenon (Fortunati et al., 2015) and recent research has resulted in positive findings. But, the majority of these interactions are mainly one-off interactions. We witness a handful of applications where their integration is in a long-term or longitudinal setup. Researchers foresee these adaptive robotic interface more applicable for long-term interaction with humans in various social settings (Mubin, Stevens, Shahid, Al Mahmud, & Dong, 2013; Leite, Martinho, & Paiva, 2013). We, unfortunately, also find lesser research on the implementation of AdSoRs that can be utilised in real life settings.

Our research aims are to design, implement and evaluate a mechanism for an AdSoR in a real life setting to challenge the issue of saturation during HRI. We believe that an adaptive robot will help towards overcoming the problem of maintaining long-term social engagement and establishing a social relationship with humans (Komatsubara, Shiomi, Kanda, Ishiguro, & Hagita, 2014; Jimenez, Yoshikawa, Furuhashi, & Kanoh, 2015). Therefore, in order to understand the field of adaptive robots and what can be considered an adaptive feature or interaction, we took on the task of conducting a systematic review on various types of robots adaptive interactions research in the field of HRI. We as a community find a range of studies where robots have been used in various interactions for various purposes, however; we are unsure what can be classified as adaptive behaviour in robots. Although, previously, researchers have reported systematic reviews on the applicability of robots in education (Mubin, Stevens, Shahid, Al Mahmud, & Dong, 2013; Benitti, 2012), long-term utilisation of robots work environments and public places (Leite et al., 2013), domestic settings (Leite et al., 2013), and healthcare (Robinson et al., 2014). But, unfortunately, to the best of our knowledge, we do not find a review based on adaptive interactions in HRI.

2.2 Methodology

We conducted an in-depth literature review on the applications of social robots reported in past research in the following manner. Firstly, we performed an electronic search on the digital platforms such as Google Scholar and Microsoft Academic Scholar. Secondly, we manually searched archives of the top journals and premium conferences in Human-Robot Interaction and Social Robotics as ranked by the Google scholar (Scholar, 2017) from the year 2006 to 2015. We searched keywords that included 'Adaptive social robots', 'Autonomous Robots', 'Adaptation and Robot', 'Applications of Adaptive Robots' and 'Adaptive Robotics Systems'. The search was limited from the year 2005 to the year 2015. Our search resulted in retrieval of articles from a number of venues such as the International Conference on Human-Robot Interaction, and social robotics, International Conference on Humanoid Robots, and Ro-Man. We also found articles published in the international Journal of Human-Robot Interaction and International Journal of Social Robotics. The rationale for choosing Google Scholar was due to the finding that the coverage and reach provided by Google Scholar was more extensive than other similar academic repositories (Meho & Yang, 2007).

2.2.1 Inclusion and Exclusion Criteria

Our criteria for selecting a research article was based on the definition of an adaptive robot as discussed in the introduction section. We did not include articles reported on robots that did not possess user specific adaptation capabilities (Fong et al., 2003) in our review. We read the paper to understand the robot's capabilities before excluding it. We also ignored articles that did not incorporate a user study or a field trial. In addition, we did not review studies that reported qualitative assessments only. Lastly, we did not review research on industrial and commercial robots. After applying all these criteria, we found 37 articles that are reported in this survey.

2.2.2 Coding Scheme

We followed the organisation as reported by Leite et al. to review all of these research papers. We divided the papers based on their social application domain. We are reporting four different domains: 1) Healthcare and Therapy, 2) Education, 3) Public domains and work environments, and 4) homes. Another rationale for choosing these four domains was based on their popularity and usage (Beer et al., 2017). We also find an overlap for some reported articles on adaptive social robots in the case of health care and in-home social domains. For instance, a case where a robot has been used with elderly for assistance at home or at a therapy centre. In such a case, we selected the domain where the study was conducted.

We reviewed articles on robots being applied in various domains on the basis of following criteria. The criteria involved looking for the adaptation features (context, emotion, personality, or memory) and the capabilities (displaying gesture and gestures or communicating through dialogue) of the robot. We present our discussion on the existing adaptive strategies and types of robots used in the research along with a summary of the studies conducted with these robots.

2.3 State of the Art Social Adaptive Robots

In this section, we present the state of the art on the way adaptation is implemented in HRI, organized by application domains.

2.3.1 Health care and therapy domain.

Socially Assistive Robots (SARs) is a commonly used term as introduced by Tapus et al. in the field of HRI and SR. It revolves around all types of robots that can be used to assist people with special needs. In the last decade, we find a number of applications of SARs where they have used with children suffering from autism (Miyamoto et al., 2005), and elderly with Alzheimer's and dementia (Huschilt & Clune, 2012). We are here, particularly interested in reviewing such applications where these SAR's have been designed with a user-specific adaptation mechanism. In this section, we present our results on the existing user studies on these robots where they displaying various adaptive capabilities.

François et al. (2008) presented a conceptual model for a robot capable of adapting its behaviour based on the detected playing styles of an autistic individual in real-time. The interactions with a robot were classified into two classes: gentle and strong depending on the amount of force with which a participant touches the robot. An experimental study was conducted with 5 children to assess the effectiveness of the interaction styles through checking the criteria of both gentle and strong touches. They also measured the number of interactions correctly recognised by an Aibo robot. Experiments performed with the conceptual model showed that the Aibo robot was able to classify the interaction in the real-time. It was also able to adapt to the interaction through changing its own behaviour and therefore, also changing the interaction with the user.

Tapus et al. (2008) and colleagues designed a socially assistive adaptive robot capable of engaging post-stroke users in rehabilitation exercises. A behaviour adaptation system was designed for the ActiveMedia Pioneer 2-DX mobile robot that enabled it to select a behaviour after taking information about user's personality and sensory data. The sensory data involved user detection, navigation and speech recognition. The robot detected speech from the microphone and human user movement was captured through the use of a motion capture device. An experimental study was later conducted with 12 adults to test the system. It was one of the first experimental studies that showed that personality adaptation by a social robot can have a positive effect towards improving user's task performance.

Robins et al. (2008) and colleagues tested the temporal behaviour-matching hypothesis, which predicts that a child will adapt and match robots behaviour during children robot interaction. They used the KASPER robot controlled by Wizard of Oz to play Drumming Call and Response and Gesture Imitation games with 18 children. During both game playing sessions, they measured the effect of the robots response to child's behaviour through controlling the timings and gestures on an interaction of the child with a robot. Results supported the temporal behaviour-matching hypothesis during HRI. As the child adapted according to the robots behaviour in both tested conditions. The conditions included introducing timing delays between a childs and a robots turn taking activity, and when robot displayed various non-verbal behaviours.

Tapus (2009) proposed an adaptive SAR that was able to provide assistance to people suffering from Alzheimers disease. A novel adaptation mechanism that enabled the robot to maximise user's task performance on a cognitive task was presented. The robot was capable of praising or motivating the user based on the user performance that included user reaction time and a total number of correct answers. It was also able to adapt its dialogue based on the updates in the game difficulty levels. A within-subject experimental study was conducted with 9 participants for 6 months in which each participant played the Song Discovery or Name That Tune game in the presence of the Torso robot and a music therapist. The results of the study showed that the adaptation mechanism based on the task performance was successful as the participants recognised the songs with the same probability in both conditions. In addition, a robot was also able to encourage task performance and attention training.

Boccanfuso & OKane (2011) contributed a low-cost social robot, CHARLIE, capable of playing turn-taking imitation game and was also able to also perform face and hand tracking after adapting to a childs non-verbal actions during game play. The authors trained hand and face classifiers through collecting data from children aged 4 to 11. A proof of concept experiment was also performed to check the face and hand tracking during the gameplay. Results showed that hand detector averaged 86% and face detector averaged 92% across all sessions and users.

McColl & Nejat (2013) programmed Brian robot to autonomously provide feedback on user's meal eating behaviours. The robot was capable of selecting an appropriate behaviour based on the following sensory inputs. These inputs include meal tray, utensil tracking, and user state. Meal tray and utensil tracking provided information about meal consumption and position or movement of the utensil on the tray. The user state performed user recognition. Based on these inputs, the robot adapted its dialogue according to user meal consumption or other aforementioned inputs. An experimental study was later conducted with eight individuals to investigate user engagement during meal time. The results showed that participants enjoyed interacting with the robot during the meal.

Wainer et al. (2014) and colleagues programmed an autonomous KASPER robot capable of playing the video game with children diagnosed with autism. KASPERbot was capable of sensing based on the events within the game, planning on different responses based on the sensory data, and acting through gesture, facial expressions, and speech. The authors conducted an evaluation with six autistic children in order to measure children enjoyment and collaboration during the human child and robot-child game-play. Results showed that children were happily willing to play the game with the robot. However, they enjoyed and collaborated more when playing with the human. The authors conjectured that the results might have been influenced due to the novelty factor.

R. Q. Stafford et al. (2014) programmed an autonomous Charlie robot capable of responding to touch, recognising faces, generate speech, and navigating from one room to another in order to manage the health care of elderly people. An adaptation mechanism was implemented that enabled the robot to adapt according to the user profile. The robot scheduled visits, reminded about medications, and also measured blood pressure. The authors conducted a technology acceptance based study with 25 elderly people. Results showed that participant reacted positively towards the use of an assistive robot.

Coninx et al. (2016) and colleagues presented an adaptive robot capable of switching between multiple activities during a single interaction. The adaptive NAO played turn-taking quiz, creative dance and collaborative sorting games on a tablet with the child. Three children diagnosed with grade I diabetes participated in the user-based evaluation. The objective of the evaluation was to access the effect of an adaptive NAO towards richer and more personalised user experience and potential consequences of such an interaction on childrens self-management skills. Results showed children employed the activity switching mechanism actively to customise their interaction with a robot. However, due to limited no of subjects, quantitative findings can be regarded as of preliminary nature.

Ref.	**Robot**	**Study Design**	**Robot Capabilities**	**Adaptive Features**
Francois et al., 2008	Aibo Anthro: Yes	Subjects: 5 autistic children No. of Interactions: 1 Interaction Type: autonomous Measures: interaction style recognition Method: video analysis	provides feedback after detecting the interaction style	user-interaction style based dialogue adaptation

Tapus et al., 2008	ActiveMedia Pioneer 2-DX mobile robot anthropomorphic: No	Subjects: 19 adults No. of Interactions: One-off Interaction Type: autonomous Conditions: (robot vs. Human) Measures: task performance Method: Video analysis	updates dialogue, and human user movement	User's task performance-based adaptations
Robins et al., 2008	KASPER anthropomorphic: Yes	Subjects: 18 children No. of Interactions: one-off Interaction Type: WoZ controlled robot Measures: 1) Duration of the turn taking pause between a child and robot during both games. 2) Child imitating reactions duration. Method: video analysis	plays the game and display gestures and facial expressions.	game-based adaptations
Tapus et al., 2009	Torso anthropomorphic: Yes	Subjects: 9 adults No. of Interactions: once every week for six months Interaction Type: autonomous Conditions: (robot vs. Human) Measures: task performance Method: Video analysis	Display Gestures, Facial expressions, and Utters Speech	User's personality-based adaptations

Boccanfuso et al., 2011	CHARLIE anthropomorphic: Yes	Subjects: 3 children with grade I diabetics No. of Interactions: one-off Interaction Type: autonomous Measures: Speed and accuracy of robot's hand and face detection abilities Method: Video analysis	playing an imitation game and perform face tracking	Detecting faces and imitating hand movements
McColl et al., 2013	Brian anthropomorphic: Yes	Subjects: 8 elderly people No. of Interactions: One-off Interaction Type: autonomous Measures: user engagement Method: Questionnaires	updates dialogue and gesture based on the users eating behaviour	Dialogue and gesture-based adaptations
Wainer et al., 2013	KASPER anthropomorphic: Yes	Subjects: 6 autistic children No. of Interactions: One-off Interaction Type: autonomous Conditions: (robot vs. Human) Measures: Enjoyment, collaboration Method: Video analysis	Display Gestures, Facial expressions, and Utters Speech	Game event-based adaptations

Stafford et al., 2013	CHARLIE anthropomorphic: Yes	Subjects: 25 Elderly No. of Interactions: 3 interactions/child between one and two months Interaction Type: Autonomous Measures: User experience Method: Questionnaires	Speech generation, face recognition, understanding touch sensors, navigation to users room	User profiling (scheduling visits, reminding medications, blood pressure measurements)
Coninx et al., 2016	NAO anthropomorphic: Yes	Subjects: 3 children with grade I diabetics No. of Interactions: 3 interactions/child between one and two months Interaction Type: autonomous with WoZ controlled speech Measures: User experience Method: Questionnaires and logged Data	Switching between activities, display gestures, dances, and Utters Speech	User profiling (name, age, performance, preferences) User emotions detection Memory adaptations

Table 2.1: Summary of Adaptive Interaction studies in the health care and therapy domain.

2.3.1.1 Discussion

All of the reported studies as shown in Table 2.1 resulted in positive results based on their measures such as technology acceptance, user experience, social engagement and task performance. However, it is evident that the diverse user groups from children to adults to elderly in all of these studies points towards more research with these robots to consolidate these results because the effect of age and gender on users attitudes has been highlighted in the past research (Kuo et al., 2009). In addition, limited users have been used in all of these studies. The reasons for the limited number of users is understandable as it is may be difficult to find participants with special needs. Still, we see that adaptive robotic interactions have yielded results that point towards implementing more interactions of this kind such as user profiling (R. Q. Stafford et al., 2014), user emotions (Coninx et al., 2016) and personality (Tapus, 2009).

Another important aspect is that all of these studies except one have used anthropomorphic robots to implement adaptive interactions for this domain. This aspect might be related to the type of adaptation presented in these articles. As most of these adaptations revolve around user's game or task performance, gestures and personality, therefore, we conjecture that anthropomorphic robots might have been considered as the best possible choice for such adaptations due to their social dimensions. In addition, such robots have full capabilities to interact with human users in a holistic way. All of these adaptations have had a positive effect on user's perception however, it is yet to find that how these adaptations will respond when they are tested with a larger number of users and during long-term interactions, once the novelty factor wears off. We must also acknowledge that high levels of humanlike appearance can create expectations on behalf of the user that remain unfulfilled. A review of healthcare robots (Broadbent et al., 2009) asserts that the embodiment of an assistive robot is a delicate issue. The review also suggested that a humanoid-like appearance may not necessarily be the most appropriate one for health care robots as in one of the past studies, participants preferred a robot without a face (Cesta et al., 2007). Conversely, there is also a relationship between the embodiment of a robot and its adaptive behaviour as the humanoid embodiment could drive more human-like behaviours in the robots.

Another critical aspect is that most of these studies such as (Wainer, Dautenhahn, Robins, & Amirabdollahian, 2014; Boccanfuso & OKane, 2011; Robins, Dautenhahn, Te Boekhorst, & Nehaniv, 2008) have utilised video analysis to measure user experience during the interaction. However, we need to consider a general scheme for coding such interactions so that future studies may be compared with each other. The nature of these empirical studies was such that due to the lower number of participants, a control group (such as interaction with humans or proxies) was not used to compare interaction with the robot.

The significance of sensory data used by the robot to adapt is another important issue. The sensory user data capturing mechanism used in the afore-listed studies has been taken from the robot's built-in cameras or cameras located somewhere inside the experimental setup. The data has been mainly used to calculate 1) user emotions, 2) facial recognition, and 3) hand movements. We understand that in case of user emotions, researchers need to focus on the loss of data during interaction. The data loss could occur due to user's moving head during the interaction. Similarly, the data to understand user personality has been taken from the pre-questionnaires conducted before the user study. However, research needs to be focused on more dynamic ways on collecting information of such kind. One of the ways could be enabling the robot to ask questions and perceive their personality traits.

2.3.2 Education

Educational robots have been utilised successfully as a tool to teach programming in schools in the developed world (Williams, 2003). In the future, researchers have envisioned robots to be not just used as a tool to make students understand certain concepts, but help teachers in different ways (Mubin et al., 2013). The usage of robots who will help teachers perform repetitive tasks such as replying to repetitive questions and also act in various social roles (friend, buddy, companion) in an educational setting is a growing phenomenon. In this section, we are interested in reviewing existing AdSoRs that have been utilised in education.

Salter et al. (2007) conducted a trail using Roball, a spherical mobile toy robot, capable of recognising sensory data patterns and later adapting to individuals behaviour patterns during the interaction. The robot was initially programmed to display two behaviours: wandering, and obstacle avoidance. Two different studies were conducted at the laboratory without children before taking Roball in a real life setting. Following the first two studies, three adaptive behaviours were added to the robot and later were tested in a real time environment with children. Results showed that human interaction perceived through proprioceptive sensor has the ability to inform behavioural adaption. However, in order to devise well-informed adaptive robotic systems, we need to consider several other adaption strategies.

Gonzalez-Pacheco et al. (2011) presented an autonomous Maggie robot that was programmed to play different games (Peekaboo, Guessing the character, Hangman, Tic tac toe, and Animal Quiz) with children in order to promote edutainment. Maggie robot constitutes Voice System (ASR, TTS), Vision System (Object Identification), Radio Frequency Identification (RFID), Touch Sensors, built-in tablet screens along with interactions via smart phones, and engagement gestures. The robot was capable of adapting its interaction based on the game scenarios presented to it. A series of experiments were conducted where children participated and played various games such as Peekaboo, Guessing a Character, Tic-tac-toe, Hangman, and Animal Quiz with the Maggie robot. Preliminary findings show that children got more involved and comfortable with the robot as it displayed more interaction capabilities.

J. B. Janssen et al. (2011) presented a study in which children played an adaptive game to learn arithmetic with the NAO robot. The robot was capable of adapting its behaviour based on the mistakes performed by the user during the game. It was a between-subject study design in which NAO was able to change its behaviour in two different conditions. In condition 1, when the user made a mistake, the game complexity was maintained while in condition 2, the game complexity was reduced. Children played the game for three times with the robot and their intrinsic motivation was measured. Results showed that in condition 2, participants showed the higher level of motivation.

Szafir & Mutlu (2012) presented a design of an adaptive robot that was capable of monitoring and improving user involvement during the interaction. The robot acted as an educational assistant using eye gazing, head nodding and gestures as its behavioural features. The agent voice control was also used to gain attention during the interaction. The task of the robot was to narrate the story to the user. The robot capable of adapting based on the user engagement levels and the information about the engagement level was taken from the EEG device. A between subject evaluation study was performed with 30 participants, where three groups, each consisting of 10 participants interacted with a robot capable of reacting in one of the three different ways: 1) low immediacy, 2) random immediacy and 3) adaptive behaviour. These behaviours were calculated based on engagement level of the users. Results showed that the use of adaptive agent was able to significantly improve attention and performance in a narrative task. In addition, they also found the gender difference in terms of motivation during the interaction with the adaptive version of the robot. Females motivation was significantly higher than of males.

Kühnlenz et al. (2013) developed an adaption mechanism in which an EDDIE robot adapts according to users mood and then portrays a similar emotional state. The adaption mechanism involves two different ways of expressing an emotion: implicitly, or explicitly. In explicit scenario, the robot asks the user questions and responds me too, whereas in implicit, the robot generate facial or verbal emotions based on the mood of user measured through questionnaire before the interaction. A 5-step (pre-questionnaire, social sub-dialog, bonding game, picture labelling and post questionnaires) experimental evaluation was conducted later with 84 participants where a robot displayed emotional adaption in four different conditions (Full Emotion Adaption (FEA), Implicit Adaption (IEA), Explicit Adaption (EEA) or no adaption) to measure helpfulness. Results showed that participants ranked FEA highest followed by EEA, IEA and no adaption in order.

Brown et al. (2013) conducted a study with the DARWIN robot that was able to adapt its behaviour during a mathematics test conducted on the tablet device. The robot was capable of adapting both verbal (positive and supportive feedback) and non-verbal (gestures) behaviours during the interaction based on user game performance. A total of 24 students participated during the study to test whether the use of an adaptive educational robot can increase test performance through performing aforementioned adaptations. Results showed that the test completion time was recorded lesser in the case where robot provided supportive feedback, gestures as compared to the no robot interacting with the user during the study.

Ros et al. (2014) implemented a mechanism for the NAO robot that enabled it to adapt according to children dance moves in a long-term interaction study. The NAO robot was capable of updating both of its verbal and non-verbal behaviour based on the current and previous user-state during the interaction. The user-state involved the history of user's dance movement, its profile (ID, name, age, gender) and also current body configuration. The authors conduct a long-term study that involved 18 sessions with 12 children in which the robot taught different dance movements to the children. Results reported a high level of engagement during the interaction and also emphasised on the need of implementing new ways of adaptations to impact the long-term social engagement.

Leite et al. (2014) addressed the problem of sustaining engagement long-term interaction during children robot interaction. She presented an emphatic model for an iCAT social robot capable of understanding and sharing the feelings of the user and later evaluated the model during a chess game. The iCAT robot was capable of emphasising with the children through providing positive reinforcements in terms of facial expressions and dialogue. They conducted an experimental study with 16 children who played the Chess game with iCAT for five times during five weeks. Results showed that iCAT was able to sustain engagement throughout all five sessions.

Uluer et al. (2015) presented a semi-autonomous Robovie-3 (R3) tutor capable of teaching Turkish sign language. R3 was controlled through WoZ while the vision module was functioning autonomously to recognise different signs. The authors conducted an evaluation across three groups (18 graduate students, 6 children with typical hearing, 18 hearing impaired children) to measure robots recognition ability and its effect on users learning performance. The recognition rates for each sign showed by Robovie were consistent (higher than 90%) for all the three groups. In addition, robots has a positive influence on users performances through all groups.

de Greeff & Belpaeme (2015) presented a human teacher and robot social learning scenario in which a robot learns different words by playing a word meaning association game on a surface table. The human teachers begin with choosing the topic and uttering a word, in response, the robot finds the category that strongly associates with the uttered word and communicates the information back to the teacher. The teacher in response provides feedback and the robot adjusts category of the word accordingly. Robots social learning was evaluated through a user study in which 41 subjects participated and played the role of a human teacher and robot displayed two different conditions. In the social condition, the robot showed non-verbal social behaviours (fixating gaze) to inform about learning preference and followed the script whereas in a non-social condition the robot only followed a script. The purpose of the study was to measure robot's learning performance, participant's choice of topic to teach, participant's gaze behaviour and overall user experience. Robot showed better learning performance in a social condition. The gender difference was observed for robots learning performance for both social and non-social conditions. Male participants didnt differ in terms of learning performance in both conditions while female participants were more engaged in the social condition.

Ref.	**Robot**	**Study Design**	**Robot Capabilities**	**Adaptive Features**
Salter et al., 2007	ROBALL anthropomorphic: No	Subjects: 12 children No. of Interactions: one-off Interaction Type: Autonomous Measures: Accelerometers, Tilt sensors. Method: video analysis	moving and avoiding abstavles	user-playing patterns based adaptation
Janssen et al,. 2011	NAO anthropomorphic: Yes	Subjects: 20 children Conditions: between-subject (Personalized vs Group level versions) No. of Interactions: 3 Interaction Type: semi-autonomous Measures: motivation Method: Questionnaires	generate context-aware dialogue during game.	game-event-based adaptation

Szafir et al., 2012	Virtual Agent	Subjects: 30 children - 10 per group Conditions: between-subject (low immediacy vs random immediacy vs adaptive behaviours) No. of Interactions: one-off Interaction Type: Autonomous Measures: user attention and task performance Method: Questionnaires	updates dialogue, controls voice and displays gestures.	user-interaction-engagement based adaptation
Kuehnlenz et al., 2013	EDDIE anthropomorphic: Yes	Subjects: 84 adults Condition: Full Emotion Adaption vs Implicit Adaption, Explicit Adaption vs or no adaption No. of Interactions: one-off Interaction Type: autonomous Measures: robot's helpfulness Method: Questionnaires	displays emotional verbal and facial expressions	emotion-based adaptations

Brown et al., 2013	DARWIN anthropomorphic: Yes	Subjects: 24 children Conditions: between-subject (verbally interactive robot vs non-verbally interactive robot vs verbally and non-verbally interactive robot vs no-robot) No. of Interactions: one-off Interaction Type: semi-autonomous Measures: user engagement Method: Questionnaires	displayed supportive dialogue, gestures during the game.	game-event-based adaptation
Ros et al., 2014	NAO anthropomorphic: Yes	Subjects: 12 children No. of Interactions: 18 Interaction Type: Autonomous Measures: social engagement Method: video analysis, Questionnaires	updates both verbal (text-to-speech) and non-verbal (LED's, head, poses, and dance moves)	user-profiling, memory based adaptation
Liete et al., 2014	iCAT anthropomorphic: Yes	Subjects: 16 children No. of Interactions: 5 Interaction Type: Autonomous Measures: engagement, perceived support and social co-presence Method: Questionnaires	Emphatic dialogue, facial expressions	user profiling (name, performance), user emotions-based Adaptation

Uluer et al., 2015	ROBOVIE anthropomorphic: Yes	Subjects: 3 groups (18 graduate students, 6 children with typical hearing, 18 hearing impaired children) No. of Interactions: One-off Interaction Type: semi-autonomous Measures: learning performance Method: Video analysis	playing game, display Gestures, LED lights and Utters Speech	Gesture specific, LED's specific and speech/sound adaptations
Greeff et al., 2015	LightHead anthropomorphic: Yes	Subjects: 41 adults Conditions: social vs non-social No. of Interactions: one-off Interaction Type: semi-autonomous Measures: Robots learning performance and gaze behaviour Method: Questionnaires and video analysis	playing turn taking language game with a human teacher in order to learn words.	gaze-based adaptation, user-performance based facial and verbal expressions

Table 2.2: Summary of Adaptive Interaction studies in Education domain.

2.3.2.1 Discussion

Table 2.2 summarises the type of adaptations implemented and evaluated in various educational settings. Our review results show that most of the studies conducted in the education domain are based on one-off interaction. In addition, the lesser number of users are involved in these studies. We understand that in a case of long-term interactions, it is difficult to have a huge number of participants, however, we need to conduct studies with more participants to consolidate our results. Another important aspect resulted in two of the presented studies (Szafir & Mutlu, 2012; de Greeff & Belpaeme, 2015) was the effect of gender during human-robot interaction. As it was reported that female participants were more social as compared to male, therefore, more research on the perception of adaptive robots on genders need to be performed. We conjecture that research needs to reflect on the way a robot can adapt its behaviour based on the gender of the user.

We also find an overlap on the types of adaptation implemented during these studies. Games have been used in most of these studies such as (J. B. Janssen, van der Wal, Neerincx, & Looije, 2011; Brown, Kerwin, & Howard, 2013) as a medium of communication on different devices and most of the adaptation revolves around adapting robot's dialogue based on the game events. Other adaptation factors were based on user's performance, or game outcome (Leite et al., 2014). In addition, a few studies (Uluer et al., 2015; de Greeff & Belpaeme, 2015; Leite et al., 2014) have also utilised non-verbal adaptation based on the sensory input received through the facial scan. We believe that more research needs to be conducted on implementing these adaptations grounded on a theoretical framework. In addition, adaptive robot interactions have not been utilised based on various aspects such as memory, user's personality. Most of the adaptations by the social robot revolved around reacting on interaction events or understanding user-emotions through facial scans.

Anthropomorphism is another interesting aspect as all of these studies presented in table 2.2 have utilised social robots. We conjecture the reasons can be based on the type of domain as a robotic tutor will be envisioned in a shape of a living being. It has also been showed in one of the studies where children were asked to design robots to used in an educational setting (Obaid et al., 2015). However, more research needs to be conducted to address the issue revolving around the appearance of an AdSoR in education.

2.3.3 Work Environments and Public Spaces

The penetration of social robots in work environments and public spaces is growing. We find the use of robots for advertisement at the shopping mall (Kanda et al., 2009), as waiters at the hotels (Asif et al., 2015) and also at the banks in Japan (Yamamoto, 2014). In addition, we also find the utilisation of commercial robots being applied in various ways to the public spaces in the developed world. As we witness the applications of robots in such settings, it is certainly important to implement novel means of adaptations for these robots to apply them in public spaces in various ways. Therefore, it is also significant to reviews the applications of various adaptive interactions in these scenarios.

Hoffman & Breazeal (2007) presented an adaptive action selection mechanism for a robotic teammate during a collaborative task. They proposed the use of an educated anticipatory action selection in an agent (robot) based on expectations of each others behaviour. However, the action sequence was based on an assumption that human collaborator will follow a roughly consistent set of actions. In order to compare the performance of the selection process, a reactive agent was also implemented. Both, a robot with anticipatory adaption and reactive mechanism were tested with a human teammate in a game based task where the robot and a human-mate worked collaboratively to build 10 carts. The humans role was to bring parts (a floor, a body, two kinds of axles, and two kinds of wheels) to the workspace and robots role was to attach the car parts using the tools (the welder, the rivet gun and two wrenches). Results showed that participants performed significantly better in the adaptive anticipatory case compared to the reactive case.

Svenstrup et al. (2008) conducted a field trial by placing a FESTO robotic platform at a shopping mall capable of identifying, tracking and following individuals in a natural way. The robot randomly roamed around the mall until an individual is identified. Once detected, the robot started smiling and playing the jingle-bell song and tries to follow the person until lost at a safe distance. Once done, it updated its expressions and starting roaming again. Results of the trial conducted with 48 participants showed that people had in general positive attitude towards the use of a social robot in this environment. However, some reservations were reported with respect to maintaining the distance between the human and the robot during the interaction.

M. K. Lee & Forlizzi (2009) presented a Snackbot robot capable of autonomously navigating in the hall and delivering snacks to the people. The authors addressed the challenge of maintaining engagement with the robot in terms of user interest during the long-term interaction. They implemented an adaptation mechanism that enabled the robot to adapt according to the user preferences that includes snake choice and snake usage patterns. An experimental study was also later conducted at a workplace with 21 participants where the robot acted as a delivery person for a span of 4 months. Results of the long-term service robot interaction revealed that participants attached social roles to the robot that were beyond the delivery person. For instance: The interaction triggered new behaviours among employees such as drawing social comparisons and even jealousy (Lee et al., 2012).

Kanda et al. (2010) conducted a field trial with a Robovie-IIF robot capable of detecting people and guiding them by providing directions in the shopping mall. The robot was programmed to identify users based on the RFID information and also provide information based on the previous interactions. A total of 235 participants interacted with the robot during 25 days that result in a sum of 2642 interactions. This interaction were later coded to measure user-experience along with questionnaires that were used to measure user perception of the robot. Results showed that the participants encouraged the use of robots in the mall for aforementioned purposes.

Shiomi et al. (2013) and colleagues presented a semi-autonomous, speech recognition was controlled by the human, Robovie-II, and Robovie-miniR2 robots capable of acting as an advertising agent in the mall. Both Robovie robots were also capable of displaying speech utterance, gestures, and non-verbal behaviours according to persons action. The authors conducted a field trail in order to measure the effect of robots presence on peoples participation and overall advertising process. The trail consisted 256 individuals who interacted with three robotic conditions GUI based robot, Robovie-miniR2, and Robovie-II. In order to measure the effect on users participation and overall effect on advertising, three observations (total interacting users, total printed coupons, and interaction initiation) were coded through video analysis for all three conditions. Results showed that a maximum number of people interacted, printed coupons, and initiated interaction first with Robovie-miniR2, after Robovie-II and GUI in a descending order.

Sekmen & Challa (2013) present an autonomous mobile robot capable of learning through adapting the behaviour and preference of the interacting user. The authors contributed a learning model that enabled the robot to update itself every time a user interacts with a robot. In order to update the model according to user state and preference, a Bayesian learning mechanism was implemented. The learning model utilised various robots capabilities such as face detection and recognition, speech recognition and localisation, natural language understanding, Internet information filtering, and navigating. Following the learning model, the robot acted as a tour-mate on a university campus and 25 students at the university were recruited to evaluate the robotic tour-mate in two different conditions (adaptive, and non-adaptive). In general, results showed that participants preferred an adaptive tour-mate robot to a non-adaptive one.

Rousseau et al. (2013) reported an IRL-O- interactive omnidirectional robot platform that can be used to combine both verbal and non-verbal modalities in order to perform engaging interactions with people in controlled and real-world settings. The IRL-0 was programmed to autonomously engage with people in different interaction scenarios through displaying these verbal (voice (V)) and non-verbal (Facial Expressions (FE), Arm Gestures (AG), Head Movements (HM)) modalities. The robot initially detected legs of the user, once detected the robot walks towards the user and maintains a socially acceptable distance from the user. Later, it asks user for assistance and validates it. If the user has been previously detected and has refused robot's offer for assistance, the robot convinces it. If not convinced, the robot says goodbye and goes back to its original position. The IRL-O robot's interaction modalities were later tested in a within-subject experiment with 35 participants comprising of four conditions. In condition 1, it had V and FE and in condition 2, it had V, FE, and H and in condition 3, it had V, FE, AG and lastly in condition 4, it had all four modalities. The participants were asked to give their preferences on these modalities on IRL-O based on their use. Participants found V, AG, FE, HM to be useful by 100%, 77%, 31%, and 50% respectively. A field study was later conducted with IRL-O by placing it in a work environment for two weeks where 381 users stopped and interacted with one of the modalities of the robot at a museum. The user preferences on each modality were measured through coding which modality made the user stop and interact with the robot. Results showed that voice and facial expressions were the key reasons for the user to stop and interact with the robot.

Aly & Tapus (2013) described architecture of a NAO humanoid robot capable of autonomously adapting according to users personality and later displaying combined verbal and non-verbal behaviours. An evaluation study was conducted with 35 participants through following steps; 1) NAO robot was able to autonomously identify participants personality through dialogue, 2) the robot asks the participant to choose a restaurant from a given list, 3) during the response delays, the robot generated appropriate speech and gestures. During the interaction, the participant did ask details about different restaurants. The goal of the study was to find if a robot that matches users personality and displayed appropriate gesture was perceived more expressive. Results showed that personality played an important role as user perception and preference was influenced by the personality portrayed by a robot.

Keizer et al. (2014) programmed an iCAT robot that played the role of a socially aware bartender robot. The robot was capable of detecting customer, tracking multiple customers, and taking their orders. To enable these capabilities in iCAT, the authors presented two implementation mechanisms for a Social Skill Recognizer (SSR). The input parameters for SSR includes location, facial expressions, gaze behaviour, and body language of all the users in the environment. An experimental evaluation was conducted with 37 adults to compare the two implementations of the SSR. In one implementation, they implemented a rule-based SSR where rules were hard-coded in the system and in another, they programmed a trained SSR. The purpose of the evaluation was to measure the detection rate, initial detection time, drink serving time and a number of engagement changes during the interaction. Results showed that trained SSR was found be more responsive in terms of a number of engagement changes however, no significant differences were observed.

Shiomi et al. (2015) presented an autonomous wheelchair NEO-PR45 robot capable of adapting its speed and speech using the preferred speed information and pre-defined small talk base on user's position and registered map information. A within-subject experimental study with 28 elderly people was participated in three conditions. In condition 1, the wheelchaired robot moved automatically at a fixed speed. In condition 2, the robot performed both speech and speed based adaptations. Lastly, in the third condition, the caregiver wheeled the participant. The purpose of the study was to measure the degree of comfort, enjoyment and easiness to make the request to the robot. Results based on quantitative questionnaire and interviews showed that adaptive wheelchaired robot was rated higher in terms of enjoyment and easiness than the simple wheelchair robot and caregiver.

Kato et al. (2015) and colleagues conduct a field trial to compare three different adaptive behaviours of a social robot at a shopping mall. The behaviours include: 1) the robot autonomously decides whether to approach the user based on the intention estimation 2) simply proactive, where the robot aimed at approaching everyone in the experimental field and 3) passive, where a robot waits until a visitor asked an inquiry. The authors measured the amount of robots successful, failed and missed attempt to initiate interaction with people who intended to interact with it. Results showed that the ratio of success in the proposed condition is significantly higher than simply proactive and passive condition.

Dang & Tapus (2015) presented an autonomous NAO robot capable of playing Operation Board game in order to measure players stress level through collecting players heart rate and performance during the game. The game was able to generate true and false alarms based on user action categories as normal and stressful alarms. In addition, the robot adapted the coaching style according to players personality. The authors focused on the role of NAOs personality adaption through verbal reactions in order to motivate people and help them improve their performance. A within-subject experimental study was conducted with 17 participants through following 5 different steps (Introduction to the game, players personality identification, recording of heart rate baseline, game play, rate different coaching styles of robot) in four different conditions (with or without Robot in normal alarm system and with or without Robot in stressful alarm system). Results showed that the players performed better when the robot coached them. There was a correlation between participants personality and their preference about the robots personality. The heart rates exceeds in case of false alarms and it was also reported that introverted player have greater heart rates as compared to extroverted ones.

P. Liu et al. (2016) presented a model for an autonomous generating socially acceptable pointing behaviours during an interaction with the robot. In order to understand the behaviours, human-human interaction was monitored and interviews were conducted which resulted in three categories of pointing behaviours. The behaviours include gaze only, casual pointing or precise pointing. Through these observation, a behaviour selection model was implemented that enabled the Robovie robot to automatically calculate appropriate deictic behaviours through interacting with the user that included speech recognition and user tracking system. An evaluation of the behaviour selection model was later conducted at a shopping mall in which a total of 33 participants were hired to measure the naturalness, understandability, perceived politeness and overall goodness of the robot's deictic behaviours. Results show that participants perceived behaviours polite and natural while the understandability was not perceived well.

2.3.3.1 Discussion

Our discussion on the summary of adaptive interactions in public spaces and work environments as shown in table 2.3 will address the following issues. Anthropomorphism is one of the issue that needs attention during adaptive interactions in public spaces. As one of the studies (Shiomi et al., 2015) showed that the medium sized robot was preferred to be the best choice of interaction at a shopping mall. However, we need to perform more studies on the validation of the right size for a robot. In addition, all studies have utilised wheeled and anthropomorphic robots however, we need to explore robots with legs and find user preferences on the use of such robots in public spaces.

Navigation was found to be one of the critical issues during HRI at public spaces and work environment. The adaptation of a robot on when to navigate towards an interacting user and how to address the user is a challenge. Similarly, the distance between the robot and the user has been guided by (Sisbot et al., 2005) however, studies have still reported intimidation on the user side (Svenstrup et al., 2008). Therefore, we need to conduct studies involving more participants to consolidate existing findings.

The robot's adaptation has been majorly focused on the user preferences in most of these studies. All of these studies have resulted in positive findings however, it also brings attention towards the utilisation of other sorts of adaptations during HRI in these settings. For instance, A human user advertising about a certain product can forget about a frequent user but if a robot adapts on these parameters, it may result in positive findings. Therefore, we believe more research needs to be conducted on implemented various ways of adaptations to find its benefits. As one of the reported study also showed that the engagement can be improved when robot adapted based on the social situations and another also reported the effect of adaptive version of a robot in comparison with a non-adaptive version.

Another issue is based on the behaviours a robot should portray after adapting based on the type of input. We need to conduct more studies to compare the effect of various verbal and non-verbal behaviours in a long-term setting in work and public environments.

Ref.	**Robot**	**Study Design**	**Robot Capabilities**	**Adaptive Features**
Hoffman et al., 2007	Virtual Agent	Subjects: 32 adults No. of Interactions: one-off Interaction Type: autonomous with WoZ controlled speech Measures: Time taken to complete a task. Method: logged Data	attach car parts through anticipating actions	User and game based adaptations

Svenstrup et al., 2008	FESTO anthropomorphic: Yes	Subjects: 48 adults No. of Interactions: one-off Interaction Type: Autonomous Measures: user experience Method: Questionnaires and interviews	detecting individual, playing music and showing expressions	user-identification based adaptation.
Lee et al., 2009	SnackBot anthropomorphic: Yes	Subjects: 21 adults No. of Interactions: four-months field study Interaction Type: Autonomous with WoZ controlled speech recognition Measures: user experience Method: Questionnaires and interviews	utters context-aware speech, and deliver food through navigating inside the hall	user-preference based adaptation.
Kanda et al., 2010	Robovie-IIF anthropomorphic: Yes	Subjects: 235 adults No. of Interactions: 25-days field trial Interaction Type: Autonomous with WoZ controlled speech recognition Measures: enjoyment and social interaction, visitor's perception Method: Questionnaires and video analysis	identify users, providing shopping information, route guidance, inquire personal information	user and memory-based adaptation.

Shiomi et al., 2013	Robovie-II Robovie-miniR2 anthropomorphic: Yes	Subjects: 256 (only interacting users) Conditions: GUI, Robovie-miniR2, and Robovie-II. No. of Interactions: field-trial Interaction Type: Autonomous with WoZ controlled speech recognition Measures: Robots learning performance and gaze behaviour Method: Questionnaires and video analysis	utters speech, gestures, and non-verbal behaviours according to persons action.	User specific gesture, dialogue based adaptations
Sekmen et al., 2013	Pioneer 3-AT mobile robot anthropomorphic: No	Subjects: 25 adults Conditions: adaptive vs non-adaptive. No. of Interactions: one-off Interaction Type: Semi-autonomous Measures: user preference Method: Questionnaires	detecting and recognising face, and speech, understanding natural language, filtering information from Internet and navigating through following the map	Speech and user-based adaptation.
Aly et al., 2013	NAO anthropomorphic: Yes	Subjects: 35 children No. of Interactions: one-off Interaction Type: WoZ controlled robot Measures: engagement Method: Questionnaires	utters speech and display gestures	personality-based adaptations

Rousseau et al., 2013	Robovie anthropomorphic: Yes	Subjects: 381 visitors No. of Interactions: field-trail Interaction Type: Autonomous Measures: user preference on robot's behaviours Method: video analysis	detecting users, navigating toward the user, facial expressions, head movements, and arm gestures	user-identification based adaptation.
Keizer et al., 2014	iCAT anthropomorphic: Yes	Subjects: 37 adults Conditions: Rule-based Social Skill Recognizer (SSR) vs trained SSR. No. of Interactions: one-off Interaction Type: Autonomous Measures: detecting customer, detection time, system response time, drink-serving time, number of engagement changes. Method: Questionnaires and video analysis	detecting customer, track multiple customers, serve drinks and take orders.	Speech and user-based adaptation.

Kato et al., 2015	Robovie anthropomorphic: Yes	Subjects: 26 visitors Conditions: intention estimation algorithm vs proactive vs non-adaptive No. of Interactions: field trail Interaction Type: Autonomous with WoZ controlled speech recognition Measures: interaction intention success rate, Method: video analysis	detecting visitors and interacting with them.	user-interaction-intention based adaptation.
Shiomi et al., 2015	NEO-PR45 anthropomorphic: Yes	Subjects: 28 adults Conditions: simple-robot vs adaptive robot vs human-caregiver. No. of Interactions: one-off Interaction Type: Autonomous Measures: ease of user, enjoyment, degree of comfort Method: Questionnaires and interviews	provides speech based feedback and adapts speed based on user preference.	Speed and speech-based adaptation.

Dang et al., 2015	NAO	Subjects: 17 adults Conditions: Robot vs no-robot in normal alarm system and stressful alarm system No. of Interactions: one-off Interaction Type: semi-autonomous Measures: interaction preference Method: video analysis	generate true or false alarms through speech and gestures	user-stress-level and personality based adaptations
Liu et al., 2016	Robovie anthropomorphic: Yes	Subjects: 33 adults No. of Interactions: one-off Interaction Type: WoZ controlled robot Measures: naturalness, understandability, perceived politeness and overall goodness of the robot's deictic behaviours Method: Questionnaire	displays deictic behaviours through gaze, casual and precise pointing	user pointing behaviours based adaptations

Table 2.3: Summary of Adaptive Interaction studies in work environment and public spaces.

2.3.4 Domestic Settings

Smart homes is a commonly used term these days. Researchers believe that in the future, we will have robots at home helping us in various ways such as butler or influencing our behaviours (Kidd & Breazeal, 2008; Srinivasa et al., 2010). They will act as chefs, caregivers and cleaners. We find an example of a commercial robot Roomba, a vacuum cleaner robot (Forlizzi & DiSalvo, 2006) however, we don't believe it is an adaptive social robot as it does not fall under the definition as specified earlier and also in the literature (Fong, Nourbakhsh, & Dautenhahn, 2003; Fink, Mubin, Kaplan, & Dillenbourg, 2011). In addition, in order for a robot to integrate at homes, it needs to have a social mechanism. We speculate that an ASR is an excellent candidate for the aforementioned job. Therefore, before such robots are designed it is important to research about their effect on user's perception and their overall experience.

Torrey et al. (2006) studied the effect of an adaptive dialogue during HRI. The robot was autonomously programmed as a chef that was able to adapt according to individuals cooking expertise (novice, expert). Two appropriate conditions were designed for individual expertise. In one condition suitable for experts, the robots asked the participants to identify cooking tool by their name and in another condition suitable for novices; the robot not only named the tools but also described them in couple of sentences. Two different experiments were conducted to measure information exchange and social relation for two different conditions during users interaction with a robot. Results showed that appropriate adaptive dialogue improved information exchange for novices but didnt show an effect on experts. Adaptive dialogue didnt effect social relation however, when time pressure was introduced to finish a task, the adaptive dialogue improved social relation for both novices and experts.

Torrey et al. (2007) presented a robot chef capable of adapting dialogue based on robots awareness about human gaze and task progress. A trial was later conducted with two groups (experts and novices) possessing expertise in cooking, to measure the effect of adaption on task performance and communication between a user and a robot during four different conditions, 1) Question Only, the robot was able to respond to individual questions through a text based interface, 2) Gaze Added, the robot made decision on the bases of participants gaze activity, 3) Delay Added, the robot made decision based on task progress, and 4) Immediate Added, the robot provided information immediately without considering users task progress. The task was similar to Torrey et al. (2006). Task performance results showed there was a main effect for expertise; experts took less time to finish the task than novices. In addition, the effect of condition was not significant, however, participants made fewer mistakes in Immediate Added condition as compared to other three conditions. Communication measures results showed that there was significant effect for condition and expertise. In addition, the effect of condition was not significant, however, participants asked significantly higher number of questions for Question only than the Delay Added condition. In general, results didnt find any benefits with respect to addition of gaze awareness during human-robot dialogue. In addition, researchers suggested the need of more research on gaze awareness.

Gross et al. (2011) presented a CompanionAble robot to help elderly in home environments. It was a anthropomorphic robot that comprised a touch-based graphical interface, two OLED displays as eyes to express emotions, a tray to carry objects in the house and a docking battery for recharging. The robot was capable of recognising, detecting and tracking the user at various places in the house (Gross et al., 2015). The work continued for three years and later an adaptive version of the robot was programmed and evaluated. The version comprised a robot-based health assistant that measure the different parameter such as pulse rate, oxygen level and give suggestions if the physical workout is required of the user. In addition, a better eye display was developed that enabled the robot to display its internal emotions (boredom, being a surprise, listening or sleeping) on user's movement with the eyes. Moreover, the robot also had a "stroke sensor" on its head that was able to distinguish various activities of the user (slap on the head, stroke or tickle). A user trail was conducted with this version with 9 participants living independently in their private apartments to measure the acceptance of robot as a social companion.In general, participants felt safe around the robot but had reservations about leaving the robot alone in the apartment.

Overall, the integration of robot's technology into homes was appreciated.

Cooney et al. (2014) presented a sponge robot (a small humanoid) capable of providing enjoyment to individuals who play with it by hugging, shaking and moving it in various ways. The authors addressed the challenge of understanding how a robot can realise what actions people perform during play and also how this information can enhance enjoyment. In order to solve the problem, typical full body gestures collected through an observational study were mapped onto the robot. The robot then suggests different enjoyable ways through understanding the interaction. The sponge robot was evaluated with 20 Japanese participants for the different condition (rewards vs. suggestions and naive design (behaviours based on intuitive knowledge) vs. proposed design (meaningful motions, rewards, suggestions)) in a within-subject design in order to measure how to play, perceived variety, control, intention, and enjoyment. Results showed that participants rated rewards contributed significantly to perceived variety and enjoyment whereas suggestion was rated significantly to how to play. However, no significant difference was observed for control and intention. In addition, proposed system did provide more enjoyable and interactive play than a naive design.

Youssef et al. (2016) contributed towards a research question that how communication protocols based on knocking could be developed between a human and a robot on a sociable dining table. In order to construct a communication protocol for a robot, a human and wizard of Oz controlled dish robot interaction behaviours were observed on a dining table. Based on the observatory study, an actor-critic algorithm was developed that enabled the robotic dish to adapt and then move on a dining table according to humans knocking patterns. Twenty participants in a between-subject design study evaluated two different knocking behaviour adaption of the dish robot. Results showed that the participants succeeded in establishing a successful communication protocol with the robot. However, the significant difference between the number of agreements and disagreements on knocking behaviour adaption between a robot and human were found.

Ref.	**Robot**	**Study Design**	**Robot Capabilities**	**Adaptive Features**
Torrey et al., 2006	Pearl anthropomorphic: Yes	Subjects: 49 adults No. of Interactions: One-off Interaction Type: autonomous Measures: Information exchange and social relations. Method: Questionnaire and Video analysis	utters context aware speech	Dialogue-based adaptations
Torrey et al., 2007	Pearl anthropomorphic: Yes	Subjects: 66 adults Conditions: Question Only, Gaze Added, Delay Added, Immediate Added No. of Interactions: one-off Interaction Type: autonomous Measures: Performance, communication and subjective evaluation Method: Video analysis, Questionnaires and Interviews	utters context-aware speech with gaze movement	Dialogue-based Adaptations

Cooney et al,. 2014	SPONGE anthropomorphic: No	Subjects: 20 adults (within-subject) Conditions: (Nave design vs. proposed design) and (reward, vs. suggestion) No. of Interactions: One-off Interaction Type: autonomous Measures: total interacting users, total printed coupons, and interaction initiation. Method: video analysis	provides rewards and suggestions based on understanding human gestures	Dialogue Adaptation Based on Human Gestures
Gross et al., 2015	CompanionAble 'Max' anthropomorphic: Yes	Subjects: 9 elderly No. of Interactions: trial for three days Interaction Type: semi-autonomous Measures: technology acceptance Method: Interviews	display emotions, recognise, detect and track person, give recommendations, understand haptic input	User-preference and emotion based adaptations
Youssef et al., 2015	DISH anthropomorphic: No	Subjects: 3 groups (18 graduate students, 6 children with typical hearing, 18 hearing impaired children) No. of Interactions: One-off Interaction Type: semi-autonomous Measures: learning performance Method: Video analysis	understanding knocking behaviour and moving on the table.	User-knocking behaviour based adaptations

Table 2.4: Summary of Adaptive Interaction studies in Home.

2.3.4.1 Discussion

The summary of adaptive interaction studies in domestic settings as shown in table 2.4 shows that research on the applications of AdSoR is at a preliminary stage and we need to conduct more research on different ways of implementations for AdSoR that can be utilised in various ways at home. Anthropomorphism is an important issue as we found both animate robots being referred (in different sizes). We speculate that the task that needs to be accomplished will define the look of a robot. A robot as a caregiver can be envisioned to be bigger in size as it would be required to carry objects at home (Gross et al., 2015). Similarly, Children might imagine robot size to be small as they would want to use it in different playful interactions (Obaid et al., 2015).

In general, adaptation during HRI at home is beneficial as it has resulted in positive findings in terms of user's attitudes. However, we need to be careful while choosing the type of adaptation for different tasks and types of the users. As it was identified in one of the studies that the gaze-based adaptation didn't yield in benefits (Torrey et al., 2007). However, more research needs to be conducted that evaluates robots possessing different adaptive capabilities at home to consolidate these results. In addition, the type of adaptation can also depend on the area to be used in the house, for instance, an adaptation mechanism for the dinning table would be different from the one in the living room (McColl & Nejat, 2013). Similarly, a robotic chef at homes in the kitchen can also have various ways of adaptation. All these aspects are missing and needs to researched.

Most of the studies have reported results based on one-off interactions and we believe these results might be effected in case of a long-term interaction. We believe that the reported benefits of adaptation on personality adaptation, user-mood adaptation needs to be validated during long-term interaction and also for different kinds of task and with varied user expertise. A task should define the type of adaptation as suggested from the results a non-verbal behaviour adaptation can or cannot work for the type of the task. Similarly, the effect of age of the participant should also reflect on the type of adaptation.

2.4 Adaptivity in Virtual Agents

Adaptive virtual agents (AVA) can be termed as decision-making engines that are capable of perceiving the information from the environment. Based on this information, they make decisions. An example of AVAs can be found in commercial games where they tend to increase players involvement through the use of an algorithm that enables the agent to adapt according to the user game playing patterns. For instance: Adcock & Van Eck (2012) rationalize that games possessing adaptive features could provide effective learning experience for the users. This rationale was based on the following learning theories including play theory (Rieber, 1996), problem solving (D. Janssen, 1997) and intrinsic motivation (Malone & Lepper, 1987). Researchers have also utilized these theories to build effective applications that enables virtual agents to take adaptive roles to enhance users interest and engagement.

Franklin (2013) presented an adaptive interaction design for teaching mathematics using a game with a semi autonomous virtual agent. He designed his agent that was able to manage emotions during the learning process through human game interaction. The design includes four different goals, seeking interest through immediate feedback from the agent, reducing fear through providing an increase to the opportunities for success during the game play, reducing panic and anxiety through providing a collaborative learning environment and lastly reducing rage through increasing play and care.

Sampayo-Vargas et al. (2013) designed and evaluated a game with both adaptive and non-adaptive features to help school student learn about Spanish cognates. The adaptive feature in his game includes adjusting the level of difficulty on the basis of users performance. He reported a study with more than 200 school going students who played the game to learn Spanish cognates. The results showed that the performance was better in case where the game had embedded adaptive feature as compared to the simple non-adaptive version of the game. Other studies have also shown that educational computational games have showed comparatively better performance in terms of learning and motivation (Klinkenberg, Straatemeier, & Van der Maas, 2011; WILSON & Revkin, n.d.).

Leemkuil & De Jong (2012) also presented an idea of a knowledge based computer simulation game, which used advice about next steps to complete the learning process from the agent during the game play. The study was also based on the comparison between the game play with an advice giving feature and no advice feature. The results showed that there was no performance differences found with the adaptive advice feature. The possible reason for not finding the performance difference was because students were too much dependent on the advice rather than doing the task themselves. We can safely say that the performance may depend on the variability of the adaptive behaviour.

In general, past research provided an evidence that the inclusion of adaptability in a game does lead to increase of interest and performance in most cases (Sampayo-Vargas et al., 2013). It has also been shown that due to the inclusion of adaptive advice, individuals tend to loose the plot in a game based scenario (Leemkuil & De Jong, 2012). The use of controlled adaptability during its implementation in a game is one of the research gaps such that the participants are not misled or deceived during the game play. The use of an adaptive virtual agent to teach mathematics through a game (Franklin, 2013) does give us indication of increase of interest and performance.

2.5 Summary and Conclusion

In this chapter, we presented an in-depth literature review on the adaptive interactions reported in the field of HRI. The purpose of the chapter was to identify the gaps with respect to the developments of adaptive robots in the field of HRI. The studies analysed showed that a significant amount of research is being conducted on adaptive social robots and they have been utilised in different social domains. Overall the results of these studies have reported positive findings in terms of user attitudinal preferences. However, it also reflected towards addressing a number of area such as low level of autonomy in robots, non-emotion based user profiling, and implementing personalisation in robots. Most of these studies are based on short-term interactions, therefore, longitudinal type research needs to be conducted to consolidate different previous findings. Additionally, our review suggested a need for designing novel ways of adaptations in robots in different social domains. Furthermore, the type of adaptation also varies depending on the social domain. For instance; the role of a robot adapting based on user personality may be regarded higher in education domain as compared to in the public space. In essence, the requirements of adaptation in robots may vary based on the environment.

In general, adaptation is highly desired in HRI and we also find evidence of the positive effects of incorporating adaptivity in the non-HRI context on user's attitudes and task performance across different application domains. Therefore, we need to implement and evaluate various AdSoR in various domains based on user characteristics, emotions, and personality of the user. In particular, we need to implement robot's adaptation mechanism that focuses on creating robot's memory based on more sophisticated means.

CHAPTER III

Understanding behaviours and roles for social and adaptive robots in education: Teacher's and Children's views

The goal of this thesis is to inform a mechanism for the social robot to perform various user-specific adaptations in an educational context. We believe that implementing such a mechanism in social robots would result in sustaining social engagement and promoting children's learning during a long-term HRI. It is, therefore, significant to understand the perspective of both teachers and children on various kinds of adaptation by the social robot in the Education domain. Consequently, we report results from two studies conducted with teachers and children to understand the acceptability of different adaptable social behaviours and roles displayed by the social robot to humans in different domains and environments [1].

[1] Two peer-reviewed conference papers have been resulted from this chapter
Ahmad, M. I., Mubin, O., Orlando, J. (2016, October). Understanding behaviours and roles for social and adaptive robots in education: teacher's perspective. In Proceedings of the Fourth International Conference on Human Agent Interaction (pp. 297-304). ACM.
Ahmad, M. I., Mubin, O., Orlando, J. (2016, November). Children views' on social robot's adaptations in education. In Proceedings of the 28th Australian Conference on Computer-Human Interaction (pp. 145-149). ACM.

3.1 Introduction

Researchers in the past have conducted a number of interview-based studies with end users to understand about different behaviours for the robot in different social domains. For instance; Dautenhahn et al. (2005) and Mahani & Eklundh (2009) conducted interviews with potential end users to derive social roles. Similarly, in the education domain, researchers have also emphasised the benefits of involving students and teachers in the design process. For instance: Breazeal (2009) also described that the teachers play an integral role in how children learn using aids, devices in class and perceive information. Even though devices can function independently, the teacher must direct the flow of the session/curriculum and the way it supports student learning. Therefore, HRI researchers have given importance to understanding the view of teachers on understanding the possible interactions for robots.

In the past, a number of studies have been undertaken to understand the impact of educational robots on users and stakeholders. Serholt et al. (2014) conducted interviews with 5 female teachers from four different European countries on their views on the use of emphatic robotic tutors in the classrooms. Results showed that the teachers considered the robot to be a disruptive technology that would result in spurious behaviour management of students. Similarly, Serholt & Barendregt (2014) interviewed students to identify information on their attitudes towards the possible future of social robots in education. We, however, argue that the data collected in the previous studies may have been effected due to the hypothetical knowledge of participants on the use of robots. Most recently, Kory Westlund et al. (2016) also questioned about the hypothetical knowledge of teachers in one of the studies dealing with interviewing teachers (Serholt et al., 2014). Kory & Breazeal and colleagues conducted a long-term study with a Tega robot for two months. They deployed an autonomous Tega robot to interact with children in the classroom during the school time. Their results based on interviewing teachers didn't find robots to be a disruptive technology. Similarly, Obaid et al. (2015) conducted an exploratory study on children's contributions to the design of a robotic teaching assistant. Obaid et al. compared the design samples of both children and interaction design students and showed that robot designs are influenced by individuals knowledge about robots. Therefore, it can be inferred that the hypothetical knowledge of the end users may effect the results. Another limitation of the previous studies lies in not giving individuals a long-term interaction experience with a robot.

Keeping the aforementioned limitations in mind, we believe that the next steps are to conduct studies where children and teachers first participate in a long-term interaction with a social robot with a view to inform researchers about different appropriate social behaviours adaptations for an AdSoR in education or in classrooms. Our contribution is in the longitudinal aspect, we argue that the knowledge of students in previous studies is truly hypothetical and usually, students have not interacted with a social robot.

In this chapter, we present two studies conducted with teachers and children from different schools in an urban city. The focus of our study is to understand the views of both teachers and children on the most appropriate and effective robot's adaptation in education. In order to avoid the influence of prior knowledge of teachers and children on results, the teachers interacted with a humanoid robot and they were informed about all of its possible and probable capabilities whereas children interacted with the humanoid robot in a long-term interaction. As discussed by Huber et al. (2014), we also believe in order to implement adaptive robots to sustain long-term engagement, it is very important to involve teachers to guide and help us design appropriate and effective behaviours for robots. In addition, it is also indicated in literature (Kennedy et al., 2015) that the robot who over tried through giving feedback at inappropriate time negatively influenced childs learning. Therefore, we need to be careful and should make informed decisions before implementing behaviour adaption strategies and features for social robots who are going to interact in a way children will be accepting of.

3.2 Method - Study with Teachers

The entire study and associated protocol was approved by the host university's ethics office (approval number H11429).

3.2.1 Participant

We interviewed teachers from both primary and high schools in an urban city, where the first language of communication was English. All of these teachers had training and experience in teaching languages including French, Italian and English. We selected language teachers because we intend to implement an adaptive social robot that can support children's language learning and development. Eight female teachers participated in the study. Out of eight teachers, we had four teachers each from both primary and high school. The age-range of teachers was between 25 and 45. None of the teachers had prior experience interacting with the social robot. The interview sessions were held at the corresponding schools and lasted for 30 minutes each. All of the interviews followed the same procedure in both schools.

3.2.2 Procedure

We began our study by distributing information sheets and consent forms to the teachers. Once teachers have completed and returned their consent forms, we gave them an introduction about the scope of the project and answered any questions.

The study was later conducted with teachers during one-to-one interactions. Each study session lasted for 30 minutes and had three different steps which are detailed further below: 1) a 5-minute video (https://www.youtube.com/watch?v=ho9i1moUJos) on HRI and SR followed by a 3-minute video (https://www.youtube

.com/watch?v=2STTNYNF4lk) on NAO, 2) a 5-minute of interaction between teachers and NAO, 3) Interviews with teachers on how NAO can contribute towards language learning through adapting to children behaviours. The data collection was completed in two days. On the first day, we visited high school followed by our visit to the a primary school on the second day.

Firstly, we showed teachers an introductory video on SR and HRI followed by another video on NAO robot showing various capabilities such as gestures, object recognition, tactile sensors and speech recognition. The purpose of showing these videos was to avoid an effect of previous and hypothetical knowledge about robots and also to overcome any fears that robots can takeover their jobs. We later asked teachers about possible confusions or questions on NAO and HRI in general. Secondly, we asked teachers to interact with NAO in three different scenarios. In Scenario 1, teachers interacted with NAO to get familiar with its Speech Recognition capabilities. Teachers were asked to speak with NAO by asking a basic set of introductory questions such as: (*What is your name?, How are you?, How is your day progressing? etc...*). When asked an unknown question, NAO repeated their question. In Scenario 2, teachers were able to interact with NAO by showing a range of emotions using different facial expressions *(happy, sad, angry and neutral)*. The aim of this scenario was to understand about emotion detection capabilities. Teachers were asked to show emotions to NAO, after detecting these emotions, NAO named the emotion. In Scenario 3, NAO showed different gestures (clapping, waving, bowing and dancing). Teachers were asked to speak the gesture name to NAO and in response, NAO displayed the gesture. Lastly, we interviewed teachers in order to understand current practices followed by them at schools for language learning and how can NAO contribute towards enhancing child's learning. In addition, we asked teachers about their opinion on different adaptive behaviours and roles that a robot can display to influence child learning and long-term engagement in a language learning scenario.

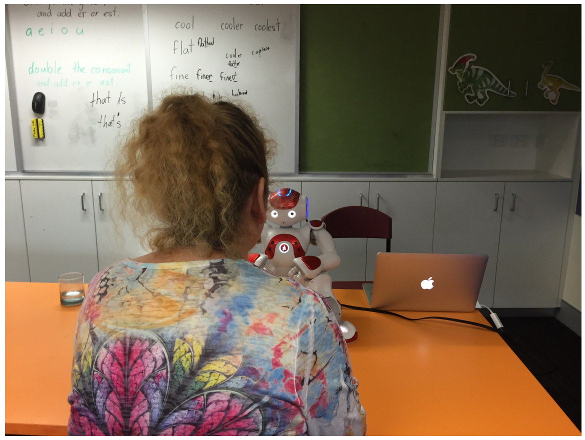

Figure 3.1: Setup: A teacher interacting with NAO.

3.2.3 Setup and Materials

We used NAO robot designed and developed by Aldebaran robotics. It is a humanoid robot measuring 58 cm in height. NAO is an interactive and adaptable robot partner. It provides researchers with a platform to design various applications driven by their creativity and requirements.

We conducted our study in a quiet room at both primary and high schools during school time. We were provided with a small table and several chairs inside an empty classroom. The NAO robot was placed in a sitting position in front of the participant teacher in order to get a the clear view of the participant's face for detecting emotions. The researcher sat in front of the participant and was involved in showing videos on the notebook and conducting interviews. The NAO robot autonomously generated behaviours for three different scenarios: *Speech recognition, emotions detection and displaying gestures*. The setup is shown in figure 3.1. We took consent from the teacher to use their picture in our thesis.

3.2.3.1 NAO Robot and implementation of Scenarios

The NAO robot was autonomously programmed to display three different capabilities: Speech Recognition, detecting emotions and displaying gestures.

We used *Google* speech to text API (Zhang, 2016) to convert teacher's speech to text. The NAO robot was programmed to respond according to basic questions of the teachers. The basic responses ranged from introducing itself to greeting each other. It was also able to repeat sentences of the teachers.

We also trained data on basic human emotions (Happy, sad, angry and neutral). We programmed NAO to capture user facial expressions and detect their emotions through using an algorithm (EVP, 2015). The algorithm uses openCV library to localise the mouth area to detect emotion of a user. The image captured by NAO is re-sized to 28*10 pixel containing only person's mouth and surrounding areas. The image is then converted into grey-scale and flattened into a vector of length 280. A logistical regression programme then takes the vector and determines the emotional state of the user.

We implemented state of the art existing behaviours in NAO robot in *Choregraphe*. The gestures included bowing, clapping, touching, whipping, hugging and dancing. The Choregraphe programme was later used to generate python code.

3.2.3.2 Interview Questionnaire

The questionnaires involved understanding different language learning strategies used by the teachers with their students and how NAO can adopt these strategies and contribute towards further supporting this learning. In addition, questions focused on what novel adaptive behaviours and roles NAO should play to prolong children robot engagement? Importantly, these questions generated data which gave direction to implementing novel scenario and behaviours for future robots in an effective way to support children's language learning. It was also important to enable teachers to interact with one of the Humaniod robots (NAO) in order to develop their understanding on the robot's current capabilities. The questionnaire was designed according to the existing characteristics of a social robot (Fong et al., 2003). Other questions were used from (Serholt & Barendregt, 2014). The list of other questions from our interviews are:

1. *What teaching approaches, resources and technologies do you use to teach oral and written language development?*
2. *How do you think a robot can contribute towards efficient language learning?*
3. *How do you want a robot to show different gestures during a one to one interaction?*
4. *How do you want a robot to display a personality according to a child?*
5. *How do you want a robot to react to children emotions?*
6. *What kind of role a robot should play to improve learning?*
7. *How do you want a robot to store child's memory?*
8. *How to ensure long-term engagement between NAO and a child?*

3.2.4 Data Analysis

All of the interviews were audio recorded after receiving consent from the teachers. We performed content analysis to analyse interviews communication (Downe-Wamboldt, 1992). One of the researchers listened to and transcribed the interviews. The author then noted possible themes and patterns in teachers responses. These patterns were later used to define main themes and sub-themes for this paper. Quotes were selected from the relevant responses.

3.3 Method - Study with Children

The entire study and associated protocol was approved by the host university's ethics office (approval number H11429).

In our study with children, we followed the procedure as described by Serholt & Barendregt (2014). However, we conducted a long-term cHRI at the school before conducting focus group activity with children to understand their views on the different adaptations for social robots in education.

3.3.1 Participants

We requested school to provide us with participants from Years 5 and 6 (ages 10-12 years). In total, 12 students participated in our study: 6 were in Year 5 and 6 were in Year 6 and there were 4 boys and 8 girls. None of the participants had experience working with NAO robot. We chose 10-12 year old children because at this age they are able to develop ideas from abstract concepts but are still open to exploring new ideas (Druin et al., 1998).

3.3.2 Procedure

Our study was conducted in two steps: 1) a long-term interaction between a child and NAO robot, and 2) a focus group activity.

Long-Term Child-Robot Interaction: All of the participating children played snakes and ladders board game for 3 times over a span of 9 days with NAO robot. Each child played the game on first, fifth, and ninth day respectively. Each time, NAO asked the child to make the first move after an introductory dialogue. The child played the game with the robot for a maximum of 10 minutes or if there was winner, whichever came first. If there was no winner, NAO ended the game and called it a draw.

Focus Group Activity: On the tenth day, the students were divided into 4 groups; each comprising of 3 children.The duration of focus group discussions was recorded to be between 20 and 30 minutes on average. We asked the groups to discuss following themes: 1) NAO as a teaching assistant, 2) Interactions with NAO, 3) NAO displaying/understanding emotions, 4) NAO interacting with gestures, 5) NAO displaying a personality?, 6) NAO remembering/recalling events?, and 7) NAO making/reacting to mistakes. During the discussions, children were enabled to inform about their opinions on afore-mentioned themes and how do they think a NAO robot can be helpful for various educational purposes. All of the focus group activities were audio recorded.

3.3.3 Interaction Scenario

We programmed NAO robot to autonomously play the snakes and ladders game with children through taking turns, however, speech recognition was controlled via Wizard of Oz (WoZ). The snakes and ladders game was programmed autonomously to role the dice for NAO after the child's turn. NAO was capable of autonomously selecting emphatic behaviours based on the game and user's states. The behaviours consisted text-to-speech, displaying gestures, recognising and reacting to user's emotions and expressions, and keeping track of user's performance. The text-to-speech and gesture-based interactions were chosen based on the game state during the gameplay. The NAO robot was programmed to recognise and react to user's emotions based on facial expression after every 30 secs. In the subsequent interactions, NAO reacted to existing game performance through recalling previous game events such as snake near 100, or an early ladder.

NAO robot began the one-to-one interaction (WoZ controlled) with each child through asking introductory questions: (Hello, I am NAO, What is your name?, "Child's Name", Nice Name, How are you today?, Today we are going to play snakes and ladders game, you can go first). Once the game was started, the robot autonomously generated appropriate phrases based on the dice outcome, snake and ladder within the game for the child. If the child was performing exceptionally well, for instance; moving up a ladder, then the robot displayed a range of praising gestures such as high five, thumbs up, and/or clapping. Upon winning or losing, the robot also congratulated or wished the child all the best for the next time through combining speech with a range of gestures such as bow down, showing sadness, show surprise, or joy respectively. NAO was also programmed to deduce information about childs affective state such as: detecting childs emotions based on the facial expressions. We also used an online API developed in python (Indico, 2016) that enabled us to determine an emotion expressed in an image on a human face. NAO gave a verbal feedback based on childs emotional state and position within the game. For instance; if the child was at a lower number then the robot and the child facial expressions were detected as sad, NAO said, “Dont loose hope, anything can happen within the game”. We also implemented a memory adaptation mechanism. On the first day, the robot kept track of the game performance of other participants. Starting from the second session with each child, the robot kept track of their names, how they and their friends performed (snake/ladder/star during the last game-play, no. of moves did they take to win the game) during the last game sessions. For instance, if they had a snake near 100 during last game session, the robot said, “I remember, you also faced a snake near 100 last time, but don’t worry, you can still win the game”.

3.3.4 Setup and Materials

We used the NAO robot designed and developed by Aldebaran robotics. The setup of the study required using two spaces. The first space was for children to engage with the NAO by playing snakes and ladders game. We conducted our study in a quiet room (school library), and as shown on the left side of Figure 3.2, below. this section of the room was divided into two portions. On one side, the child independently interacted with NAO that had been placed ,along with a tablet device, on a seat in front for the child. On the other side, one of the researchers was controlling the speech recognition capabilities of NAO that the child was engaging with. Secondly, we were provided with a classroom with a table and chairs as shown in Figure 3.2, on the right, where the researcher conducted a focus group activity with the same children. All of the focus group activities were audio recorded.

Figure 3.2: Setup: A child interacting with NAO (left) and the focus group (right).

3.4 Results

3.4.1 Teacher's View

We present the main themes resulting from the qualitative analysis performed on the interviewed data. First, we present teachers views on how NAO robot can contribute towards child's language learning and development. Second, we present about the behavioural and role adaptation for NAO according to teachers in Language language. We also present sub-themes with respect to adaptation in teacher's perspective. Thirdly, we present teacher's opinion on how NAO can maintain long-term engagement. To keep the identities anonymous, primary school teachers are labelled with the range of 1 to 8, where P1, P2, P3, P4 are from the primary school while P5, P6, P7, and P8 are from high school. Table 3.1 shows the languages taught by the corresponding participants.

Language	Teacher
French	P7,P4
Italian	P5,P6,P3,P8
English	P1,P2

Table 3.1: Teachers and Corresponding Languages

3.4.1.1 Robots and its Contributions towards language learning

Teachers from both primary and high schools pointed out that the robots can help with vocabulary learning, grammar, and correcting pronunciation. They can also play games with children on language learning. The teachers also wanted children to practise their speaking skills with the robot. Some of the teachers from both schools reported:

We can possibly use a robot for word learning, comprehension and games based on questions and answers to motivate students to learn languages (P4).

The robot can speak a grammatically correct sentence and a child can repeat it (P2).

I think, the girls and the student will find it quite comfortable to interact with, they can practise their speaking skills (P6).

Teachers from the high school emphasised the benefits of programming robots that can help them with marking both objective (multiple choice questions) and subjective (essays, reports) assessments. On the other hand, teachers from both schools referred to the benefit of having robots to fulfill drill types of tasks. For example, they stated that students usually keep asking same questions on their subject in the class. It can sometimes get frustrating for them but as the robot can repeat it as many times so it can be a great support in this regard.

Students normally ask same questions repeatedly, the robot can answer these questions in the classroom while I am working with other students (P8).

In addition, the teachers from high school commented that it would be great if a robot can autonomously find learning material for them. In general, teachers from high school wanted robot to help with finding material for their class, help with marking student's assignments, and should also be able to repeatedly explain a certain concept to the students.

3.4.1.2 Robots with Gesture Adaptation

Teachers from both schools recognised the use of gestures as one of the essential parts of teaching languages. Robot depicting several appropriate gestures to explain a concept or to give feedback can be helpful and may lead towards efficient learning and engagement.

I am quite animated during my class. Gestures such as pointing, indicating, arm moments, bending, hands on heads works best with juniors (P5).

Most of the teachers also mentioned that they use a number of effective gestures to engage students in learning. The gestures included: *pointing, arm moments, hands on heads, greetings, hands up, hands down, use of finger to point the wrong answer, thumbs up, attention, listen please, finger going to the eye, saluting, throwing, dancing, and smiling.*

The teachers from high school in particular mentioned that different gestures can be associated with different languages. They indicated that the gesture for greeting in the French language is different from other languages. The robot can have a different teaching style based on the language. One teacher from high school (P7) was of the view that gestures can be culturally driven so a robot needs to adapt accordingly. Cultural adaptation can also motivate child towards efficient learning. She reported:

We have a different way of saying good morning in french. We teach students regarding individual situations while teaching languages. The Robot can combine both gesture and speech capabilities to improve child's learning (P7).

3.4.1.3 Robots with Memory Adaptation

Teachers from both primary and high school advocated memory adaptation as they stated in language learning recalling or remembering an experience can be helpful and can result in a positive effect on the child's learning. They also mentioned that memory can be used as a great tool to show extra care towards a child and a child can be amused as well as motivated towards learning. One of the teachers from primary school (P3) reported:

We can show extra care through recalling previous interactions. If this can be made possible, children will not do the same mistake over and over again. Children will be forced to think about new ways of adaptation (P3).

One of the teachers from high school (P6) also reported that memory can be a great tool to test student's performance. She reported:

Robot with child's memory can be a great idea. If a student was able to say a certain word after two weeks, the robot can later ask the same question on the next days. It can, therefore, judge the child and also encourage through providing positive reinforcement (P6).

3.4.1.4 Robots with Emotion Adaptation

All of the teachers from both schools acknowledged the importance of recognising emotions through facial expressions. However, they were worried that facial expressions are not the only way to recognise emotions because they are not the only identifier of a child's mood or emotional state. Teachers particularly from primary school focused on the importance of detecting the emotional state of the child through dialogue. They proposed a set of conversations that a robot should have in order to realise on the emotional state of the child.

The use of facial expression to detect emotion is important, However, we need the robot to sense mood through dialogue. For example, if a child is bored, then it can *provide positive reinforcement or if the robot can detect anger, it can play a calming role* (P3).

One of the teachers from primary school was of the view that in case, a child is not in a good mood on a particular day due to some unavoidable experience, then the robot should be able to judge it and perform certain steps to overcome or change the emotional state of the child.

The dialogue-based mood detection is more important than facial expression. The robot should react non-judgmental and should also try to get them through a list of procedures and calm them accordingly (P2).

Teachers from primary school also informed that robots having feelings will improve engagement as children will not only consider robot as the toy, however, one of them (P2) was concerned that extensive emotional adaptation might intimidate children of a certain age. Although, teachers from high school didn't mention any of such concerns.

Robots with emotions makes them real and makes the student realise that robot has feelings and it is a real interaction. However, robot's understanding emotion can be intimidating depending on the age of the student (P2).

3.4.1.5 Adaptive Social Roles for Robots

Teachers from both schools pointed various roles for the robot that it can display and adapt during children robot interaction. Table 3.2 shows a taxonomy of the frequency of various roles that in view of all the teachers can be appropriate or misleading for children. We coded these frequencies from teacher's statement.

In general, teachers informed that robots need to adapt different role for the different scenario. Most of them said that robot needs to encourage and motivate children through persuasive dialogue. They also said that the robot needs to find a balance between a submissive and dominant role and should rather play a democratic role.

One of the teachers, however, was also of the opinion that dominance is important in order to be taken seriously.

If a student is hesitant to do something, then Robot should be able to persuade him/her through positive reinforcement (P5).

Student would like the idea that the robot adapts and play a buddy role in the classroom. It can be a motivator for children to learn languages (P7).

Robot displaying a dominant role can be overwhelming for the child (P2).

All of the teachers pointed attention towards the use of a robot with a good sense of humour as learning a language can get boring at times. The robot needs to play funny roles in order to motivate children and keep them engaged during the learning process.

Robot needs to act funny and tell jokes to students, for example, when explaining the meaning of a word in a vocabulary learning exercise. This will keep the child engaged and motivated throughout the process.

All teachers considered calm and compassionate roles to be best suited for the case when the child is making repetitive mistakes. The robot needs to first recognise the repeated mistake and adapt through changing the dialogue or its tone in real time.

You need to play a calming or a soothing role through the use of dialogue and gesture in order to motivate and help the child improve learning (P8).

3.4.1.6 Robots with Personality Adaptation

Teachers from both primary and high schools were of the opinion that it is not necessary to adapt according to the general definition of personality (extrovert or introvert) of the child. They commented that child's personality might vary on the day according to the set of events that might have happened during the school time. The robot, on the other hand, should be able to recognise child's personality through dialogue. Teachers reported as:

Roles	Participants in favour
Assertive	4 (P1, P4, P7, P8)
Funny	8
Encouraging	8
Persuasive or Motivating	8
Bribing	1 (P2)
Dominant	1 (P5)
Submissive	1 (P6)
Diplomatic	5 (P1, P3, P5, P6, P7)
Cooperative	8
Assistant or a Helper	8
Buddy or a Friend	8
Competitive	8
Calm and Compassionate	8

Table 3.2: Roles for a Robot in Language Learning

Personality adaptation through dialogue can be an efficient way, in a personalised interaction, personality can be a handful, for example, the robot needs to show patience in case of an introvert child (P5).

We need to detect the personality of the student, the personality can change over time so we need a mechanism to detect the personality. We can detect that through dialogue, for instance, we can detect hesitation then encourage the student through rewards or positive gestures (P6).

One teacher from high school also directed towards finding the balance between an extrovert and introvert personality. She reported:

A balance between an extrovert and introvert is required. Perhaps a more outgoing one as it would probably encourage more interaction (P7).

3.4.1.7 Robots for Long-term Engagement

All Teachers expressed concerns about maintaining engagement between children and the robot during a long-term setup. They were certain that in order to achieve sustainable engagement, they should be equally involved in preparing robot interactions with the children. They were afraid that if teachers are not given control of the robot, it could very well happen that the robot stays alone at the corner in a classroom. One teacher from high school mentioned:

If the teachers know how to control and use the robot, then we would be able to achieve long-term engagement (P5).

Another teacher from high school pointed out that teacher should be provided with an interface that enables them to update or change lessons over the period of time.

Teachers can keep the robot involved, so that it not just at the corner and always in the loop, we need to use the robot as a part of a lesson (P6).

One teacher from the primary school also mentioned that if there are a variety of programs and the systems updates itself and also adapts to a child's need, it would be possible to achieve long-term engagement. However, none of the other teachers expressed such opinion.

You should have a variety of programs so that the robot is not doing the same thing over and over again. The robot should mimic what a human would do at a certain moment (P2).

3.4.2 Children's View

In order to analyse data taken from focus groups performed after a long-term interaction between the child and the NAO robot, we performed content analysis on the transcriptions generated from audio recordings. We present our results generated from the qualitative analysis performed on the themes being discussed during focus group activity. First, we present children's views on how NAO robot should act in a classroom as a teacher or teaching assistant. Second, we present views on the role of various capabilities (Gestures, Emotions, Memory) that NAO should perform and that can keep children engaged during long-term interactions. To keep the identities anonymous, the groups are labelled as G1, G2, G3, and G4 respectively.

3.4.2.1 The Role of Robots as teaching assistants or teachers:

In general, the response of children groups on using robots as teachers/teaching assistants was positive. Children wanted robots to read books, talk about grammar, demonstrate science experiments and show them how to build new things. However, a group of children feared that *the robot can malfunction such as having a virus, therefore, it should not be used in classrooms (G3).*

Children also discussed about correcting mistakes during a learning interaction. Two groups of children didn't want NAO to reveal the answers in case a child is not unable to understand a concept, but give tips to reach an answer. On discussion about feeling scared around NAO, Children, in general, mentioned that they will only feel scared if, *NAO robot performs unexpected actions and/or moves/stand up quickly and/or gives fast reactions (G4, G2)."*

3.4.2.2 Gesture Based Adaptations:

Upon discussing the role of gestures during learning interactions, children reacted positively towards NAO adapting gesture based animations on task outcome or while explaining concepts. All of the groups of children considered the use of hand gestures as one of the most important aspects of teaching. One of the groups reported: *It can use his hands to depict emotions such as hands on eyes can display sadness, hands opening near the lips will show happiness, blinking eyes with blue colours can also depict sadness (G1).*

3.4.2.3 Memory Based Adaptations:

All groups of children appreciated a robot with memory capability. They expected robots to keep track of previous interactions with them. 3 out of 4 groups expected robots to recall memory events to build an association with them. A group mentioned: *NAO should remember previous communication. It is very important to have memory to create a social relationship (G4).* One child in a group mentioned that in order to keep the interaction natural, the robot should make mistakes while recalling events. They were of the view that robots should also forget things and be corrected. One of the children from one of the groups countered the argument saying that it is not expected from a machine to forget an event. One of the groups said: *The robot should not remember everything and it should act natural and make mistakes. (G2)* Another group said: *Robot should not make mistakes because it is not expected to do that, it is like a SIRI/Google Voice with a body (G3).*

3.4.2.4 Emotions Based Adaptations:

Children wanted the robot to not only understand their emotions but also inform them about its own feelings. All children in one of the groups reported: *He can change his eye colour, and he should tell me how is it feeling from inside and I can tell him how am I feeling (G1).* Children also mentioned that robot should communicate with children to understand their emotions. It can vary based on the time of the day. It needs to have a dialogue mechanism. Children also wanted the robot to show emotions other than happy, sad, angry, or neutral. They expected the robot to act worried or confident based on a certain situation. The group reported as: *if the robot is acting confident, it will give me confidence (G3).* In addition, they also didn't want the robot to act sad, but only when it finds out a child is hurt. They reported: *It should be happy most of the times, it should be gentle with us, it should be sad only when someone gets hurt.*

3.4.2.5 Personality based Adaptations:

Children reacted positively towards the concept of personality adaptations by a robot but had different views on adaptation methods. Children in one group (G3) wanted the robot to adapt according to their personality while another group of children (G2) reported that robots should adapt opposite to their personality so as to enhance to their disposition. One group reported: *NAO should try to encourage or motivate the child to a certain limit to engage, in case, the child has a shy/introvert personality (G4).* They also mentioned that the personality can vary, therefore, the robot should have a dialogue mechanism to first understand the personality and then adapt accordingly. The other reported: *NAO should be able to recognise which personality is on right now If I am awkward I would not want to be social. (G3)* Another reported that: *NAO should try to boost my confidence when I am shy. (G2).* They also said robot should adapt according to their strengths and weaknesses. They commented: *robot needs to know what I like and dont like and then adapt to it accordingly (G4).*

3.4.2.6 Voice based Adaptations:

Two groups (G2,G3) of children reported that they were unable to recognise the gender of NAO robot and were also of the view that NAO speaks really fast. Children in these groups also reported that in the classrooms, the robot needs to speak slowly. They were also of the view that robot needs to adapt its voice according to the given situation. For instance, if the robot is playing a helpful role, he needs to sound helpful. The voice pitch should adapt to the situation. On the other hand, children in other groups didn't report these observations.

3.4.2.7 Robots for Long-term Engagement and Sustainability:

Children reported that NAO needs to continuously update its comments or dialogue. After three interactions with NAO, all of them wanted to have different type of interactions with the NAO robot. One children in a group reported: *yes I would want to play again with NAO. But he needs to have more comments and actions (G1).* All of the children in a group (G3) wanted the presence of teachers in classrooms as they should be able to control the robot for their long-term deployment. They reported: *We need teachers to control the robot. just imagine, if a child can hit the robot and run away, he cannot do the same with or in front of a teacher. (G3).* One group of children (G4) also pointed practical issues with robots that can scare them, they questioned what if the robot has a virus? This needs to be addressed and the presence of teachers is, therefore, mandatory.

3.5 Discussion

We have learnt a number of lessons and takeaways for HRI researchers from these studies. The key lessons we learned about implementing adaptive social robots in education, as reflected in the teacher's and children's feedback are as follow:

1. The robot should be able to answer repeatedly asked questions in the classroom: There is a need to implement a mechanism in which a robot can respond to repeatedly asked questions during one-to-one or group interactions with children. In order to achieve it, we need to implement ways for a robot to perform memory adaptations. Chang et al. (2010) have described various characteristics of robots that can be important during language learning. Repeatability was described as one of the most appropriate features for language education.

2. We need to design dialogue based adaptation mechanism in order to adapt to user emotions and personality in real-time: Human emotions and personality are correlated to each other. The current emotional state or mood of humans can influence the portrayal of personality . Therefore, as indicated by the teachers and children, it is significant to design and implement a real-time adaptation mechanism for a robot to detect user's personality based on its emotional state. Previous studies (Hayes & Riek, 2014; Mileounis, Cuijpers, & Barakova, 2015) have shown that if a robot can adapt according to the personality of the user, it can positively influence learning. However, most of the studies in which a robot is able to adapt users personality are conducted through asking participants to complete a questionnaire in order to detect their personality (extrovert or introvert). Therefore, it remains an open research question on how to detect personality in real-time. In addition, another possible strategy is to use a broad categorization of emotions that move beyond a simple extrovert and introvert identification. On the other hand, Mood or emotion adaptations based on dialogue have been studied in human-computer interaction (Pittermann et al., 2010). However, we find fewer work on dialogue-based emotion adaptation and its effect on user perception and performance for educational robots and HRI in general.

3. The selection of robot role's during children robot interaction in real-time can be based on memory adaptation: Most of the teachers emphasised performing real-time role adaptation. One of the suggested ways was through the use of memory during children robot interaction. In literature, we find a number of studies where robots have played different roles of a friend (Emmeche, 2014), competitor or cooperative (J.-H. Lee et al., 2015), and persuader (Chidambaram et al., 2012). All of these studies have shown positive results with respect to robot's effect on user perception, engagement or learning. However, there is a need for designing mechanism such that a robot can adapt its role in real time according to a given situation. One of the ways is the use of memory to perform role based adaptation. For instance: in case of repetitive mistakes on a given task, the role may choose to play a supportive role or may also criticize the user.

4. Robots keeping track of a child's memory can in-turn motivate children to improve learning performance: Teachers mentioned that children make mistakes regularly. The robot with a capability to show an adaptive behaviour through the use of recalling previously made mistakes can motivate children learning. As teachers mentioned that it can enable children to think new ways of adaptations in order to impress or outsmart the robot. Children also supported robots adaptations based on their memory Literature in HRI has also emphasised on the significance of memory adaptation. In a recent survey conducted on social robots for long-term engagement, Leite et al. (2013) highlighted that memory based adaptation as one of the unexplored areas in HRI and have conjectured that possessing a memory can make social robots more flexible and personalised to particular users. Recently, researchers have also shown that robots with memory can affect user performance as it enhances their likability and empathy (Hastie et al., 2016). Therefore, teachers opinion are also in line with the current research recommendations in HRI.

5. Culture-based adaptation can be significant during language learning tasks: Our results from teacher's view study highlighted the significance of gestures during language learning. They also mentioned that robot should adapt when teaching different languages because gestures are culturally driven for different languages. We find less work on cultural based adaptation in HRI. Studies have been conducted to show the significance of culture during HRI. Researchers have indicated the understanding of gestures is dependent on the culture and even within one culture, interpretations can differ for different situations (Zheng & Meng, 2012). In addition, it has also been described that different individuals from different cultures perceive personality of the robot differently (Isbister & Nass, 2000). Therefore, it is important to study the effect of cultural adaptations on user's engagement and performance during HRI.

6. Consider designing an easy to use interface for teachers to update new lessons for long-term engagement: Our results from both studies emphasised the importance of teacher's involvement in order to keep the robot engaged and involved during the learning process for long time. In order to keep teachers involved, it is important to design interfaces that allows them to easily manage robots. The robot needs to be adaptive, however, the content and curriculum need to be revised or changed after a certain period, which can be done by the teacher only if he/she has appropriate control. In literature, Chang et al. (2010) has described that the robot needs to be flexible enough to allow teachers to adjust and design appropriate robot-supported instructional activities for relevant teaching and learning requirements. In addition, Orlando (2014) has showed that teachers today have confidence with technology and also possess a diverse range of technology expertise. Therefore, research needs to be conducted on the development of such interfaces that are easy to use for teachers.

7. Robots need to perform voice adaptations based on different social situations: Our results showed that children wanted robots to adapt their voice tone according to a certain situation. In their opinion, it can lead towards natural cHRI. Recently, Lubold et al. (2016) has conducted a study with an undergraduate student to measure the effect of voice-adaptation and social dialogue by a robotic learning companion on user's perception. Results showed that a social voice-adaptive dialogue has a significant effect on social presence as compared to a simple social dialogue. Therefore, we also need to evaluate the effect of a robotic tutor that can adapt its voice pitch and tone during a social dialogue on children's learning and engagement.

8. We need to address malfunctioning or technical issues of robots for their long-term sustainability in education: We expect the robots to be deployed in various social domains, however, we find less research on addressing robot's sustainability. As children pointed that someone can hit the robot and run away, therefore, we need to implement mechanisms to ensure long-term deployment. One of the ways, as mentioned by children, is to actively involve teachers or test a robot that gives warning to a child in order to avoid children's negative attitudes towards a robot.

3.6 Conclusions and Limitations

In this chapter, we presented our results on teacher's opinion on how robots can contribute towards language learning with children through performing a series of adaptations. We also presented our results on children's opinion on how robots should perform different adaptations. Both teachers and children reacted positively towards robots adapting to emotions, memory, and personality. In addition, children pointed towards sustainability issues with a robot in classrooms and mentioned on the importance of teacher's role towards their long term deployment in education. Moreover, we showed that there is a need to implement easy to use interfaces for teachers to enable them to upload new content for the robot. This can lead towards long term engagement of robots in education.

We didn't conduct the study longitudinally so we were not able to overcome the biases of the teachers. However, during the study, we tried to overcome pre-existing biases through the use of videos and the actual presence of the robot. The teachers agreed to have understood the capabilities of the robot and confirmed that their responses won't be driven based on any biases or past experience. Another limitation to our work is that the teachers were able to interact with limited capabilities (e.g. the robot only responded to limited questions, only recognised basic emotions, and displayed few gestures). We also understand if teachers were able to interact with NAO that showed more capabilities, we would have received more classified information from the teacher. Another argued limitation of our study can be the total number of participants. However, one of the previous studies conducted with teachers on their views on robots in education also had the equal number of participants (Serholt et al., 2014). In addition, our participants were specifically language teachers therefore we can interpret our results with some reliability.

CHAPTER IV

Understanding the Effect of Different User-based Adaptation; A Long-term Study

In this chapter, [1] we focus on understanding the effects of different adaptations portrayed by the social robot that will sustain social engagement for an extended number of interactions. We report on a study conducted with three groups of children who played a snakes and ladders game with the NAO robot to understand the aforementioned effects. During the game, the NAO performed the following adaptations: 1) Game based adaptations, 2) Emotion based adaptations and 3) Memory-based adaptation. The rationale for choosing emotions and memory based adaptations was based on the observations highlighted by the teachers and childrens in Chapter III.

4.1 Introduction

We witness applications of social robots in various environment such as elderly care (Kachouie et al., 2014), domestic (Ma et al., 2014) and work environments (Leite et al., 2013). There has also been substantial growth in and growing interest in the applicability of robots in education. An extensive review on the acceptability and

[1] This Chapter has been published as a journal article
Ahmad, M. I., Mubin, O., Orlando, J. (2017). Adaptive social robot for sustaining social engagement during long-term childrenrobot interaction. International Journal of HumanComputer Interaction, 33(12), 943-962.

use of robots in education emphasised the need for adopting robotic behaviours and curricula to the user, and that further research about socially acceptable behaviours for the robot is needed (Mubin et al., 2013). There are currently a number of studies where researchers have evaluated various robotic systems that adapt according to users through displaying various socially accepted behaviours. For instance: Kwon et al. (2010) evaluated a robot used to help and assist korean children learn English languages and Komatsubara et al. (2014) explored the use of robots in classrooms to facilitate science learning. The importance of adaptability in education extends itself to beyond social robots. For instance: dos Santos & Osório (2004) presented an intelligent adaptive virtual agent that adapted according to user's preferences and interest and applied it to distance learning. Buche et al. (2003) applied an adaptive *MASCARET* model in order to create an intelligent tutoring system. Similarly Mitrović & Djordjević-Kajan (1995) also presented a machine learning approach for model student behaviour in an intelligent tutoring system.

Despite a remarkable amount of successes reported in literature, children tend to loose interest in interacting with the robot as time progresses (Komatsubara, Shiomi, Kanda, Ishiguro, & Hagita, 2014; Kanda, Hirano, Eaton, & Ishiguro, 2004; Jimenez, Yoshikawa, Furuhashi, & Kanoh, 2015; Coninx et al., 2016). In other words, children's social engagement gradually declines. The challenge of maintaining engagement has also been addressed in Children-Computer Interaction (cHCI) (Lyra et al., 2013). Similarly human-robot engagement has also been studied with elderly (Khosla et al., 2016). The decline of children's interest in, or engagement towards interacting with a robot, has been reported to occur after the 1st week of interaction Kanda et al. (2004), or from the third session onwards (Salter et al., 2004). According to C. L. Sidner et al. (2005), "Engagement is the process by which interactors start, maintain and end their perceived connection to each other during the interaction". Different factors have been used to measure user interest or engagement. These factors include verbal and non-verbal behaviours through analysing interaction videos (Dautenhahn & Werry, 2002), and the use of questionnaires to rate social engagement, perceived support and social presence (Leite et al., 2014).

Literature suggests that children's engagement can be sustained through implementing a robotic system that can adapt according to user states as specified in chapter II. The state of a user can be based on several factors including emotions, memory, or personality. Following these theoretical discussions, different methods have been applied by various researchers to address the decline of social engagement of students during Children-Robot Interaction (cHRI). Most recently, Leite et al. (2014) evaluated the role of supportive behaviours (providing advise) portrayed by the philips *iCAT* robot while playing a chess game with children aged between 8 to 9 years for five weeks. Each child played the game for five sessions. Their findings show that the inclusion of empathy (an ability to understand and share the feelings of another) during cHRI was able to sustain long-term children engagement. Empathy refers to the ability to understand and share the feelings of another. One of the short-comings of their study was the reliance on self reports yielded from children as they are challenging to administer with children (Druin, 2002; Pasch, 2010). In addition, young children aged between 9-12 years have an intrinsic tendency to please adults (King & Yuille, 1987). Another limitation was that the study didn't answer which specific emphatic behaviours resulted in sustaining children engagement as it did not have any conditions to compare different emphatic behaviours.

Researchers have also employed methods, including understanding and reacting according to user's affective states, in order to address the decline of social engagement during cHRI. For instance: Jimenez et al. (2015) presented an emotional expression model for the *Ifbot* robot to facilitate learning. The identification of complex emotions through recognising facial expressions is another known challenge and in response. Cuadrado et al. (2016) has presented an emotional model based on recognising emotional states from keyboard and mouse interactions. Memory Adaptations is another method, whereby a robot adapts according to previous events through remembering and recalling them. Memory based adaptation method, unfortunately, has received less attention. Recently, Leite (2013a) emphasised the need for implementing robotic models to implement memory adaptations. Baxter & Belpaeme (2014) have also predicted that the future of social HRI lies in the past. They have highlighted the significance of pervasive memory and its potential impact on the long-term HRI. Recently, Hastie et al. (2016) reported that the augmentation of memory during robot's interaction with a child significantly improved child's performance in a treasure hunt game exercise. However, to the best of our knowledge, we don't find numerous long term cHRI studies been undertaken that focus on the robot that adapts based on memory and consequently its effect on children engagement has also not been studied.

Despite all these methods, one of the pivotal questions that remain unanswered is: Which individual adaptations portrayed by a robot can result in maintaining long-term engagement during cHRI? In other words, we need to conduct studies where different adaptations as conditions can be evaluated in a long-term scenario. As adaptivity is significant for various types of education technologies, therefore, we also speculate that results of such studies can be extended to other similar technologies such as virtual agents. In this study, we attempt to address the aforementioned question. We conducted a long-term cHRI study at a school where children played a game of snakes and ladders with the NAO robot. The robot was programmed to display three different adaptations: Game (control group), Emotions and Memory. We wanted to learn which adaptations result in sustaining social engagement and which ones are the most effective adaptations. We conducted video analysis and coded both verbal and non-verbal communications of our sessions to measure social engagement and also interviewed participants on the last day of our study to get their preferences on different adaptations.

4.2 Background

4.2.1 Measuring Engagement

Engagement is a collaborative activity between two individual entities that combines both verbal and non-verbal interaction behaviours (C. L. Sidner & Dzikovska, 2002). It has been discussed in Human-Robot Interaction (HRI) literature that it is possible to engage with a robotic system with no conversation. It does not mean however that engagement is plausible without communication; indeed it is not. Gestures, facial expressions are other forms of communication that drives engagement (C. L. Sidner & Dzikovska, 2002; C. Sidner, Kidd, Lee, & Lesh, 2004). It is reported that non-verbal behaviours during a HRI comprise of eye contact, gaze, facial expressions and gestures. Similarly, verbal indications comprise of word utterances, vocalisation and sentences (Dautenhahn & Werry, 2002).

Eye-contact and gaze are the most commonly used units for measuring social interaction and communication as it is described in literature that humans repeatedly make eye-contact during communication especially while listening (Argyle & Dean, 1965). Any verbal communication that does not involve significant eye-contact is not considered a complete communication. In addition, several studies in HRI (Okita et al., 2011), and Children Computer Interaction (cHCI) (Rajagopalan et al., 2015) have used eye-contact and gaze as a measure of engagement and social interaction.

According to C. Sidner et al. (2004), gestures are an important part of non-verbal communication and also reflect on the amount of engagement between two individuals. Similarly, facial expressions including smiles (Castellano et al., 2009) also reflect on the level of engagement and interaction during a social communication. For instance, Castellano et al. (2009) showed that the amount of smiles at the iCAT robot increased with time.

Verbal Interactions including vocalisation (mumbling, whistling, or yelling) and word or sentences occurrences are also common variables to measure engagement and social interaction in both cHRI (Serholt & Barendregt, 2016; Kanda, Hirano, Eaton, & Ishiguro, 2004) and cHCI (M. I. Ahmad & Shahid, 2015).

4.2.2 Long-term Children-Robot Interaction

It is imperative to ascertain the implications of long-term cHRI in the educational domain, where typical interactions are not one-off but extend for several sessions. As learning is a long-term process (Wittrock, 1974), a one-off interaction is not a suitable case in an educational settings. One of the aspects of a long-term interaction is to understand its length. The duration of a long-term cHRI can be understood from the discussion in Human Computer Interaction (HCI) literature. We find reports on longitudinal studies lasting for five weeks (Karapanos, 2013). However, recently, it has been argued in the literature on social robots that a long-term interaction is not only dependent on the duration in terms of the number of days or weeks, but, it can also depend on the number and duration of sessions (Leite et al., 2013). According to Leite et al. (2013), "An interaction can be considered as "long-term" when the user becomes familiarised with the robot that her perception of such robot is not biased by the novelty effect anymore". We further believe that the duration of the long-term interaction may vary on the capabilities and overall scope of the robot and its interaction scenario. A robot with limited abilities and interactions can cause fading of novelty in less time.

In the past, a number of long-term studies have been conducted with children in schools and have resulted in declining children interest and engagement during cHRI. Salter et al. (2004) conducted a study with the *WANY* robot capable of obstacle avoidance and moving in the environment. The study was undertaken with

8 children for five weeks and showed that children lost interest in the interaction from the third session. Kanda et al. (2004) conducted a field trial with a humanoid robot, *ROBOVIE*, capable of identifying and teaching English skills to children at a school. The motivation and engagement of children declined at the end of the first week. Kanda et al. (2007) also conducted another field trial for two months with 37 children who interacted with the *ROBOVIE* robot capable of identifying a child, displaying more behaviours towards a child who interacted more, and trusting the child with it's secret. Results show that children kept engaged and interacted with the robot after two weeks. Komatsubara et al. (2014) also conducted a field trial with a *ROBOVIE* robot capable of identifying, performing gaze movements and teaching science concepts through asking questions to children at a school during break time. Children lost interest after the second week. The conjectured reasons were a non-flexible answering design. In addition, an extensive overview on several long-term interaction studies conducted with various social robots across various social domains can be found here (Leite et al., 2013).

There are also studies which have shown that the robot was able to successfully maintain engagement after long-term cHRI. For instance: Kozima et al. (2009) evaluated a *KEEPON* robot capable of displaying non-verbal behaviours such as gaze movements and displaying emotions with 27 children for 20 sessions. Results showed that *KEEPON* played the role of the mediator through displaying non-verbal cues and was also able to maintain engagement. In addition, Leite et al. (2014) also conducted a study with an iCAT robot capable of playing the game of chess through displaying emphatic behaviours. Their result showed that engagement was sustained throughout the five sessions. In summary, it can be inferred from these studies that a robot can maintain engagement if it tries to develop a social relation with a child or have a good number of autonomous behaviours.

Researchers have also reported pros of utilising an adaptive social robots in the education domain as specified in Chapter II.In addition, adaptivity is also a highly sought after design feature in most aspects of Human Computer Interaction (Cheng et al., 2013) as it promotes usability and better task performance. Moreover, Its applications can also be found in educational settings in HCI (dos Santos & Osório, 2004; Buche, Querrec, De Loor, & Chevaillier, 2003). For instance: Hassani et al. (2016) has recently presented an intelligent virtual environment that was utilised to improve learner's speaking and listening skills.

Keeping this background in mind, to the best of our knowledge, research that focuses on understanding the impact of different adaptations on the social engagement during long-term cHRI is not available and is needed. Therefore, our contribution is about understanding the effect of various adaptations performed by an Adaptive Social Robot on maintaining children Long-term engagement.

4.3 Study

The entire study and associated protocol was approved by the host university's ethics office (approval number H11429).

The purpose of our study was to evaluate the effect of three different adaptations performed by a social robot on child's engagement during a long-term cHRI. We evaluated three different adaptations for the NAO robot: 1) Game adaptations as the control group, 2) Emotion adaptations, and 3) Memory adaptations. Each adaptation was considered as an experimental condition. The rationale for our choice of memory and emotion adaptation came from a study conducted with teachers to understand their views on various adaptations portrayed by a robot M. Ahmad et al. (2016c). The results of study showed that teachers commented that memory and emotion adaptations can in turn motivate children. In addition, Baxter & Belpaeme (2014) has also emphasised the importance of memory adaptations. Moreover, Conati (2002) and Leite et al. (2014) have showed that the augmentation of the understanding of emotions by the robot can also effects social interaction.

We conducted our evaluation for three sessions to measure the aforementioned effect. We chose to run our evaluations for three sessions because literature shows that children loose engagement from the third session onwards (Salter et al., 2004). In addition, as discussed by Leite et al. (2013), an interaction can be considered long-term depending on the novelty wear-off time. We understand that children's novelty factor will wear off in the first two interactions because the communication scope for the robot in a snakes and ladders game is limited.

Our Hypothesis (H) based on the observations reported by (Mubin et al., 2013) and on the result of the study conducted by (Hastie et al., 2016) and (Komatsubara et al., 2014) is as follows:

H1 - Children's engagement will sustain for both Emotions and Memory adaptations as compared to the control group where there is no "user" based adaptation over sessions.

H1a - The engagement in terms of gazes, facial expressions, verbal responses and gestures will remain constant from 1st - 3rd session.

H2 - Children's engagement will be significantly higher for both Emotions and Memory adaptations as compared to game adaptations.

H2a - The engagement will decline for gaze, facial expressions, verbal responses and gestures for game adaptations.

H3 - Children's engagement will be higher for emotion adaptations as compared to memory adaptations for gazes, facial expressions, verbal responses and gestures.

4.3.1 System Description

Our system as shown in figure 4.1 comprised of a NAO robot and a *Snakes and ladders* game running on an *Android* tablet developed in Unity 3D. We implemented a server program responsible for communication between both the NAO robot and the game. The robot's behaviour was selected by following a decision making algorithm based on the game state and user's affective state and also depending on the adaptation type. In case of Game adaptation, we only used game state as an input. Whereas, in addition to game state, we also considered user affective state that included emotion recognition mechanism and other affective state that included face recognition mechanism for the emotion and memory adaptation respectively. This input was later passed on to the Behaviour selection and adaptation mechanism unit, where, it employed a decision making approach to select behaviours such as gestures and dialogues depending on the selected adaptation type from the behaviour processing unit. The behaviour was then requested from the database and later the NAO robot portrayed it. In addition, in case of memory adaptation, the game state or different events that happened in the game were stored in the database to keep track of the user profile.

In this section, we discuss different adaptations that were displayed by the NAO robot during children-game-robot interaction. We also present the interaction scenarios of our experiment and the experimental protocol.

4.3.1.1 Snakes and Ladders Game

We have updated the *Snakes and ladders* game implemented using Unity 3D game engine as shown in Figure 4.2 as we added stars as a third element alongside snakes and ladders on the game board. The rational for adding stars was to make game more interactive and challenging. The game rules are simple and easy to understand. Each player takes a turn rolling the dice, based on the number on the dice, the player makes the move accordingly. If the number 6 appears on the dice, the player gets another turn. Each player may face a star, a ladder or a snake while progressing towards 100 (finish line). On each snake, the player goes back to the tail of the snake on the game board. On each ladder, the player jumps to the number where the ladder

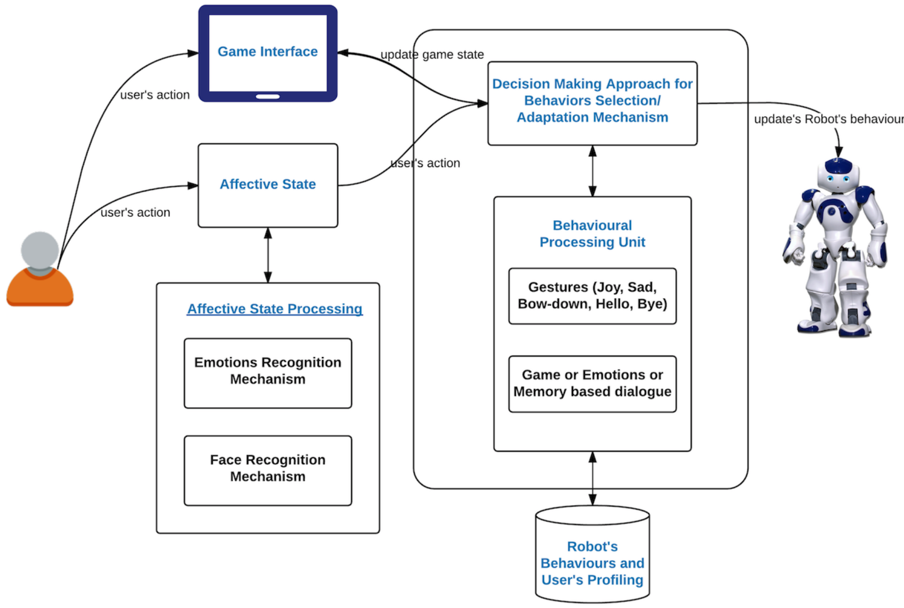

Figure 4.1: System Architecture

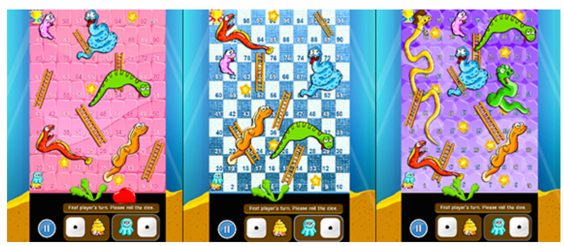

Figure 4.2: Snakes and Ladders game

is headed. On each star, the game randomly decides if the player should go from one to six steps forward or backwards. The player that reaches the 100 mark first, wins the game.

4.3.1.2 Game Adaptations

NAO robot was able to generate a supporting behaviour ranging from text-to-speech, and a range of gestures (thumbs up, high five, heads up, heads down, and hand gestures) by following the decision making mechanism. On each six (rolled on the dice), snake, ladder or star, NAO robot was programmed to generate an appropriate phrase based on the game state. We implemented various decision making scenarios:

On each six, NAO said: "Wow, you have six, you are playing extra-ordinary". On consecutive sixes, NAO said: "Another six, this is great, looks like you will win today" or alternatively portrayed a High Five gesture and said: "High Five, you are doing really well". If a child was behind the robot, NAO tried to encourage the child by saying: "Hey, you need some catching up, you are still on <<PLAYER CELL>> I might win today", if the child was ahead of robot, NAO said: "Hey, I need some catching up, I am still on <<NAO CELL>> you are going to win". If the child performed exceptionally well, NAO was able to show positive gestures such as thumbs up, and high fives.

We also implemented three different game boards as shown in Figure 4.2 with increasing difficulty level for the children. The robot adapted its dialogue based on the game board with every session. In the first board, we only positioned one snake near 100. This was followed by two and three snakes near 100 in the second and third game boards respectively. The robot was programmed to autonomously detect and adapt to different game boards. At the beginning of each session, the robot also informed every child about the new level in the game.

4.3.1.3 Emotion Adaptations

The NAO robot was able to detect children's emotions through analysing their facial expressions. Once an emotion has been detected, NAO was programmed to perform one out of the four actions based on the game state and/or the child's current emotion. Firstly, it generated a supportive dialogue about how it felt about the emotional state of the child. On detecting Happy emotion, NAO said: "I am so happy to see you smiling", "You are looking happy, I think you are enjoying the game". If the robot detected sadness, surprise, or angry, NAO said: "Are you sad?" or "Don't loose hope anything can happen in the game", "You look surprised, It's great, I can tell for sure you are enjoying the game", and "Are you angry, you can still win the game" respectively. Secondly, it displayed two different emotions (joy and sad) through showing gestures. We used the gestures as described here (Häring et al., 2011). Häring et al. (2011) presented sadness and joy based emotional expressions in two different ways. The expressions were designed for the *NAO* robot using body movement and eye colors. Thirdly, in case the child is behind the robot in the game, and if NAO detects that the child is sad, we implemented a mechanism that enabled a child to receive higher numbers on the dice to encourage the child. Lastly, it followed the decision-making mechanism as implemented for game adaptation.

In order to detect a child's emotion, we have used two mechanisms. Firstly, we trained data on basic human expressions (smiling and not smiling). We programmed NAO to capture user's facial expressions and detect their emotions using an algorithm EVP (2015). The algorithm uses the library from Mordvintsev & K (2013) to localise the mouth area to detect emotion of a user. The image captured by NAO is re-sized to 28*10 pixel containing only the person's mouth and surrounding areas. The image is then converted into grey-scale and flattened into a vector of length 280. A logistical regression programme then takes the vector and determines the emotional state of the user.

We also used an online *Indico* API developed in python Indico (2016) that enabled us to determine an emotion expressed in an image on a human face. The API returned a dictionary with 6 key-value pairs. These 6 key-value pairs represented 6 different emotions (happy, sad, fear, surprised, angry, and neutral). The API returns the probability values to inform the emotion on the human face, however, the probability values with less than 0.05 should be discarded. We didn't use fear emotion and ignored fear values in the 6-key pair because we didn't expect the child to feel any fear during the game play. In order to measure the accuracy of this API, we tested the API and found that if the probability value is greater than 0.25, it was highly likely that the suggested emotion is correct. Therefore, we only accepted the key-pair with the value greater than 0.25.

We used both mechanisms to decide on the emotional state of the child. However, our algorithm preferred smile detection higher than happy detected by the indico API. We detected the emotional state of the child after every 30 seconds during the gameplay. We also implemented a mechanism in order to avoid repetitive detection of the same emotions through comparing the previous occurrences with the current ones. In addition, in an attempt to calculate most likely user emotion through facial scan and avoid data loss, we programmed our vision algorithm to initially detect child's mouth. In case, it could not detect it, we took an image again and re-called the api to give us 6-pair values (happy,sad,surprise,neutral,fear,angry) of the user emotion. However, we ignored the fear values.

4.3.1.4 Memory Adaptations

We implemented memory adaptation based on the following situations: 1) Identifying children, 2) Identifying Children's friends through dialogue, 3) keeping track of child's and his/her friend's game performance and results, 4) storing information on the number of moves a child took to win the game, 5) remembering how many times a child's snake was near 100, 6) if the child got a ladder, 7) How many times the child got two 6's simultaneously, 8) How the child responded to greeting questions. We stored all this information in our local database. We used the principle of recalling and remembering an event to implement memory adaptations.

We again computed a decision making mechanism within the game that enabled NAO to generate memory-based adaptive dialogue based on the state of the game. For instance: on a negative star near 100, the robot said "You are on a Star, unfortunately you are going backwards on <<PLAYER CELL>> I remember you got a negative star last time as well", and on a ladder, the robot uttered, "Wow, this is great, you have a ladder. you are very lucky. I remember you had a ladder last time as well" or if the ladder brought the player close to 100, the robot displayed a High Five gesture along with saying "High Five, you are playing extraordinary". On winning the game, the robot said, "Remember! It took you <<NO OF MOVES>> to beat me last time, Can you tell me what is the secret behind your success".

4.3.2 Interaction Scenarios

We programmed NAO robot to autonomously play the game with children, however, speech recognition was controlled via a Wizard of Oz (WoZ) setup. One of the researchers had implemented a program to reply to basic preconceived questions during introduction and game-play phases. For instance, it involved responding to "thanks", "hello" and "how are you" respectively. The researcher responded to participant's queries through pressing a button on the WoZ program. The robot stayed quiet in case, where the child as questions out of its scope. NAO was capable of autonomously performing three different aforementioned adaptations. Each adaptation comprised of different behaviours such as: speech recognition, text to speech, displaying gestures, recognising and reacting to user's emotions and expressions, and keeping track of user's memory. These characteristics have been described in literature (Fong et al., 2003) for an Adaptive Social Robot. The interactions varied for all these adaptation types.

The Interaction Scenarios was divided into three sections: 1) Introduction, 2) Game play and 3) Game end. NAO began introduction through one-to-one interaction with a child by asking introductory questions: (Hello, I am NAO, What is your name, $<< NAMEOFTHECHILD >>$, Nice Name, How are you today?, How is your day progressing?, Today we are going to play snakes and ladders game, Have you played before?, you can go first) for game-based and emotion adaptation for all the session. In the case of memory adaptation, during the first session, NAO followed the same interaction style but also asked the child about his friends and then spoke about their game outcome. In the following sessions, the robot also identified the child before beginning the game and also reminded them about its previous responses through the dialogue in the introductory phase. The number of interaction between the child and the robot however, remained constant for all three adaptation types.

Once the game has started, the robot generated appropriate text to speech based on the dice outcome, snake, ladder within the game through following the afore-explained decision making mechanism. If the child was performing exceptionally well, for instance, getting a quick ladder, the robot then displayed a range of gestures such as high five, thumbs up, and clapping. In the case of emotion adaptation, in addition to game adaptation, NAO was programmed to conceive information about child's affective state such as: detecting child's emotions based on the facial expressions. NAO gave verbal feedback based on child's emotional state and state within the game. For instance; if the child is at a lower number than the robot and the child facial expressions are detected as sad, NAO said, "Don't loose hope, anything can happen within the game". In the case of memory adaptation, in addition to game adaptation, beginning from the second session with each child, the robot generated a response based on the previous interaction with a certain child. For instance, on a snake near 100, the robot says, "I remember you got a snake near 100 last time too", in the case of a winning, the robot reminded the child about the previous game result and also mentioned that his/her friend won the game today as well.

Upon winning or losing, the robot also congratulated or wished the child all the best for the next time combining speech with a range of gestures such as bow down, showing sadness, show surprise, or joy respectively. In the last session NAO gestured "bye" along with saying "Good Bye, today is the last game playing session, I hope to see you soon."

4.3.3 Setup and Materials

We conducted our study at a quiet room as shown in Figure 4.3 and Figure 4.4 respectively inside the school library. The room was divided into two sections, on one side, the child interacted with NAO sitting on a table along with a tablet device with a seat in front for the child. The NAO robot was placed in a sitting position in front of the child in order to get a the clear view of the child's face for detecting emotions. We also placed a video camera on one of the tables to video record all the sessions. On the other side, one of the researchers was controlling the speech recognition capabilities of NAO through a Wizard of Oz setup.

Figure 4.3: Setup: A Child playing snakes and ladders with NAO (front view) - Permission has been taken to use the picture with child's face.

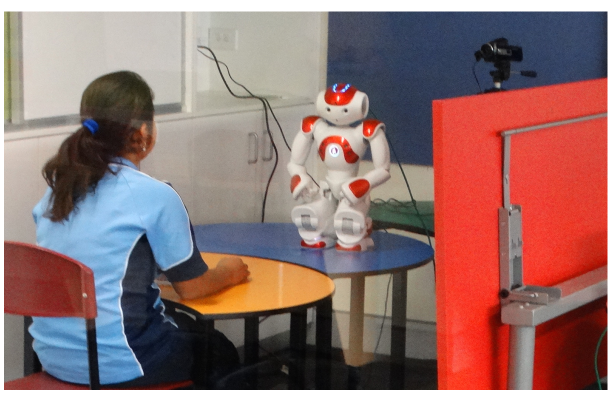

Figure 4.4: Setup: A Child playing snakes and ladders with NAO (back view)

We used NAO robot designed and developed by Aldebaran robotics. It is a humanoid robot measuring 58 cm in height. NAO is an interactive and adaptable robot

1) How did the NAO behave when you were losing the game?
2) Did you understand NAO's gestures and comments towards you? How so?
3) How did NAO behave when you played well?
4) What did NAO do that you didn't like?
5) What do you think about NAO understanding emotions during the game?
6) What do you think about NAO recalling and remembering past events during the game?
7) What do you think NAO would have done otherwise?
8) Can you mention three things you liked the most and 3 things you didn't like?
9) What other tasks would you like NAO to help you with?
10) Do you have any suggestions to improve NAO?

Table 4.1: Interview Questionnaires

partner. It provides researchers a platform to design various applications driven by their creativity and requirements. The game was running on a 10.1" android tablet.

We also used the questionnaires as shown in table 4.1 from Leite et al. (2014) to get their preference on different adaptation and feedback on their experience.

A consent form along with an information sheet was sent to the parents in order to receive their consent on the participation and video recording of their child during the study.

4.3.4 Participants

We conducted our evaluation at a primary school with 23 participants (16 girls and 7 boys). The participants were randomly divided into three groups. Out of 23 participants, 7 children participated in game adaptations, 7 in emotion adaptations and 9 in memory adaptations. The distribution of gender and ages is shown in Table 4.2. The study took place at a school during school timings with 5th and 6th grade children. The ages of participants were between 10-12 years. None of the children had interacted with NAO or any social robot before this study. All interaction were in English language and all children have English as their native language.

Adaptation type	Gender	Average
Game	6 females, 1 male	10.8
Emotions	5 females 2 males	11
Memory	5 females 4 males	10.6

Table 4.2: Participant's average age and gender per adaptation type

4.3.5 Procedure

Our study was a between-subject evaluation (3 conditions of robot adaptivity type) spanning a period of 10 days. The evaluation was conducted individually with one child at a time. Each child played the game with the same adaptation type of NAO robot 3 times on three different days (one session per day), for a total of 69 sessions (23 children * 3 sessions per adaptation type). We conducted our sessions on the 1st, 5th and 10th day respectively. Each group of children played the game on a tablet for one of the three conditions (game as a control group condition, emotion or memory adaptation) for the three sessions. Each session lasted for approximately 12 minutes comprising of a 1-minute introduction, a 10-minute game playing session with the NAO robot and a 1-minute end greeting session. The evaluator used a stop-watch to maintain the time consistency throughout the sessions.

The child played snakes and ladders game with the robot. In condition 1 (the controlled group), the robot adapted according to game performance (number on the dice, a snake, a ladders, and a star on the game board, state of the game). In condition 2, the robot performed game and emotion adaptation. In condition 3, the robot performed game and memory adaptation. The game playing session were 10 minutes long, in case there was no outcome after 10 minutes the robot was programmed to end a game in a draw. All of the sessions were video recorded.

On the last day, children were asked to participate in an individual interview to give their preferences and feedback on their understanding and perception of behaviours displayed by NAO during game playing sessions.

4.4 Video Analysis

To measure the effect of three different adaptations on social engagement, we conducted video analysis for the first and third sessions following the long-term interaction studies design guidelines on analysing videos as specified by (Leite et al., 2013). We also followed the guidelines and skipped the second session. In addition, we find examples in literature such as (Kozima et al., 2009) in which they video coded first 15 sessions and skipped sessions from 15 to 30 due to the time consuming nature of video coding. A total of 23x2 videos (9x2 for memory adaptations, 7x2 memory adaptions and 7x2 for game adaptations) were analysed. We divided our sessions in three time intervals: 1-minute introduction, 10-minutes game-play and 1-minute end-greetings. We were not able to code for 2 sessions for emotions adaptation and 1 session in game adaptation due loss of data for the end-greetings interval. In addition, we could not see the face of the child in the video as the robot's face came in front of the child's in one of the sessions for the memory adaptation.

Following the coding scheme as discussed by Serholt & Barendregt (2016), we are coding videos for following dependent variables: Gaze, Verbal Interaction, Facial Expressions, and Gestures. However, our coding scheme did not look into the negative indications as discussed by Serholt & Barendregt (2016). Two researchers were involved in video coding process. One of the researchers did not take part in the evaluation process. The second researcher coded 20% of the videos separately and discrepancies were resolved with consultation. The first researcher then completed the coding. We coded for both frequency and durations for the four different factors because it is important to measure how many times a child did a certain act and for how long did the child do it. We find examples in literature such as (Bartneck et al., 2007), who has also coded for durations to measure social interaction. The coding mechanism as shown in Table 4.3 was followed.

4.4.1 Gaze:

Literature shows that gaze that includes the robot's face can be considered as a sign of engagement (Argyle & Dean, 1965). We also found studies such as Bartneck et al. (2007) that have used gaze (looking at the robot) as one of the measures for social communication. Therefore, we coded the frequency and duration of the child-robot face gazing, face-table alternating (where the child saw the robot after facing the table), and face-researcher alternating (where the child saw the robot after looking at the researcher or else where) as gaze indications. We didn't count any other gaze indications as a sign of social engagement. For instance, where a child kept looking on the tablet. The transition took place when the child moved their gaze away from the robot for more then 2 seconds.

4.4.2 Facial Expressions:

Literature shows that smiles can be considered to signify social engagement with a robot (Castellano et al., 2009). Therefore, we coded frequency and duration of the different versions of smiles as shown in figure 4.5 as facial expressions. All other facial expression that do not include smiles, such as nervous, and/or confused expressions were not considered to be a sign of engagement. The duration of the smile was calculated until the change was observed in terms of expressions on the child face.

Figure 4.5: Different types of smiles - Permission has been taken to use the picture with child's face.

Gaze	**Facial Expression**	**Verbal Response**	**Gesture**
Robot face	Timid Smiles	“Hello”	Wave
Robot face-table alternating	Surprised	“thank you” “Okay”	thumbs up head shake
Robot face-researcher alternating	Flushed	“Yes” “No” “Good” “I am fine”	fist Nod High five

Table 4.3: Coding Scheme used to measure social engagement

4.4.3 Verbal Response:

All verbal responses including word or sentence occurrences, vocalisations in response to robot praise or response or different game events were coded as verbal responses. We coded for both the frequencies and durations. Most commonly used verbal responses included: “Hello”, “thank you”, “Okay, Yes”, and “No”, and “Good”. An exception to this rule was the utterance of “what” during the communication as it showed that the child was not able to understand the robot properly. We counted two different verbal responses if a pause of more than or equal 2 seconds was observed between two verbal utterances.

4.4.4 Gesture:

Literature shows that gestures are also a sign of social engagement C. L. Sidner et al. (2005). Gestures including “Wave”, “thumbs up”, “head shake”, “Nod”, “High five”, “bow” and “fist” were coded as gestures. We didn’t observe any pause when gestures were depicted by the child, all of the gestures were spontaneous.

4.5 Video Analysis Results

We conducted a repeated measure Analysis of Variance (ANOVA) with the *session* as the within-subjects factor with *two* levels and *adaptation type* as the between-subject factor using only one of the following set of Dependent Variables (DV's) 1) Gaze, 2) Facial Expressions, 3) Verbal Response and 4) Gestures. The repeated measure ANOVA was conducted with three different phase or interval for both frequencies and durations for the aforementioned DVs. The rationale for choosing different phases is based on a study (Serholt & Barendregt, 2016).

In this section, we present the result for frequencies and durations for three different internals (introduction-greetings, game-play, and end-greetings) and complete session respectively.

4.5.1 Introduction-Greetings

Results as shown in table A.1 show that for introduction-greetings phase, in case of frequencies, there was a statistically significant effect of session on gaze ($p < 0.001$), facial expressions ($p < 0.001$) and verbal response ($p < 0.001$). In addition, as shown in table A.2 we also found statistically significant effect of adaptation type per session on facial expressions ($p < 0.001$). Moreover, we found a nearly significant effect of adaptation type on verbal response ($p = 0.064$) as shown in table A.3.

The mean values plots with 95 % confidence interval as shown in figure 4.6 show that the number of facial expressions were found to be constant for memory adaptation. In addition, they nearly remain constant for emotions adaptations, i.e a minor decline was observed, however, game adaptation decline for the last session.

Results as shown in tables A.1, A.2 and A.3 show that for durations, there was also a significant effect of session on gaze ($p < 0.001$), facial expressions ($p < 0.002$) and verbal response ($p < 0.004$). In addition, we found a significant effect of adaptation type per session on facial expression ($p < 0.03$). Lastly, we also found a significant

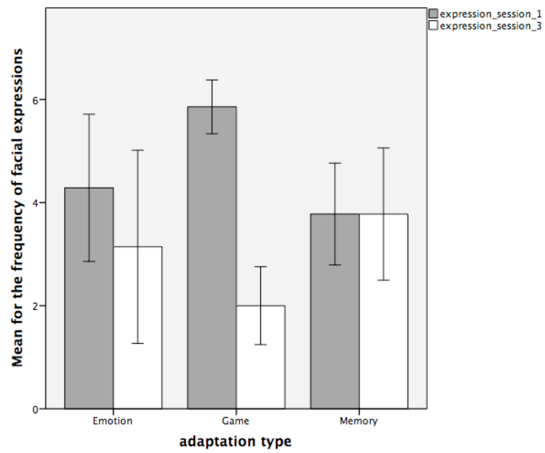

Figure 4.6: Means and 95 % confidence interval for facial expressions frequencies during introduction-greeting interval.

effect of adaptation type on gaze ($p < 0.01$).

The mean values plots with 95% confidence interval as shown in figure 4.7 show that duration of facial expressions from first to third session remained constant for memory adaptation. We witnessed a decline for facial expressions for emotion and game adaptation.

In case of durations, to further examine whether a different was significant of the adaptation type, we conducted a $Bonferroni$ post-hoc check. Results showed that Memory based adaptation was found to be significant in comparison with emotion adaptation ($p < 0.04$) and game adaptation ($p < 0.04$) in terms of gazes.

4.5.2 Game-play

Results in tables A.1, A.2 and A.3 show that in case of game-play for frequencies, there was a statistically significant effect of session on gaze ($p < 0.001$), facial expressions ($p < 0.001$) and verbal response ($p < 0.002$). In addition, we also found statistically significant effect of adaptation type per session on gaze ($p < 0.03$), facial expressions ($p < 0.006$), verbal response ($p < 0.04$) and gesture ($p < 0.02$) respectively. Moreover, we didn't find significant effect of adaptation type on all four DVs.

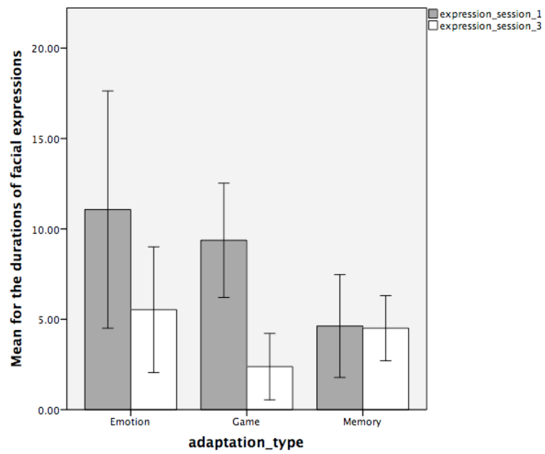

Figure 4.7: Means and 95 % confidence interval for facial expressions durations **(msec)** during introduction-greeting interval.

The mean values as shown in figures 4.8, 4.9, 4.10, and 4.11 show that the number of gazes, facial expressions, verbal responses and gestures remained constant for both emotion and memory adaptation. However, the engagement declined for game adaptation. Particularly, we witnessed an increase in the numbers for all DV's across sessions for emotion adaptation. Lastly, we also observed an increase in the number of gazes and gestures for memory adaptation from first to third session.

Duration results as shown in tables A.1, A.2 and A.3 show that there was a statistically significant effect of session on gesture ($p = 0.05$). In addition, we found a statistically significant effect of adaptation type per session on facial expressions ($p = 0.05$), and verbal response ($p < 0.04$) respectively. Moreover, we didn't find significant effect of adaptation type on all four DVs.

In terms of duration, the mean values as shown in figures 4.12 and 4.13 show that the verbal responses remained constant for both emotion and memory adaptations. However, the duration for both facial expressions and verbal responses enhanced for emotion adaptation. We also witnessed the minor and sharp decline for the facial expressions duration for memory adaptation and game adaptation.

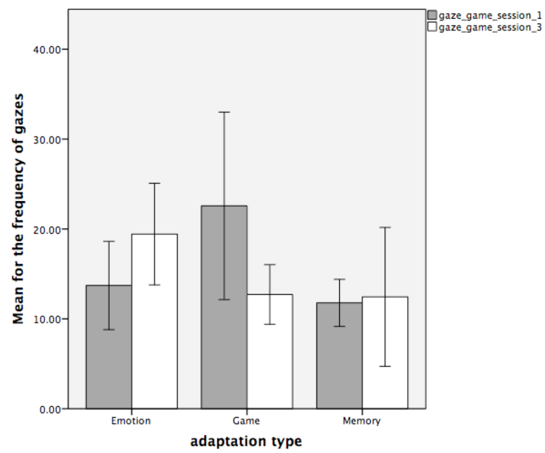

Figure 4.8: Means and 95 % confidence interval for gaze frequencies during game-play interval.

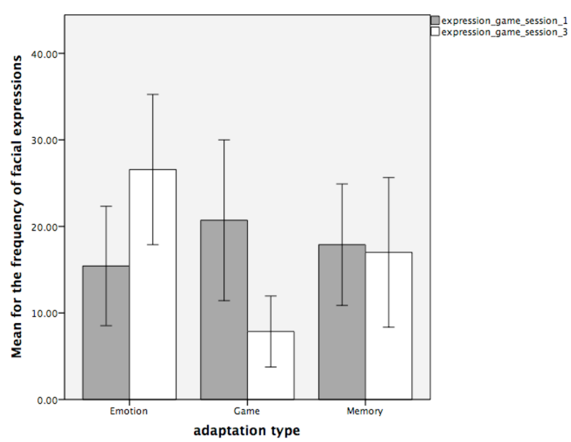

Figure 4.9: Means and 95 % confidence interval for facial expression frequencies during game-play interval.

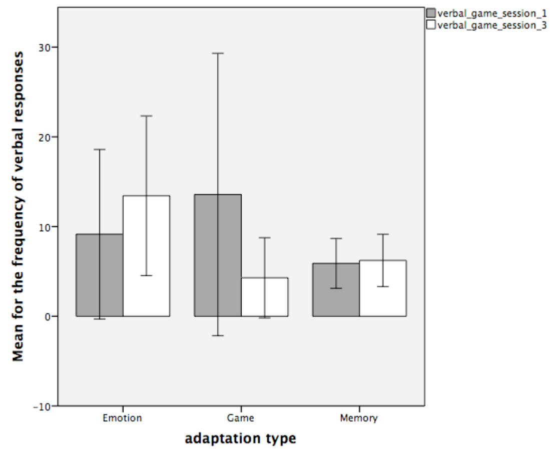

Figure 4.10: Means and 95 % confidence interval for verbal response frequencies during game-play interval.

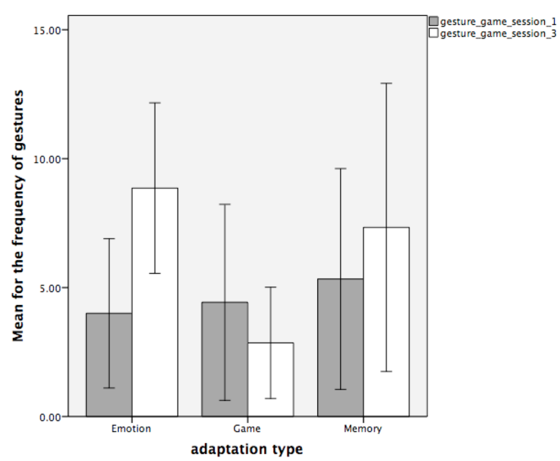

Figure 4.11: Means and 95 % confidence interval for gestures frequencies during game-play interval.

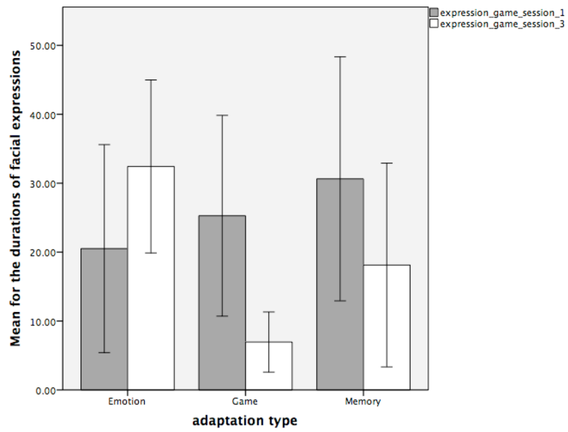

Figure 4.12: Means and 95 % confidence interval for facial expression duration **(msec)** during game-play interval.

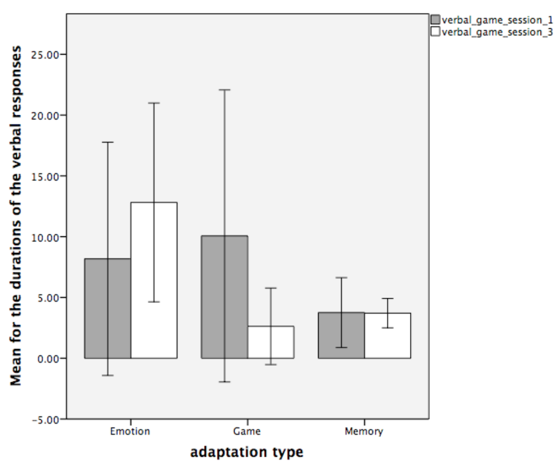

Figure 4.13: Means and 95 % confidence interval for verbal response duration **(msec)** during game-play interval.

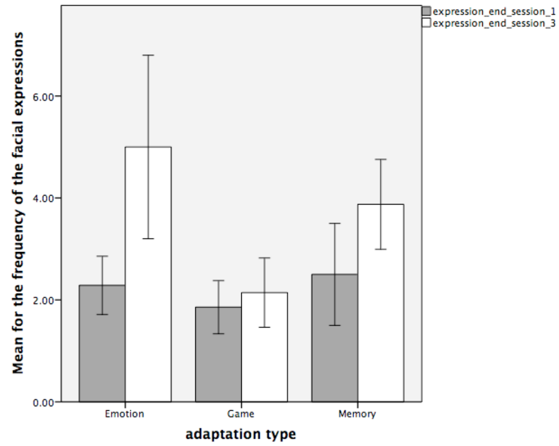

Figure 4.14: Means and 95 % confidence interval for facial expression frequencies during end-greetings interval.

4.5.3 End-Greetings

Results in tables A.4, A.5 and A.6 show that for frequencies, there was a statistically significant effect of session on gazes ($p < 0.02$), facial expressions ($p < 0.001$) and gestures ($p < 0.002$). In addition, we also found statistically significant effect of adaptation type per session on facial expressions ($p < 0.04$), and verbal response ($p < 0.05$) respectively. Moreover, we also found significant effect of adaptation type on facial expressions ($p < 0.04$).

The mean values as shown in figures 4.14 and 4.15 show that the number of facial expressions and verbal responses increased for both memory and emotions adaptations. However, the number of facial expressions remained constant for game adaptation but the verbal responses declined from the first to the last session.

To further examine whether a different was significant of the adaptation type, we conducted a $Bonferroni$ post-hoc check. Results showed that Emotion based adaptation was significant in comparison to game adaptation ($p < 0.04$) in terms of facial expressions.

Duration results as shown in tables A.4, A.5 and A.6 show that there was a statistically significant effect of session on gaze ($p < 0.001$), facial expressions ($p < 0.002$),

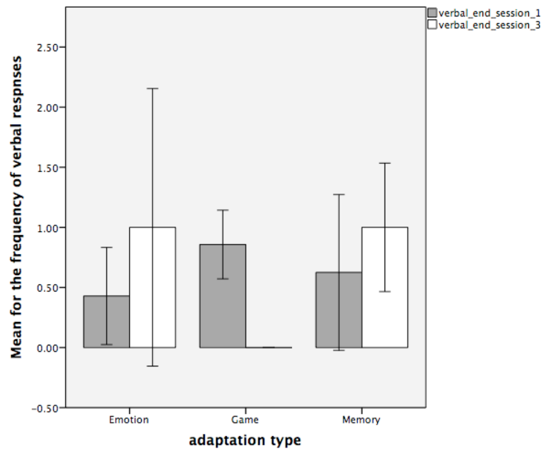

Figure 4.15: Means and 95 % confidence interval for verbal response frequencies during end-greetings interval.

verbal response ($p < 0.002$) and gestures ($p = 0.03$). In addition, we also found a statistically significant effect of adaptation type per session on gaze ($p < 0.002$), facial expressions ($p < 0.007$), and verbal response ($p < 0.02$) respectively. Moreover, we also found significant effect of adaptation type on gaze ($p < 0.001$), and verbal response ($p < 0.02$) respectively. Gestures were found to be nearly significant ($p = 0.06$).

In case of durations, we observed an increase from the first to the last session for both emotions and memory adaptations as shown in figures 4.16, 4.17 and 4.18 for gazes, facial expressions and verbal responses. However, for game adaptation, gaze remained constant but the facial expressions and verbal responses declined.

In case of duration, to further examine whether a different was significant of the adaptation type, we ran a $Bonferroni$ post-hoc check. In case of gaze, memory based adaptation was significant over emotion based adaptation ($p < 0.001$) and game based adaptation ($p < 0.001$). In addition, for verbal response, memory based adaptation was significant over game adaptation ($p < 0.04$). Emotion-based adaptation was significant over both memory ($p < 0.04$) and game based adaptation ($p < 0.05$) for facial expressions.

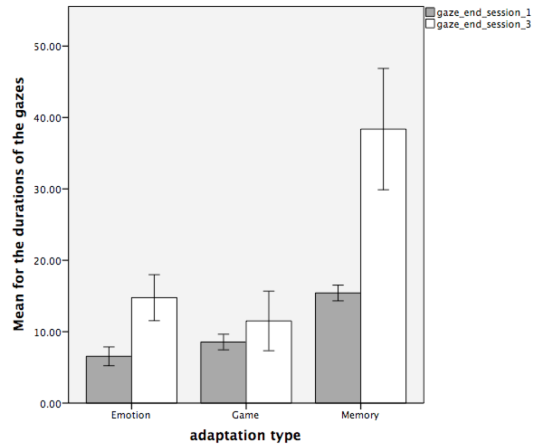

Figure 4.16: Means and 95 % confidence interval for gaze durations **(msec)** during end-greetings interval.

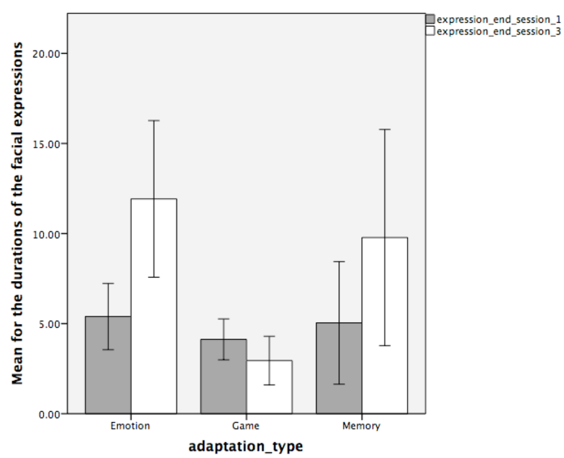

Figure 4.17: Means and 95 % confidence interval for facial expression durations **(msec)** during end-greetings interval.

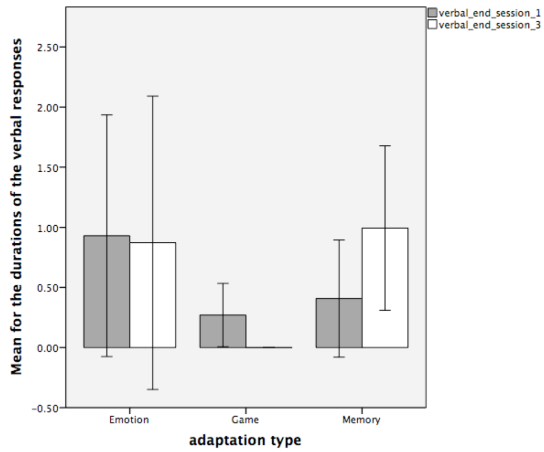

Figure 4.18: Means and 95 % confidence interval for verbal response durations **(msec)** during end-greetings interval.

4.5.4 Complete Session

Results in tables A.4, A.5 and A.6 show that for frequencies, we found a significant effect of session on gestures ($p < 0.03$). In addition, we also found statistically significant effect of adaptation type per session on gaze ($p < 0.04$), facial expressions ($p < 0.002$), and verbal response ($p < 0.02$) and gestures ($p < 0.01$) respectively. Moreover, we didn't find the effect of adaptation type on any DVs.

The mean values as shown in Figures 4.19, 4.20, 4.21 and 4.22 show that we witnessed an increase in the number of gazes, facial expressions, verbal responses and gestures for emotion adaptation. On the other hand, memory adaptation remained constant throughout the three sessions. Lastly, we observed a decline for game adaptation for gazes, facial expressions, verbal responses and gestures.

Duration results as shown in tables A.4, A.5 and A.6 show that we didn't find a significant effect of session on any DVs. In addition, we also found a statistically significant effect of adaptation type per session on gaze ($p < 0.04$), facial expressions ($p < 0.03$), and verbal response ($p < 0.04$) and gestures ($p = 0.05$) respectively. Moreover, we also found significant effect of adaptation type on gazes ($p < 0.02$).

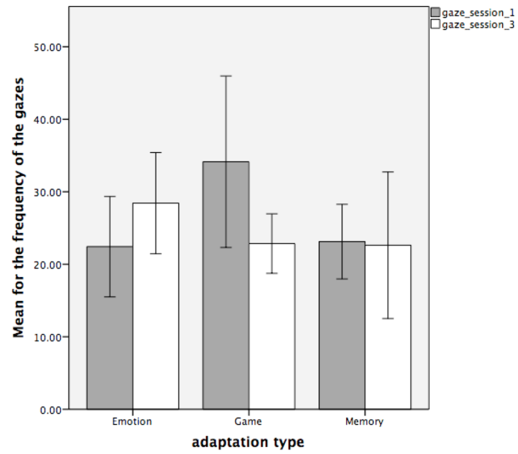

Figure 4.19: Means and 95 % confidence interval for gaze frequencies during complete session.

The mean values of durations as shown in figures 4.23, 4.24, 4.25 and 4.26 show that in case of emotion adaptation, gazes, facial expressions, verbal responses, and gestures enhanced across sessions. In addition, the memory adaptation remained constant for gazes and gestures, but, an acute decline was observed for facial expressions and verbal responses. Lastly, game adaptation declined in terms of all measurements.

To further examine whether a difference was significant for the adaptation type, we ran a $Bonferroni$ post-hoc check. Our results show that Memory-based adaptation were significant over both emotion ($p < 0.02$) and game based adaptation ($p < 0.05$) for the duration of the gazes.

4.6 Qualitative Results

We adapted interviews questionnaires from (Leite et al., 2014). The purpose of the interview data was to identify if children were able to recognise the three different adaptive robot behaviours. In addition, we also want to identify if the children were able to understand the behaviours as showed by the NAO robot? Moreover, we also wanted to receive their feedback on improving the robot's response mechanism. The

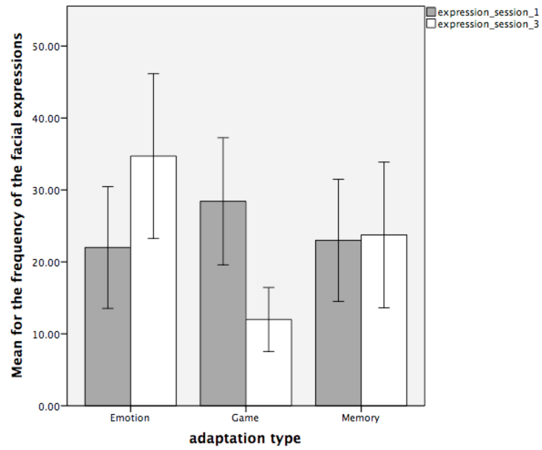

Figure 4.20: Means and 95 % confidence interval for facial expression frequencies during complete session.

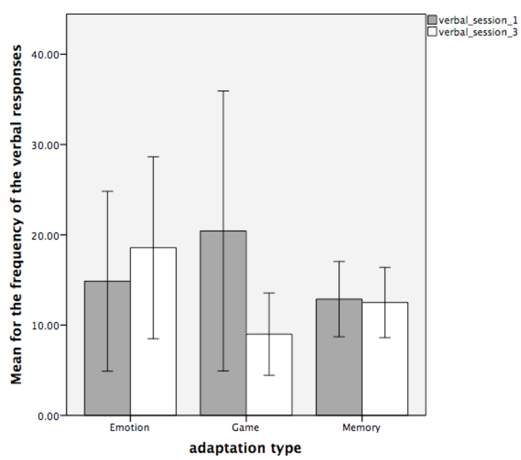

Figure 4.21: Means and 95 % confidence interval for verbal response frequencies during complete session.

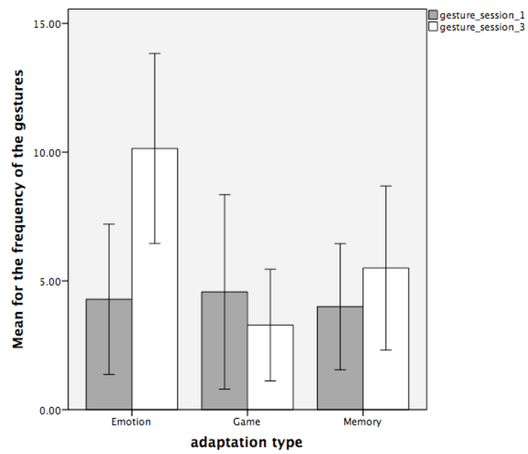

Figure 4.22: Means and 95 % confidence interval for gesture frequencies during complete session.

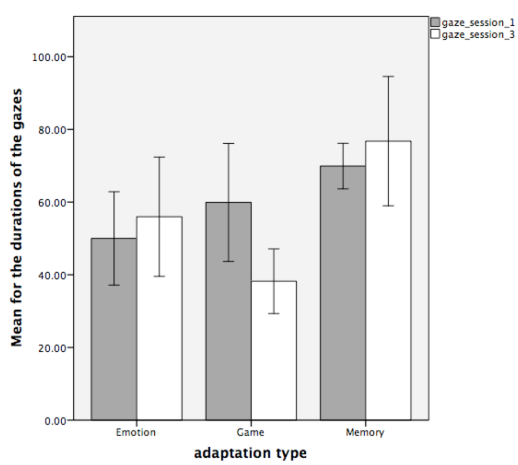

Figure 4.23: Means and 95 % confidence interval for gaze durations **(msec)** during complete session.

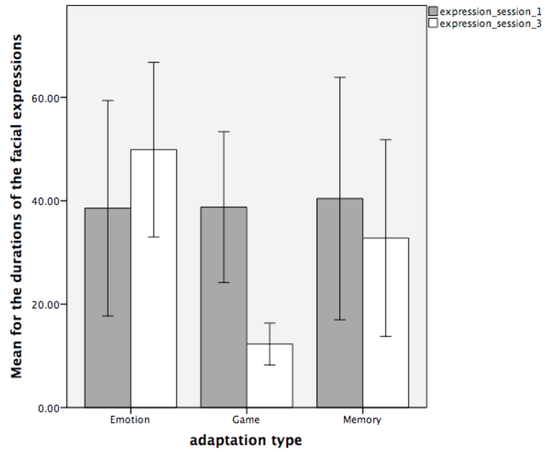

Figure 4.24: Means and 95 % confidence interval for facial expression durations **(msec)** during complete session.

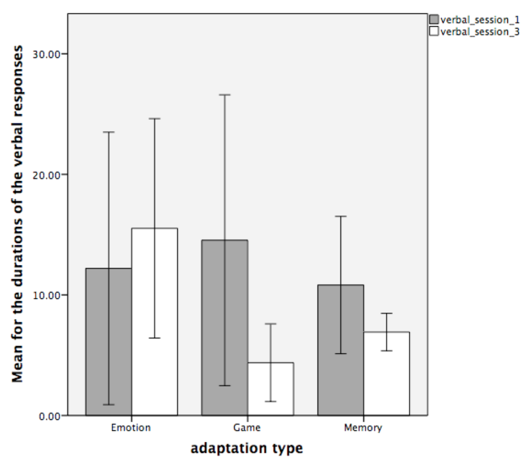

Figure 4.25: Means and 95 % confidence interval for verbal response durations **(msec)** during complete session.

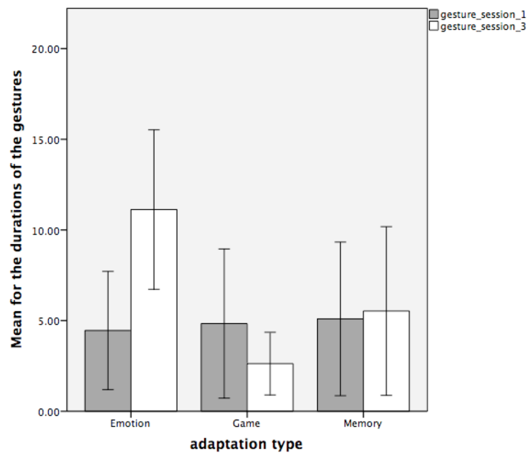

Figure 4.26: Means and 95 % confidence interval for gesture durations **(msec)** during complete session.

questionnaires are shown in table 4.1. We discarded Question (Q) 5 and 6 from the group of participants for condition 1. Similarly, Q5 from condition 3 and Q6 from condition 2 were removed from the interviews. The rationale for discarding questions was due to different type of adaptation performed by the robot for each group of children. In the analysis, we refer to different participants from different conditions as (P1,...,P7, C1),(P1,...,P7, C2), and (P1,...,P9, C3) respectively.

Results showed that in general, all of the children happily received the comments and gestures uttered or portrayed by NAO upon losing the game. Children felt that the robot was kind and emphatic towards them and it was able to form a bond with them in all three conditions. Children across all conditions reported:

NAO was good when I lost the game, he said: maybe next time you can win. (P2,C1)

NAO comforted me when I was losing but also gave me the advice to help me reach 100. Sometimes NAO would boost me after saying that he was at a higher number than me. (P7, C2)

NAO behaved in a respectful way and didn't criticise me. (P4, C3)

In reaction to Q4, Children in C1 and C2 reported that they would prefer NAO to remember them and their interactions with it. In C1, children also commented that they would like NAO to understand their emotions and would like them mention how did they (children) feel during the game-play. Children from the control group (C1) mentioned that NAO needs to add more actions and comments and ask more questions other than their day at the school. We understand that this was a result of novelty effect wearing off. In addition, an interesting result was that children complained that NAO robot speaks too fast at times and they could not understand it. It could have been due to the robotic voice of the NAO robot. Children also reported that NAO sometimes did not wait for them to respond to their response. We believe it may have happen due to the limitation with respect to no speech recognition mechanism during the study. It might also have happened because NAO would have reacted to another event happened during the game. One of the children said NAO did not show competitive behaviour because NAO was programmed to give hope and be emphatic towards the child in all circumstances. They wanted NAO to be proud of himself as well rather than just focusing on the child. For instance, children said:

NAO was talking too fast and I was not able to understand it (P4, C1)

I didn't like how NAO asked questions because sometimes he didn't wait for me to answer them. (P8, C3)

I think NAO should have been proud of himself too and not always focus on me P2,C1

All Children also in response to Q5 and Q6 reacted positively towards NAO emotion and memory detection behaviours. Most of them mentioned that recalling and remembering previous events motivated them. Upon asking, if they would appreciate if NAO could make memory mistakes in terms of recalling past events, they said, they do not expect a robot to make mistakes. They reported as:

It was really good because he knew my emotions. He told me how to play snakes and ladder in a good way. (P4, C2)

I love that robot! I have never seen a robot with so much emotion before! It was a *real person talking to me (P5, C2).*

I liked the robot, It could tell when I was smiling and when I wasn't. (P3, C2)

NAO has a very good memory and recalled them to me. (P3,C3)

In response to Q8 and Q9, one of the children (P4,C3) appreciated that NAO motivated them to compete more with the NAO robot when it was able to name their position on the game board, particularly, when NAO was ahead of them. One of the children (P2, C2) was not pleased with the emotion detection algorithm, they commented that they didn't feel sad, but, the robot thought they were sad because they were behind in the game or when we lost the game. Some children (P2,C1).(P4,C1) from the control condition also wanted the robot to recall events and remember their names and previous communications. In addition to playing games with NAO, in general, children wanted NAO to help them with learning. They want NAO to help with languages, mathematics and science. They also wanted NAO to speak in different languages and read a book.

It was really cute and understands things! IT needs to fix its facial scans to detect emotions. (P2, C2)

I thought it was very motivating. (P4,C3)

As for Q10, children in C1 suggested that NAO needs to see and detect emotions such as he should be able to identify when they are angry. The children also reported on improving some of the answering mechanism, they mentioned, at times, NAO is giving comments when they are not required. Lastly, all of the children reported their desire to play again with the robot.

4.7 Discussion

The results presented above show that overall, children's social engagement in terms of frequencies and duration of gazes, facial expressions, verbal response and gestures did sustain for both emotions and memory adaptions. Therefore **H1** was supported.

Our results for the three phases and complete session also showed that the robot adapting on the basis of emotions and previous events resulted in sustaining children engagement. Therefore, **H1a** was also validated.

We also observed a minor to a major decline in children social engagement over sessions for game adaptations for gazes, facial expressions, verbal responses and gestures during all phases except for the gaze and facial expressions during end greetings phase. Therefore, **H2** and **H2a** was also supported. We conjecture that simple adaptive emphatic dialogue is not enough and therefore game adaptation does not sustain long-term children robot engagement. We believe that children felt uninterested as the robot got repetitive with its comments mechanism. Similar trends have been also reported in the literature for other studies such as (Kanda, Hirano, Eaton, & Ishiguro, 2004; Komatsubara, Shiomi, Kanda, Ishiguro, & Hagita, 2014). Our qualitative results also indicated that children wanted the robot to have a realistic and varying dialogue and new comments. In other words, they wanted robot to have novel comments in subsequent interactions. We also understand that in future, considering the scenario of an education system, it states that effective learning situations for children are those which are personalized (Chang et al., 2010; Prain et al., 2013). Therefore, A repertoire of personalised responses by NAO can be effective, However at this point, they cannot be expected to compare to a human teacher.

The consistency of gazes and facial expressions during the end-greetings can be a result of robot performing gestures such as joy and bye in the end-greeting phase. The gestures from the NAO robot could have kept children engaged at the end phase. As it has been also shown by (Serholt & Barendregt, 2016) that gazes to the robot's greeting over the three sessions were the highest for all the participants. This could have happened because of different movement depicted by the robot while displaying different gesture.

The results of the comparison between emotions and memory adaptations showed that in general, the robot adapting based on the children emotions resulted in not only maintaining the social engagement but, also resulted in increasing the frequencies and durations of certain measurements during different phases. Hence, our **H3** was also supported.

We observed that for both game-play phase and complete session, the emotions based adaptations resulted in enhancing the frequencies for facial expressions, verbal responses, and gestures from first to the third session. In addition, in terms of durations, we observed an increase in facial expressions and verbal responses. Emotion adaptation was also significantly higher in comparison to memory adaptation during introduction and end greetings for both the frequencies and durations of facial expressions. We believe the reason why emotions were most effective is because emotion is the most basic principle of social interaction (Andersen & Guerrero, 1998). Similarly, we also find various studies that showed that emotions or emotional expressions have a positive influence on children during cHRI. For instance: Tielman et al. (2014) also developed a model for adaptive emotion expression for the *NAO* robot and studied the effect of adaptive emotion expression on the interaction behaviour and opinions of children. Their results show that children reacted more positively towards a robot with emotional expressions as compared with no emotions. In addition, our qualitative results also showed that children liked it when the robot was able to inform them about their emotion. Children were also keen to know about the robot's emotional state.

We have also observed a significant amount of variation among children in terms of overall engagement in our results for different phases during the first and third sessions. We conjecture that one of the reasons for this variation could be the choice of the suburb where we conducted our study. We had participants from different cultures and this may have resulted in the aforementioned variation. As it can also be found in literature, Shahid et al. (2014) has emphasised the importance of cultural background during the children-robot interaction and showed in comparison study between Pakistani and Dutch children that children from Pakistan were more expressive towards the robot as compared to the Dutch children. Additionally, it could also be due to difference in the gender of participants in our study. As it was found that male children responded more positively towards the *ZENO* robot as compared to female children when researchers attempted to measure the impact of affective facial expressions displayed by a *ZENO* robot on children's behaviour (Cameron et al., 2015).

In comparison with the measurements for both adaptation types, the durations of gazes were significantly higher for memory adaptation in case of introduction-greetings, end-greetings and complete session. We conjecture that in the case of emotions, some of the children may have felt shy or intimidated initially to look at the robot when an emotion was detected. This may have happened based on the introvert personality of some of the children as we didn't choose participants based on their personality. For instance: Abe et al. (2014) developed a play strategy for shy children and confirmed that shyness affected the relationship between the child and robot. In addition, our qualitative results also indicated that most of the children were engaged and motivated when the robot recalled previously held events. Moreover, we did not take time of the day into consideration because we started in the morning and the sessions went throughout the day. Therefore, a child may have been more patient in the morning session than in the afternoon session. As literature also suggests, better school performance can be an outcome during morning time (Vollmer et al., 2013).

For end greetings, memory adaptation was preferred over emotions adaptation for the verbal responses. This happened because the robot reminded the child about the friend's victory or the child's previous victory. Due to this positive findings, we understand that memory is the future of long-term children-robot interaction as also identified by (Baxter & Belpaeme, 2014).

In summary, we found that emotion adaptations were found to be most effective in comparison with memory or mere context-based snakes and ladders game adaptations in terms of maintaining long-term social engagement. In addition, the significance of memory was also found to be critical as it did sustain the social engagement between the child and the robot. Based on our finding, we believe that our results have deeper implications to the field of educational technology and HCI. They may reflect where virtual agents can be utilised in different game scenarios to maintain long-term social engagement. We also understand that our results may vary based on the type of the game for instance Chess or any similar game. However, as the outcome of our game is based on luck, therefore, we conjecture that similar types of games may end in resulting similar outcomes. Based on the findings of this study, we speculate that both emotion and memory based adaptations will also positively influence children learning such as language and mathematics learning in different educational interactions.

4.8 Limitations

We were limited to the number of participants per adaptation type but as most of the cHRI studies (Leite, Castellano, Pereira, Martinho, & Paiva, 2014; Kozima, Michalowski, & Nakagawa, 2009) are conducted in schools, therefore, it is not possible to accommodate more participants in number. In addition, we were also limited to receiving consent on video recording for our sessions.

It can be argued that our participants are not equally divided among genders. However, we requested from school the equal number of children per gender, but again due to video recording permission, we were able to receive more female participants. It should also be noted that most of the cHRI studies does have more male as compared to female (Leite et al., 2014).

The number of sessions can be another limitation of our study when it comes to being categorised as "long-term". However, our selection of a number of sessions is grounded in literature as it shows that from the third session onwards children lost engagement (Salter et al., 2004). Moreover, we believe that the scope of snakes and ladders game was limited and during first two sessions, children were able to overcome the novelty effect.

We also had a technical limitation with respect to speech recognition through the WoZ, in case, the child asked questions that were out of robot's scope, the robot stayed silent. We intend to integrate these missing questions in our Woz in our future studies.

4.9 Conclusions

In this chapter, we report our findings of a study conducted with children where they played the snakes and ladders game with a NAO robot. The NAO robot was able to perform three different adaptations based on the game state, child's emotions (facial expressions) along with game state, and memory (previous game events). We measured the effect of three different robot's adaptations on children social engagement during a long-term interaction. Our findings show that the type of adaptation does have an effect on the social engagement long-term children-robot interaction. The robot adapting on the basis of children's emotions and game state resulted in the highest level of social engagement in terms in comparison with memory and game based adaptations. In addition, memory adaptations were also found to be critical and it also resulted in sustaining long-term social engagement during a children-game-robot interaction.

CHAPTER V

Emotion and Memory Model - Long-term User based Evaluation

In this chapter, [1] we present an emotion and memory model for a social robot. The model allowed the robot to create a memory account of a child's emotional states. The robot then adapted its behaviour based on the developed memory. Our rationale for the creation of an emotion and memory model was based on the findings of our previous study presented in Chapter IV, where we found that both emotion and memory based adaptations resulted in sustaining long-term engagement during children robot interaction.

5.1 Introduction

A growing interest in the field of the social robotics has been towards the use of social robots in Education. Robots have historically been employed as a tool to teach computer programming skills in the past (Lawhead et al., 2002; Williams, 2003).

[1] This Chapter is based on a peer reviewed workshop paper and also an accepted journal paper

Ahmad, M. I., Mubin, O., Shahid, S., Orlando, J. (2017). Emotion and memory model for a robotic tutor in a learning environment. In Proceedings of the Seventh ISCA workshop on Speech and Language Technology in Education 2017, August 25-26, 2017, Djur, Stockholm, Sweden (pp. 26-32).

Ahmad, Muneeb Imtiaz, Omar Mubin, Suleman Shahid, and Joanne Orlando. 2019. Robots Adaptive Emotional Feedback Sustains Children Social Engagement and Promotes their Vocabulary Learning: A Long-term Child-Robot Interaction Study (To Appear) Adaptive Behavior.

However, more recently, due to the introduction of Humanoid robots, new possibilities have emerged to use robots in ways other than as a tool to teach skills from various subjects. As mentioned in the previous chapters, there is an emphasis on the need for a mechanism that will enable the robot to adapt its behaviour according to the characteristics of the user/student. The applications of such adaptive social robots (AdSoRs) have been enlisted in Chapter II, however, most of the studies conducted with these AdSoRs have been based on short-term interactions and did not capture children's repeated interactions with a robot. This area of understanding is needed as long term engagement with a robot is essential for understanding the role of robots in the future educational landscape. Additionally, learning is also a long-term endeavor therefore, we need to evaluate the potential of educational robots over weeks, months or years and ideally deployed in real life settings.

Researchers face many technical and social challenges during long-term interaction with social robots (Tapus et al., 2007b). One of them relates to the decline in user interest in the interaction over time as mentioned in Chapter IV. The reasons for this decrease in interest is robot's repetitive behaviour (Kanda et al., 2004), the loss of any novelty factor (Leite et al., 2014), the robot or task being boring or lack of relationship or feeling or closeness to the robot. To address these challenges, it is important to consider the definition of the term, "long-term" interaction as defined in chapter IV. We further believe that the duration of the long-term interaction may vary on the capabilities and overall scope of the robot and its interaction scenario. A robot with limited abilities and interactions can result in the fading of novelty in less time. To address these challenges, it is again emphasised to implement various autonomous adaptation mechanisms for a social robot to overcome the aforementioned effect as identified in Chapter II. For example, these mechanisms can be based on the user's emotions, memory, or personality (Leite, Martinho, & Paiva, 2013; M. Ahmad, Mubin, & Orlando, 2016b; Leite, Castellano, Pereira, Martinho, & Paiva, 2014).

The autonomous adaptation mechanism for a social robot can be implemented using different approaches. It can either be through utilising machine learning algorithms to direct the robot's behavior (Gao et al., 2017) or by following cognitive models that describe how humans create memory or how emotions are regulated in diverse situations and later applying them to generate the behaviour for the social robots (Ho et al., 2009). For instance: Belpaeme et al. (2012) proposed a model for adaptive strategies for sustainable long-term social interaction based on the theories in cognitive sciences. Trafton et al. (2013) presented a cognitive architecture named ACT-R/E (Adaptive Character of Thought-Rational / Embodied) that enables the robot to predict what a user will do in a certain scenario through understanding previous knowledge about the user. Leite et al. (2014) designed an emphatic model applied on an iCAT robot capable of playing chess with children. As stated earlier in this and previous Chapters, we find limited research on social robots that have been used as partners with students in a learning environment during long-term interactions (Leite, 2013b). Additionally, the models in HRI literature based on the memory have not been evaluated to study the impact of memory on user perception during long-term interaction (Leite, Martinho, & Paiva, 2013; Ho et al., 2009; Kasap & Magnenat-Thalmann, 2010). Moreover, there is still need for a model that can be customised and can be integrated in real-time social settings (Jeon, 2017). Lastly, the impact of a model that enables a robot to generate a behaviour based on memory is also under-studied with educational robots. We therefore, find a gap for a model for an educational robotic partner that can be employed to promote student learning.

Another important aspect while designing an AdSoR that can be utilised for the learning settings lies in the type of feedback it generates through adapting it based on user's emotions or memory. In other words, the significance of the effect of feedback on the student learning has been well-emphasized in literature (Butler & Winne, 1995; Hattie & Timperley, 2007). As the use of educational robots is on the rise, therefore, we believe that the need for evaluating the effect of different kinds of robot's feedback on student learning is needed. One of the recent relevant studies conducted with children where three different types of robot's feedback (peer-like, adult-like, and control) were evaluated in a one-off second language tutoring setup. Results showed that different types of feedback didn't affect the engagement measured through eye-contact, however, children seemed to perform more independently in the condition where robot provided peer-like feedback (Haas et al., 2017). These initial result highlighted the significance of feedback during robot tutoring scenarios, however, there remains a gap where such an impact has been studied in a long-term scenario. We therefore also attempt to study the effect of three different kinds of robot's feedback on student learning.

In this chapter, we present the design, implementation and evaluation of an emotion and memory model for a social robot. The model enables the robot to create a memory of the information retained during different user's emotional states and select an appropriate related behaviour accordingly. The model is based on the theory on how humans create memories of an emotional state/event/episode (J. LeDoux, 2007). Our goals were two fold; first, we wanted to evaluate the model through measuring its effect on children's social engagement during long-term children-robot interactions (cHRI). Second, we wanted to study the effect of model-driven adaptive behavioural feedback generated by the robot on the children's overall learning (memorisation) of vocabulary in our long-term study. To achieve our goals, we conducted a 2 week long-term cHRI study to evaluate our emotion and memory model. We programmed the NAO robot to play a version of the snakes and ladders game; the game had been modified such that learning Robot Interaction Language (ROILA) (Mubin, 2011) became an integral component of the game and children should be able to learn ROILA in a playful and interactive way. We implemented three types of robot's emotional responses based on positive, negative and neutral emotional events happening during the snakes and ladders game play. Our research question focuses on studying the effect of robot's positive, negative and neutral emotional response on a child's engagement and learning performance in a vocabulary memorisation task during a long-term cHRI. To the best of our knowledge, this effect has not been studied during the long-term child-robot interaction.

In this study, we choose to focus on vocabulary learning as the interaction task because it is one of the essential components of language learning (Moskovsky & Alrabai, 2009). Vocabulary learning helps improve listening, reading, writing and speaking skills (Rinaldi & Ispita, 2012). Additionally, studies have shown that the amount of words a child learns in early years lead to academic success in upcoming childhood years (Becker, 1977). Similarly, the number of words learnt also leads to quick understanding and learning of grammar (Goldfield & Reznick, 1990). Moreover, we choose the artificial language ROILA for this study because it was created based on the rules, syntax, and principles of the major natural languages of the world (Mubin, 2011). The choice of language allows us to mitigate the confounding factor of children having different linguistic backgrounds; while it will always be an influence, this may be lessened in the case of ROILA (Mubin, 2011) because it has no connection with other languages and dialects spoken. We chose game play as a focus because the significance of play and interaction in education has been well described (Vygotsky, 1980; Rieber, 1996).

Our work has primarily two main contributions. One of them lies in the design and evaluation of an emotion and memory model for a robotic partner or tutor that can be utilised to promote children's learning and social engagement through providing personalised feedback based on user's emotional memory. Secondly, we also measured the effect of different types of robot's emotional feedback on the children learning and engagement during a long-term children robot educational setup.

5.2 Background

5.2.1 Significance of interaction and play in Education

Different learning theories have emphasised the significance of social interaction and play on children's development. Vygotsky (1980) discussed the role of social communication and interactions during a child's learning development in sociocultural theory. Vygotsky (1980) introduced the concept of the zone of proximal development (ZPD) and defined it as "the distance between the actual development level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peer" (p. 86). In short, these theories emphasised the importance of learning under the guidance of an adult or competent peer because it acts as a catalyst for learning development. In addition, play theory (Rieber, 1996) also stressed on the value of play for learning activity in childhood.

Games as an end product presents an ideal representation of the aforementioned theories. Games naturally provide a playful environment and strongly encourage social communication. Games are defined as "an artificial constructed, competitive activity with a specific goal, a set of rules and constraints that is located in a specific context" (Hays, 2005). Games consist of several attractive features such as rules, goals, conflict, competition, challenge, multiplayer, interaction, feedback, outcome and representation (Prensky & Berry, 2001). Considering these facts, we have designed a modified snakes and ladder game that we used to teach vocabulary in a playful and interactive way during children-robot interaction.

5.2.2 Role of Emotions in Memory

The effect of emotions on human memory has been widely reported in the emotion and memory research. A body of research (Canli et al., 2000; Cahill et al., 1996) shows that emotions enhance human memory in tone, while another claims that emotions enhance central information at the cost of peripheral details (Levine & Pizarro, 2004). Several Studies, where participants were exposed to negative and neutral pictures, showed that memories of the pictures rated as emotionally intense were remembered better as compared to the neutral ones (Canli et al., 2000; Cahill et al., 1996). Another study examined the effects of different emotions (happiness, anger and sadness) on the participants and showed that different types of emotions have different effects on human memory (Levine & Burgess, 1997). We also find studies showing that pleasant emotions and usually remembered better than the unpleasant ones D'Argembeau et al. (2003); Comblain et al. (2005). In summary, the aforementioned research indicates that there seems to be a relationship between memory and emotions.

Considering these finding on the relationship between emotions and memory, we also wanted to investigate the impact of robot's emotional feedback on children's emotions and thereby how it may be able to have an effect on children's memory in terms of memorisation of vocabulary during a child-robot educational scenario. For instance: when a robot provides negative feedback on child's performance, how does it effects user's emotional state and how does it effects their memorisation of vocabulary.

5.2.3 Adaptive Social Robots in Education

In the recent literature, we find research on the applications of Adaptive Social Robots in education. For instance: an autonomous Maggie robot was programmed to play games and promote edutainment (Gonzalez-Pacheco et al., 2011). Similarly, a NAO robot was used in a study to help children learn arithmetic through playing a game (J. B. Janssen et al., 2011). In addition, An Educational Assistant robot was designed capable of improving user engagement in a learning task (Szafir & Mutlu, 2012). Moreover, it has also been used to teach Turkish sign language (Uluer et al., 2015). Most of the aforementioned studies on robots in education show that they have been successfully used mainly during one-off child-robot interactions. In general, the application of robots in a long-term interaction to support learning is understudied (Leite, 2013b).

In the most recent literature, we do find examples in which researchers have implemented various adaptations and personalisation mechanisms on the social robots. For instance; Gordon et al. (2016) presented an autonomous robot that adapted the kind of affective feedback it provided during a vocabulary learning activity. The adaptive strategy based on the reinforcement learning algorithm was evaluated with children for two months. Results highlighted the benefits of personalised feedback in terms of increasing children's engagement. Coninx et al. (2016) also presented an adaptation mechanism for the robot where it could adapt its behaviour through switching between multiple activities during a single interaction. Leyzberg et al. (2014) conducted a short-term experimental study where children solved a logic puzzle with the assistance of a personalised and non-personalised feedback providing robot. Although the mechanism for personalisation was not very sophisticated, results showed that personalisation can be beneficial for robotic educational setups. Kennedy et al. (2016) conducted a study with the robot, capable of providing two different kinds of verbal feedback during a vocabulary learning task. They compared the effect of different kinds of robot feedback on children vocabulary learning. Their results showed that they didn't observe a learning difference for different kinds of robot verbal feedback, however, in comparison with a no-robot condition, learning was found to be better for the robot condition. Baxter et al. (2017) conducted a long-term HRI study with children to promote their learning. They compared two conditions of a NAO robot where it showed personalised and non-personalised behaviour during a learning interaction. Their results showed that children learning performance on a mathematics task was better for the personalised condition in comparison with the non-personalised one.

Other examples of studies where robot depicted various kinds of adaptations or personalisations during various learning tasks can also be seen in literature (S. Lee et al., 2011; Kory & Breazeal, 2014; Westlund et al., 2017).

Keeping the trends and results of the aforementioned research in HRI, we conjecture that personalisation has, in general, resulted in a positive effect on learning. However, we also believe that to fully understand the value of personalisation or adaptions in robots, we need to conduct more long-term studies in HRI. In addition, we also conjecture that robot's adaptation can be implemented in many ways. For instance; the kind of personalisation presented in (Baxter et al., 2017) is based on the mathematics task and we believe it is significant to find how several sophisticated kinds of robot's personalisation and behaviour adaptation based on a working model may affect children learning on different tasks during HRI. Additionally, the use of memory to implement personalisation and then study its effects on user's perception and learning in a long-term setup is also understudied (Leite et al., 2013). Moreover, we also didn't find many examples in HRI literature where the effect of different kinds of robot's emotional feedback on user learning outcome and their social engagement has been studied in a longitudinal setup. Therefore, firstly, we are reporting a method to implement an adaptive robot based on memory that can be used to promote children learning skill development and engagement during a long-term children-robot interaction. Additionally, we also want to understand the effect of different robot's emotional feedback on children learning and engagement in a one-to-one long-term HRI setup.

5.2.4 Memory Systems In Human-Robot Interaction

Memory is an essential component of a social being and modelling of human-like memory in the machines has always motivated researchers in different fields. Human Memory system is categorized in two types; short-term memory (STM) and long-term Memory (LTM). STM is a limited capacity system that stores and maintains the information temporarily. LTM is a large capacity system that maintains the stored information for a long time in the human memory. LTM is further divided into declarative and procedural memory. The declarative memory is formed on the basis of the episodic memory that stores information about relevant past events (Tulving et al., 1972).

The use of employing memory has been gaining attention in HRI field and most research has been reported on utilising episodic memory (Ho et al., 2009; Kasap & Magnenat-Thalmann, 2010). Ho et al. (2009) proposed an initial memory model for a virtual or robotic companion based on the way humans retain STM and LTM. The memory was based on the context of interaction with the human user and focused mainly on storing and retrieving of autobiographic memory, a form of episodic memory. The purpose of the model was to mitigate the effect of the loss of interest during long-term interaction during HRI. Additionally, Kasap & Magnenat-Thalmann (2010) proposed a model for episodic memory to support affective interaction during long-term HRI. The model was applied during a robotic tutoring scenario and was tested during a short-term evaluation setup. The proposed models in both cases emphasised the need for using memory to facilitate long-term HRI. We also understand the findings from Kasap & Magnenat-Thalmann (2010) were encouraging however, an extensive evaluation of the system was missing. Additionally, in the recent review on social HRI has also highlighted that the benefits of the utilisation of memory are still unclear and needs attention (Leite et al., 2013).

Most recently, researchers have conducted studies where they have used memory to investigate its effect on the user interests and learning. For instance; Hastie et al. (2016) conducted a study with an emphatic robotic tutor to investigate the effect of incorporating memory in the HRI on children likeability of the robotic tutor. Results showed that there was a positive effect of adding memory as it helped the learner in finishing the task. However, the virtual robotic tutor with memory was perceived to be less likeable. Leite et al. (2017) also studied the effect of persistent memory during repeated HRI on children's perception of the robot. Their findings show that the preference on the likeability and friendliness of the robot was based on the age of the children as younger children (4 to 6 years old) prefer a robot with no use of memory. On the other hand, children aged between 7 to 10 years old preferred otherwise. We conjecture that these results may be influenced by the task and as the study was conducted in a one-off setup in one case, therefore, a long-term investigation is needed.

In general, as also highlighted by Leite et al. (2013), a need to investigate the benefits of memory during HRI is required. Additionally, we find memory models in the HRI literature for the artificial robotic companion but their application in terms of implementation and evaluation on the social robot in a real-life scenario remains an open challenge. Therefore, we present our work through understanding the ways in which humans stores and retrieves memory of during different emotional states and later try to apply this process on the robot to create memory in a HRI educational scenario. To the best of our knowledge, we didn't find an example of an implementation and evaluation of the application of creating and utilising memory based on the process of what information humans store during different emotional events (Levine & Pizarro, 2004) during a social HRI educational setup. Similarly, we also don't find an example where the effects of the use of such memory have been studied on the user learning and social engagement. Moreover, the novelty of our work lies in applying the process of storing and retrieving memory of emotional events of the user by the robot that is different from the previous work on using episodic memories (Kasap & Magnenat-Thalmann, 2010). We, particularly, refer to the robot imitating the memory of a user. Lastly, it is also understood that emotional events are often recalled better than normal events (Christianson, 2014), therefore, we believe incorporating such a process may enhance familiarity during HRI.

5.3 Emotion and Memory Model

Our model is grounded in the process of formation of emotional memories as described by J. LeDoux (2007) and also shown in the Figure 5.1. It is well-recognised that humans create memories of both positive and negative emotional experiences. It is known that the long-term memories are formed in a variety of systems but broadly can be divided into two categories: 1) Conscious (declarative) Memory and 2) Unconscious (procedural) Memory System. The conscious memory refers to the explicit memory system and 2) memories stored unconsciously refers to the implicit memory system. Human stores memories about different emotional situations in both kinds of memory systems. Emotional memory is defined as "a special category of memory involving the implicit (probably unconscious) learning and storage of information about the emotional significance of events" J. E. LeDoux (1993). On the other hand, the memory about the emotional situation (memories about emotions) is a part of conscious memory and refer to the explicit memory system. In general, the formation and retrieval of the emotional memories happens in the following manner; emotional events or experiences are processed in the human sensory system. They are later transmitted to the temporal lobe or to the amygdala in order to form either an explicit memory or an implicit memory. The memory of these emotional events is retrieved in case of an occurrence of a cue. This cue is processed by the sensory system that later leads to retrieval of both explicit or implicit memories. Based on the aforementioned science of the creation and retrieval of the memory about the external emotional event, we have created a model for social robots to store explicit or conscious memory of different emotional experiences of the user and later incorporate this memory in its future interactions with the user. The rationale for choosing user's memory of emotional events for the robot was to make it more familiar and natural

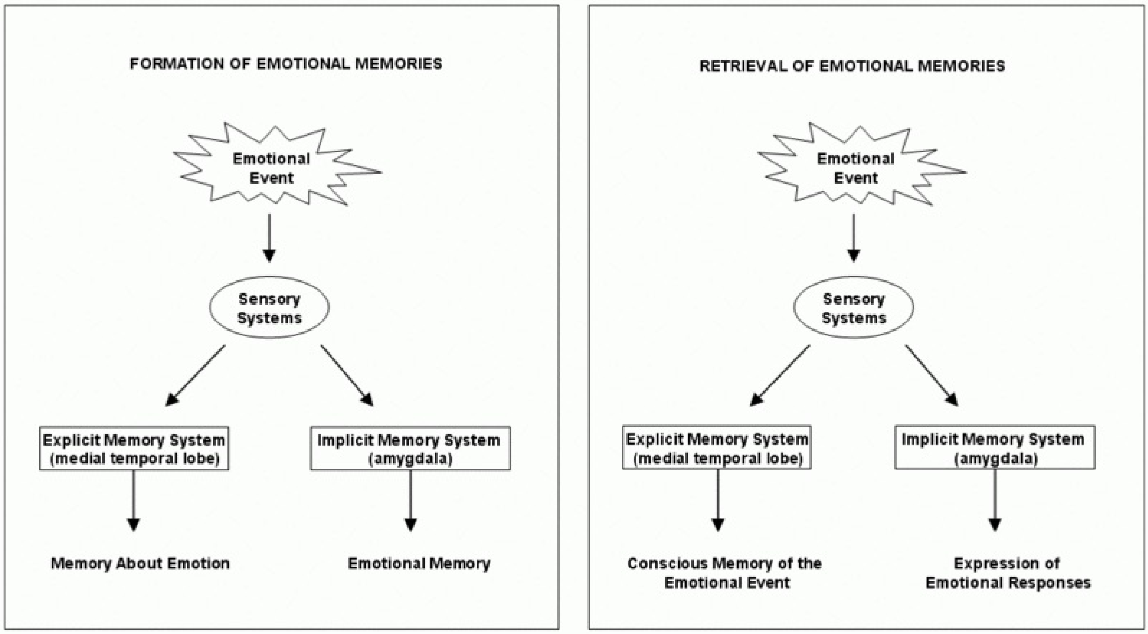

Figure 5.1: Formation and Retrieval of Emotional Memories; Image taken from (LeDoux, 2007)

as this can result in using robots as more effective tools (Dautenhahn, 1998).

As we are creating the memory of user's emotional events, it is necessary to define both positive and negative emotional events. Positive emotional events are described as events when goals are achieved or no immediate problems are encountered towards achieving the goal. Negative emotional events are registered as impediments towards a plan and causing loss to achieve a certain goal. On the other hand, Neutral events are situations that do not significantly threaten an outcome in either positive or negative ways (Bower, 1992).

Levine & Pizarro (2004) presented a review on emotion and memory research and showed that different types of information are remembered under various emotional states. A user's emotional state is directly related to an emotional situation. A positive or negative situation would refer to positive or negative states. It is therefore necessary to understand the information that should be stored in an emotional state or at an emotional situation. According to Levine & Pizarro (2004), in general, during

Discrete Emotions	Motivational State	Central Information
"Happiness"	"maintain current state; attain new goal."	"broad range of information from general knowledge and the environment"
"Fear"	"avoid or escape threat of goal failure."	"sources of threat; means of avoiding threat"
"Anger"	"remove obstacle to goal attainment."	"goal; agents obstructing goal attainment."
"Sadness"	"adjust to irrevocable goal failure."	"outcomes and consequences of goal failure."

Table 5.1: General rule for the type of Information stored during positive and negative Emotional Situations; taken from (Levine 2004).

positive emotions mainly happiness, humans store a broad range of information from general knowledge and the environment to their memory. Depending on the type of negative emotional state (sad, fearful, or anger) during an emotional situation, humans store different types of information. For example; Sadness leads to remember about the outcomes and consequences of goal failure. Anger leads to store information about goals or agents obstructing goal fulfillment. Lastly, fear leads to storing information about the source of threat and means of avoiding the threat (Levine & Pizarro, 2004). Table 5.1 summarizes general rule for the type of information stored during different emotional situations.

Based on the general understanding of what information human stores during different emotional states, we have designed our emotion and memory model, as shown in Figure 5.2, to enable a robot to create a memory of user emotional events. It is worth mentioning here that the rationale for our choice on an emotional event was based on considering two parameters; one lies with the definition of the aforementioned positive and negative events and two lies on the detection of user emotional state during the event. This is due to the reason that it is not necessary for an agent to depict positive or negative emotions who is trying to reach the goal. It may depend on the task as the task can be too easy to finish, therefore, we believe, it is important to take both parameters into consideration. The model enables the social robot to use the created memory in its dialogue and behave accordingly. The selection of the created memory during robot's future interaction was grounded on the process of the retrieval of the memory of emotional event (J. LeDoux, 2007).

The purpose of the model is to enable the robot to simulate user's memory of emotional events during various education/edutainment setups. In essence, the idea is enable robot to generate comments through using this memory. We believe this will enhance personalisation and can be used to facilitate different kinds of learning such as concepts from science or mathematics or languages during children-robot education based interactions. Our rationale is grounded on the afore-mentioned recent positive findings with respect to the use of personalisation in general during educational HRI.

Our model has the following modules: 1) Inputs, 2) Emotional Event Calculation (EEC), 3) Memory Mechanism Generation (MMG) and 4) Behaviour Selection Unit (BSU).

5.3.1 Inputs and Pre-Conditions

The model has three input types: 1) Game events, 2) User emotional states, and/or 3) Learning states. Additionally, the pre-conditions for applying the model includes marking the game event/states as positive or negative according to the aforementioned definition of the positive and negative events. Additionally, it also includes marking their impact as high or low during the game to understand the dynamics of the event. These inputs amalgamate to create an emotional event during the interaction.

5.3.2 Emotional Event Calculation

The emotional event is computed based on the type of game state or event (positive or negative or neutral), OR learning outcome (positive or negative) AND user emotional state (happy, sad, angry, fear, surprise, neutral) in the EEC module. The game event is either marked positive or negative as a pre-condition. Similarly, the emotional state is calculated through facial scan, mainly using a third party api; *indicio*. The negative emotional states are marked as Sad, fear and angry, whereas positive emotional states include Happy. Similarly, the learning outcome is either correct (positive) or incorrect (negative). To identify an emotional event to be positive or negative, we consider both the dynamics of the positive or negative game event along with the emotional state. It is done to make sure the event is coherent and according to the situation. For instance; a positive game event AND a positive emotional state would be categorised as positive emotional event. Similarly, a positive game event AND a negative emotional state would be categorised as a positive or negative emotional event based on the dynamics of the game event and also due the limitations of the existing emotion recognition methods (**?**). The dynamics of the game event refers to its either low or high impact on the output of the game. Additionally, a negative game event and positive emotional state would also be categorised as negative or positive emotional event based on the context of the game event. This means in case the impact of the negative game event is low and the emotional state is positive, it would be categorised as a positive emotional event. Moreover, the negative game event and the negative emotional state would be categorised as the negative emotional event. In our context as an exemplar, a positive event following our inputs may be one where the user moves the game piece closer to the end goal in the game; at the same time, the emotional state of the user is also positive (happy). A negative emotional state may be where user scores poorly on the learning test during gameplay and a sad emotional state is detected. Based on the afore-described criterion, the emotional event type is computed in the EEC module and it is later transmitted to the MMG module.

5.3.3 Memory Mechanism Generation

Based on the type of the event, the MMG module provides data on the type of information remembered under various emotional situations, we then send this information to our Memory Processing Unit. In this unit, we store the "central information" contained by the user under various emotional states as highlighted in the Table 5.1. For instance; in a happy event, the information about the game event and its context or the history of past similar event will be remembered. Similarly, in case of a sad event; a bad move causing the game event will be remembered. We create the memory for the robot in this unit and stores it into a database. In addition, in the case of an occurrence of same event type during same circumstances, we update our database with the new event and send the information back to the MMG module.

5.3.4 Behaviour Selection Unit

The MMG module later transmits this information to the BSU, that is responsible for selecting an appropriate behaviour or response. The BSU makes a database query with the passed information and returns the behaviour on the occurrence of the similar event. Lastly, the robot displays the behaviour. The behaviour of the robot is created based on the memory of different events. The robot use the behaviour to comment based on what is happening during the interaction. More specifically, the robot plays the role of a "commentator" and reacts positively, negatively, or neutrally on different events. In summary, the model enables the robot to maintain the memory of how humans felt at particular instants in the game or learning scenario and what were the main outcomes/events, basically during game and later behaves accordingly.

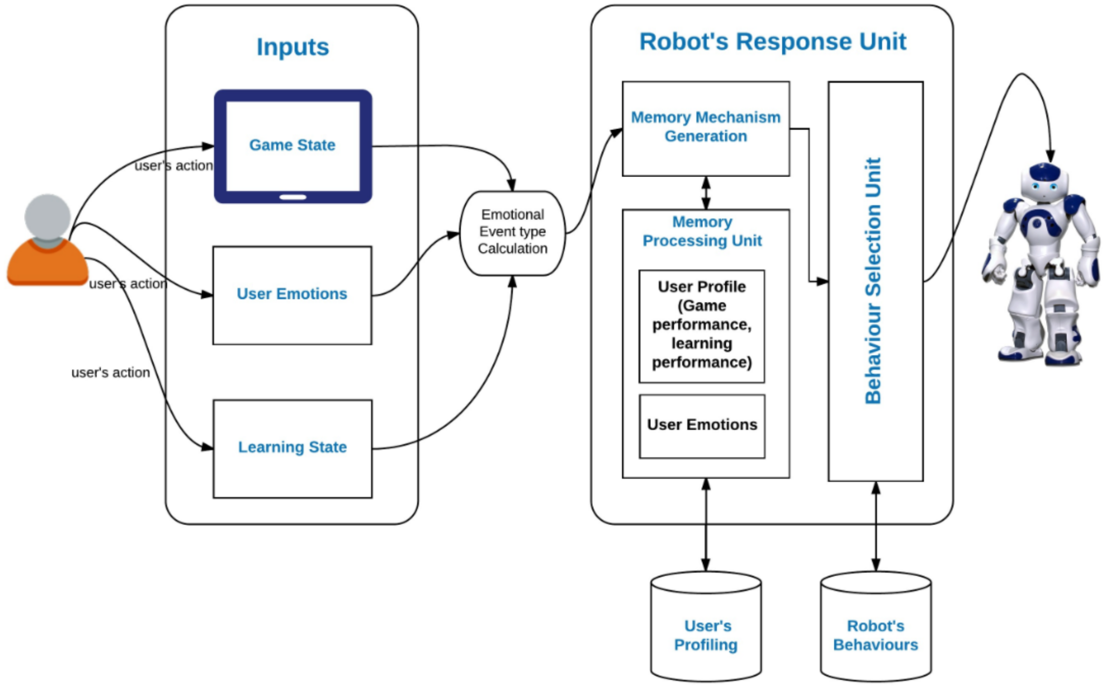

Figure 5.2: Emotion and Memory model

5.4 Research Method

Our research tried to explore two different aspects. Firstly, we wanted to understand how well our emotion and memory model for the robot performed in terms of teaching vocabulary to children and maintaining social engagement, mainly calculated through individually recording the duration of user's eye-gaze facing robot, Facial expressions (smiles), verbal responses, and gestures. during the long-term children-robot interaction. Secondly, we investigated the effect of robot's emotional feedback on the retention of children's vocabulary during a long-term interaction.

Keeping these aspects in mind, we tried to find answers to the following Research Questions (RQs) in the context of children-robot interaction:

RQ1 - Which of the following has a better effect on the child's retention of vocabulary and social engagement; a robot displaying positive (supportive), negative (critical) or a neutral emotional expression (feedback) combining both verbal and non-verbal behaviours?

RQ1a - How does a robot's positive, negative and neutral emotional expression affect the child's retention of vocabulary across sessions during a long-term interaction?

RQ1b - How does a robot's positive, negative, and neutral emotional expression affect the child's social engagement measured in terms of the duration of user's eye-gaze facing robot, facial expressions (smiles), verbal responses, and gestures during a long-term interaction?

RQ2 - What is the effect of our emotion and memory model directed robot's response on Social Engagement measured in terms of the duration of user's eye-gaze facing robot, facial expressions (smiles), verbal responses, and gestures across sessions during long-term HRI?

RQ3 - What is the effect of our emotion and memory model directed response on immediate retention of vocabulary across sessions during the long-term HRI?

Our Hypotheses based on the RQs are as follow:

H1a - A robot reacting positively to the child's vocabulary learning outcome will result in better children vocabulary retention followed by the negative and neutral feedback across sessions during the long-term interaction session.

H1b - The positive (supportive, encouraging) emotional feedback of the robot will also result in better social engagement measured in terms of the duration of user's eye-gaze facing robot, facial expressions (smiles), verbal responses, and gestures of children in a educational scenario.

H2 - The feedback generated through the process followed by our emotion and memory model would result in sustaining social engagement across all the sessions.

H3 - The feedback generated through the process followed by our emotion and memory model would result in generating overall higher immediate retention of vo-

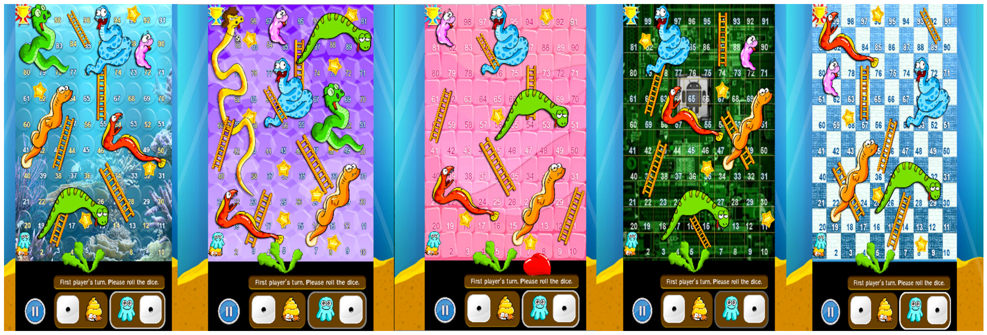

Figure 5.3: Snakes and Ladders

cabulary across sessions.

5.4.1 System Description

To test the applicability of our model, we implemented a scenario where the NAO robot plays the Snakes and Ladder game with a child during an one-to-one interaction. In this section, we present our modified version of the game. We also discuss the mechanism we used to calculate a type of emotional event along with the type of information stored in our system. Lastly, we give information on a selection of the robot's behaviours under different situations.

5.4.1.1 Snake and Ladders Game:

We modified the snakes and ladders game as discussed in Chapter IV to facilitate vocabulary learning. We updated rules of the Snakes and Ladders game as shown in Figure 5.3 and also introduced stars on the game board. On every snake appearance in the game, NAO robot was programmed to teach a new ROILA word to the child. The word will appear on the screen with an image based description as shown in Figure 5.4. The same process was repeated on each snake. In the case of a ladder, the child would take the ladder. On every star, a positive or negative number appeared on the dice suggesting the player to move forward or backwards. Lastly, the child was declared the winner by the robot when he/she reached the 100 mark.

Figure 5.4: Different words from ROILA appearing on the screen.

5.4.1.2 Applying the Model in the Snakes and Ladders Game:

To identify the type of an emotional event, we categorised both positive and negative types of game events based on their effect during the game play and also children reactions coded in one of our previous studies on similar game events (M. Ahmad, Mubin, & Orlando, 2017a) in Chapter IV. In the previous study, we coded for the significant game events such as the appearance of a snake, a ladder, a positive/negative star near or away from the 100 mark, a continuous six on the dice, continuous wastage of turns near 100, and winning or losing the game. To compute the emotional state of the player, we performed automatic facial scans as described in our previous research (M. Ahmad, Mubin, & Orlando, 2017a) and in Chapter IV. We used an online *Indico* API developed in python (Indico, 2016) that enabled us to determine an emotion expressed in an image on a human face captured through the camera. The API returned a dictionary with 6 key-value pairs. These 6 key-value pairs represented 6 different emotions (happy, sad, fear, surprised, angry, and neutral). The API returns the probability values to inform the emotion on the human face, however, the probability values with less than 0.05 should be discarded. We stored six different emotions values after every 10 seconds of the interaction. On each significant game event, we calculated the current emotional state by taking an average of the most recent six emotional states stored in our system. In essence, we computed the emotional state of the user through taking both the present and past values of the user's emotions through facial scan. Lastly, the learning state referred to the outcome of the words taught during the interaction.

In table 5.2, we briefly present our list of selected emotional events calculated on the basis of three aforementioned inputs of our model along with the type of information stored during these events in order to create robot's memory. We also include examples of the NAO robot's behaviour during these emotional situations. For instance, considering the definition of a negative emotional event, a snake near 100 will be rated as a negative event because it hampers the child from winning the game or thwarts the child from achieving the final goal. As mentioned earlier, in case of the negative event, we store the information based on the emotional state of the user. Therefore, in case of sadness, we stored information about previous game outcome after a snake that appears near 100 and the emotional state of the user. Similarly, for anger, we stored information about the number of times a snake is encountered near 100. Lastly, for fear, such as fear of losing, information is stored about the opponent (robot location on game). In addition, a ladder near 100 will be considered a positive event because it is helping the user achieve the end goal. As in positive situations, a broad range of information is stored about the environment, therefore, the information such as the number of previous ladders near and far from 100 or the past game outcome was stored. In a case, where the emotional state of the user is happy and the user receives a snake on the board, we store the robot's position to determine if the robot was ahead or behind the user. The behaviour selection of the robot uses the memory of previous emotional events to generate context-aware verbal and non-verbal responses either independently or simultaneously. As our purpose was to confirm our model's applicability, we used decision making statements to chose robot's behaviour.

We created a database of the robot's behaviour consisting of all plausible emotional events during the snakes and ladder game. On each event, the robot displayed the most appropriate behaviour by retrieving it from the database. The robot's behaviour on each event was based on the event type. The robot reacted positively through providing encouragement on a positive event. On a negative event, it reacted sadly and stayed neutral on the neutral event. In the table 5.2, for understanding, we only enlist a few behaviours [2] .

5.4.1.3 Applying the Model during the Vocabulary Learning Test:

As we used our model to generate behaviour for the robot to teach vocabulary, it was justified to apply the model during the testing phase, where the robot asked about words learned during their Snakes and Ladders game playing session.

Keeping the same criteria, we also used learning outcome, child's emotional state as the inputs for the model. We generated the behaviour of the robot through following the similar aforementioned mechanism. In table 5.3, we briefly present an example of selected emotional situations and respective verbal and non-verbal robot's behaviour. For instance; if a user remembered the word and the emotional state of the child was happy, we stored the information about the word and session. Similarly, when the children could not remember the word, we stored information depending on the detected emotional state. In the case of happiness, we stored of the user's emotional state and also about the word and the session. In the case of sadness, we stored the number of attempts to memorize the word. In the case of anger or fear, we stored information about the session where the child had answered the word correctly because the child may be angry about forgetting about the word or the child may feel threatened to match the previous performance on the test.

As specified earlier in this chapter, one of our goals was also to study the impact

[2] Snake Far 100 refers to the case, where the position of the user is at least 50 steps away from 100 (winning position).

Emotional Event	Information	NAO's Behaviour
Game Event: Snake near 100 **Emotional State:** Sadness **Event Type:** Negative	User's emotional state, Game outcomes caused by snake near 100	First Session: It is sad you have a snake near 100, I can see you are feeling <USER EMOTIONAL STATE>. Let's learn a new word. Other Sessions: You look <USER EMOTION>. It is sad that you had a snake near 100 during <SESSION NO> session and you <GAME OUTCOME> the game.
Game Event: Snake Far 100 **Emotional State:** Happy **Event Type:** Positive	User's emotional state, Robot's game state no (of blocks away from robot),	First Session: you are looking <USER EMOTIONAL STATE> and i am happy you are ahead of me. let's learn another word. Other Sessions: you are looking <USER EMOTIONAL STATE> , i am happy to remind you in the <SESSION NO> session you were <NO OF BLOCKS> behind me but you still won the game. let's learn a new word.
Game Event: Ladder near 100 **Emotional State:** Happy **Event Type:** Positive	Occurrence of ladder and number of levels skipped	First Session: You are looking <USER EMOTIONAL STATE>, and i am happy that you are <ROBOT POSITION> of me. Other Sessions: You also had a ladder near 100 in the <SESSION NO> session and i am happy to remind you that you also <GAME OUTCOME> won the game.

Table 5.2: Taxonomy for the exemplar Robot's Behaviour during Game sessions based on Emotion and Memory Model

of robot's positive (encouraging, emphatic), negative (critical) and neutral feedback (control) on the children long-term learning. Therefore, we created robot's behaviour through blending robot's emotional feedback on children learning outcome with the context of robot's emotional memory. In table 5.3 [3], we also show an example of robot's emotional behaviour for the three different types of robot's feedback across sessions on some of the event. In our study, the type of the feedback is the Independent variable (IV) and it refers to the RQ1. The phrases for the three types of robot's feedback on children's learning outcome were constructed on the feedback provided by teachers on different robot's roles in a human-robot learning scenario M. Ahmad et al. (2016c) and in the Chapter III.

5.4.2 Interaction Scenario

We programmed the NAO robot to autonomously play the game with children and teach vocabulary to them, however, speech recognition was controlled via a Wizard of Oz (WoZ) setup as shown in figure 5.5. We implemented a program to reply to basic preconceived questions during introduction and game phases learned from our previous study M. Ahmad, Mubin, & Orlando (2017a). A facilitator responded to the participant's queries through the WoZ setup. The robot stayed quiet if the child asked questions out of its scope.

Our interaction session had four phases: 1) Introduction, 2) Pre-test, 2) game and 4) Post-test.

In the introduction, NAO introduced itself and communicated with the child through a high-level dialogue. The dialogue involved inquiring about their day and the activities that they were undertaking that day. From the second session onwards, the robot addressed the children with their names and also applied contextual information such

[3] Other sessions mean second, third and fourth sessions

Event	Information	NAO's Behaviour (IV - Feedback Type)
Learning Outcome: Correct Word **Emotional State:** Happy **Event Type:** Positive	User's emotional state, word, current and/or past outcome of the word.	First Session: **Positive Condition:** 1) I am **delighted** to know you got this correct — **Gesture:** Happy or Joy **Negative Condition:** I am **Surprised** to know you got this correct. **Gesture:** Surprised **Neutral Condition:** It is **fine** that you got this correct. — **Gesture:** Neutral Other Sessions: **Positive Condition:** I am **happy** that in the <SESSION NO> session, you got it wrong but this time your answer is correct — **Gesture:** Joy, Happy. **Negative Condition:** I am **Surprised** that in the <SESSION NO> session, you got it wrong but this time your answer is correct. — **Gesture:** Surprised. **Neutral Condition:** This is **fine**, in the <SESSION NO> session, you got it wrong but this time your answer is correct. — **Gesture:** Neutral.
Learning Outcome: Incorrect word **Emotional State:** Sad **Event Type:** Negative	User's emotional state, word, no of attempts to memories the word.	First Session: **Positive Condition:** 1) It is **alright**, I am hopeful you can get it right next time. The correct answer is <WORD>. — **Gesture:** Emphatic **Negative Condition:** It is **sad**, that you didn't remember <WORD NAME>. The correct answer is <WORD> — **Gesture:** Sad **Neutral Condition:** It is **fine**. Lets see what happens next time. The correct answer is <WORD>. — **Gesture:** Neutral Other Sessions: **Positive Condition:** It is **alright** that in the <SESSION NO> session, you got it wrong and this time your answer is also incorrect. the correct answer is <WORD>. — **Gesture:** Emphatic **Negative Condition:** It is **sad** that in the <SESSION NO> session, you got it wrong, and this time your answer is also incorrect. the correct answer is <WORD>. — **Gesture:** Sad. **Neutral Condition:** It is **fine** that in the <SESSION NO> session, you got it wrong, and this time your answer is also incorrect. the correct answer is <WORD>. — **Gesture:** Neutral

Table 5.3: Taxonomy for the exemplar Robot's Behaviour during Post-test sessions based on Emotion and Memory Model

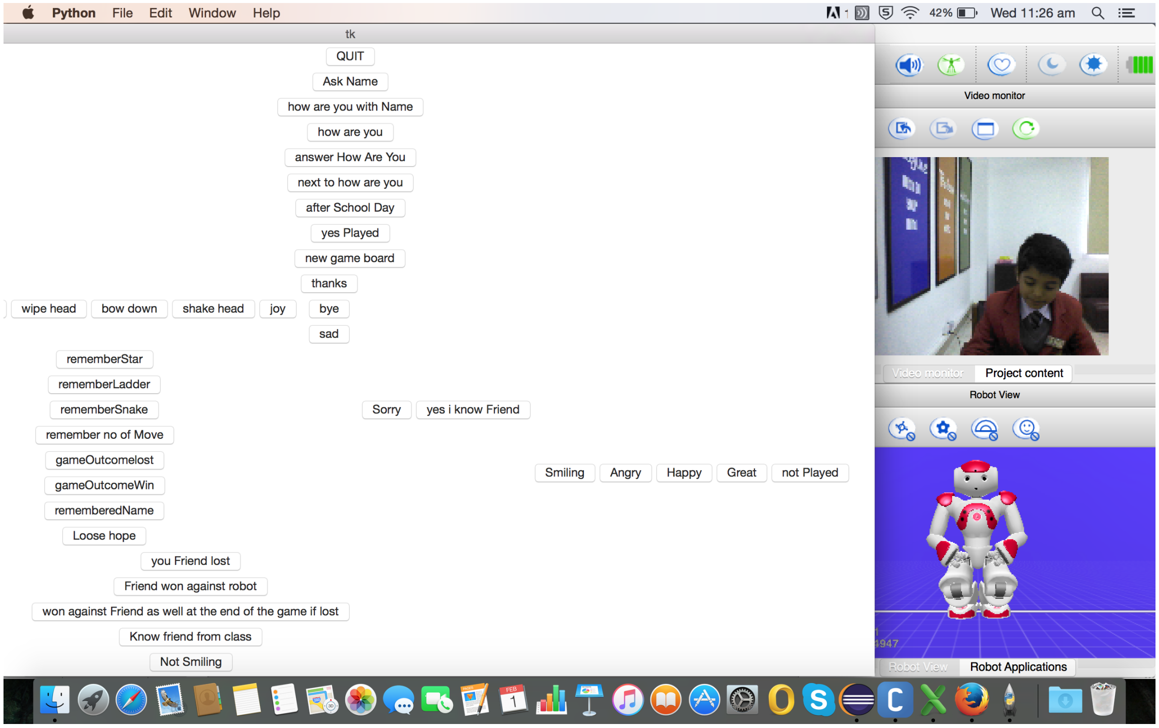

Figure 5.5: A Wizard of OZ to control NAO's Speech Recognition.

as game outcome from the last sessions.

In the Pre-test, NAO robot asked about the words to be taught during the game. In the first session, NAO asked about the words with an assumption that children didn't know the word. The robot asked "Do you know the word "Jabami"?", as the child didn't know the word, the children replied "No". The NAO robot later responded "We will learn about the <WORD NAME> shortly". From the second to the fourth session, the robot provided feedback on the previously taught words. In an event of a mistake, the robot informed the child about the correct word for the previous used words. For instance; the robot said this is incorrect and informed the child about the correct answer.

In the Game phase, NAO robot played a snakes and ladders game with the child. During the game, the child was taught six different words in each session. We coded a fixed/pre-determined pattern of turns (this means that the throws of the dice were controlled) for both child and robot during the game for every session for all the participants. The afore-explained model was applied during the gameplay to create the memory of emotional events during each session of the game. This memory was utilised after the first session.

In the Post-test, NAO tested about the words taught during the game. The NAO robot provided positive, negative and neutral feedback through applying our model. The feedback combined gestures and dialogue on the learning performance of each child. In table 5.3, we present examples of the NAO robot's behaviour for the three categories of feedback during the post-test on children learning performances. For clarification, the robot only applied the model in the post-test and one of the feedback was chosen for one group of users.

5.4.3 Setup and Materials

We were assigned a quiet room at the school that was divided into two parts with a divider as shown in figure 6.2. On the left side, one of the researchers was controlling the speech recognition capabilities of the robot. On the right side, the child interacted with the NAO robot placed on the table along with a Samsung tablet. We used the NAO robot designed and developed by Aldebaran robotics. It is a humanoid robot measuring 58 cm in height with 25 degrees of freedom.

We used 24 vocabulary words from the Robot Interaction Language (ROILA) taken from the first two chapters of the book on ROILA (Mubin, 2011). The list of the words can be found in table B.1.

5.4.4 Participants

We conducted our between-subject study with 24 children (12-males, 12 females) aged between 10-12 at a school. The mean (M) and standard deviation (SD) of the ages were M: 10.52 and SD: 0.60 respectively. All of the participants were bilinguals.

Figure 5.6: Setup (Left); A Child playing snakes and ladders with NAO (Right).

None of the participants had previously interacted with a robot.

The between-subject factor was the 3 x feedback type of the robot during the post-testing phase. Each group comprised 8 children with equal ratio for the gender. The participants were assigned randomly to each group.

5.4.5 Procedure

Our study was setup as a long-term between-subject evaluation that spanned for two weeks. The study was conducted individually with one child at a time. Each child played the snakes and ladders game with the NAO robot 4 times for 4 days (one session per day) over the course of two school weeks with a gap of two days during sessions, for a total of 96 sessions (24 child * 4 sessions). Each group of children attempted the post-test on a tablet in one of the three conditions (robot's displaying positive, negative, neutral emotional expression on child's learning outcome) for two school weeks. We conducted our sessions on the 1st and 4th days of the school weeks for two weeks respectively. Each session lasted for approximately 24 minutes and had five steps: 1) a 2-minute introduction, 2) a 4-minute pre-test, 3) a 10-minute game playing session, 4) a 4-minute break, and 5) a 4-minute post-test. The facilitator used a stopwatch to maintain time consistency throughout the sessions.

Pre-test: Following an introduction as explained in the interaction scenario, the robot initiated the session through asking about unknown words from the ROILA language displayed on a tablet device as shown in figure 5.7. The robot asked the meaning of six words during the first session. The rationale for selecting 6 words in each session came from a pilot study conducted with 5 participants, also aged between 10-12 years. The procedure and robot's feedback mechanism in the pilot study was similar to the first session of our present study. Our findings showed that on-average children were able to remember 3 out of 5 words (Mean: 3.6 and Standard Deviation (SD): 1.14). Therefore, we selected six "new" words per session. In each session, six new words were added to the test. Therefore, 6, 12, 18 and 24 words were tested for in the first, second, third and fourth session respectively. The pre-test was an auditory-visual word identification task McCullough et al. (1992) as shown in figure 5.7. The visual used in the test to represent a word was identical to the one used in the game playing sessions.

Game Play: Each child played the snakes and ladders game following a pre-defined pattern of dice outcomes for four times. Each child faced a snake inside the game six different times. On each snake, a new word was taught to the child. Therefore, in each session, six new words were taught to the participants. The child gets exposure to one word only once during the game.

Post-Test: After a 4-minute break,the child participated in the post-test to determine the accuracy of words learned during the session. The same procedure as the pre-test was repeated in this phase. The post-test was identical to the pre-test, containing the same words. We chose identical words for both pre and post-test to maintain the consistency of test results. More precisely, 6, 12, 18 and 24 words were tested for in the first, second, third and fourth session respectively. All of the test results and mistakes were logged in the database.

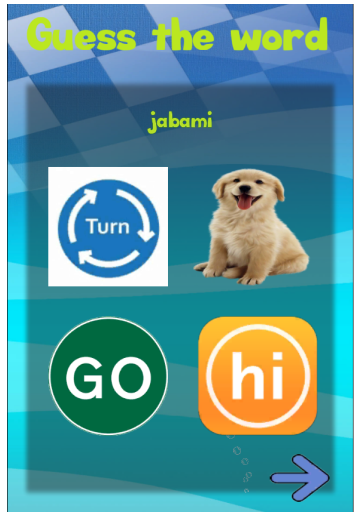

Figure 5.7: A Word "Jabami" is asked and four options are displayed on the screen.

5.4.6 Measurements

To measure the effects of our model towards promoting vocabulary learning and sustaining engagement during long-term HRI, we looked the following Dependent variables (DVs):

1. Children's immediate retention of new words during the session (total number of words remembered in the post-test of every session).

2. Children's social engagement across sessions measured through individually calculating the total duration for eye-gaze facing robot, smiles, verbal response, and gestures during each session of the interaction. In essence, each of the afore-listed variable was reflected sign of social engagement.

To measure the effects of different robot's emotional feedback on the memorisation of vocabulary and social engagement across three conditions during long-term HRI, we looked at the following DVs:

1. Total number of words remembered during all the sessions (total number of words remembered in the last post-test on the last day).

2. Retention of old words across sessions (total number of words remembered during the pre-test taught in the previous sessions).

3. Social engagement across sessions measured through individually calculating the total duration for eye-gaze facing robot, smiles, verbal response, and gestures during each session of the interaction.

To measure social engagement, we conducted a video analysis to code following DVs as shown in table 5.4. We followed the same coding scheme as reported in the chapter IV to measure social engagement. We conducted video analysis of all the four interaction sessions. A total of 96 videos were analysed and the duration of the videos ranged between 20 to 24 minutes. As the number of words increased with every session, therefore, the duration of pre-test and post-test were also increasing. We divided our sessions in four time intervals: introduction, pre-test, game-play and post-test.

Following the coding scheme as reported in the chapter IV, we coded videos for following dependent variables: Gaze, Verbal Interaction, Facial Expressions, and Gestures. Two researchers were involved in video coding process. Both of the researchers did not take part in the evaluation process. The second researcher coded 20% of the videos separately and discrepancies were resolved with consultation. The first researcher then completed the coding. We coded for both duration for the four different factors. We also witness similar examples in literature (Bartneck et al., 2007), who has also coded for duration to measure social interaction. The coding mechanism as shown in Table 5.4 was followed.

The justification for using gaze, facial expressions, verbal responses and gestures is grounded in the literature (Argyle & Dean, 1965; Castellano et al., 2009; C. L. Sidner et al., 2005) as all of these DVs have been considered as a sign of social engagement during cHRI.

Gaze	Facial Expression	Verbal Response	Gesture
Robot face	timid Smiles	“Hello”	Wave
	Surprised	“thank you”	fist
	Flushed	“Okay” ,“Yes”	Nod

Table 5.4: Coding Scheme used to measure social engagement (Ahmad et al., 2017).

5.5 Results

5.5.1 Emotion and Memory Model Results

In this section, we present finding of how children engagement shaped during the game phase to find answers to the **RQ2** and **RQ3**. We also present our findings on the effect of children immediate retention of words taught during the game session.

5.5.1.1 Social Engagement

To find answer to **RQ2** and to test **H2**, we conducted a repeated measure ANOVA with the session as the within-subjects factor with four levels on the game play phase for the following Dependent Variables (DVs): children's gaze facing the robot, facial expressions (smiles), verbal response, and gestures. It is also important to note there was no difference in the quantity of the interaction during the game phase, therefore, we don't find a need to normalize our results. The quantity of interaction refers to the total duration of interaction during the game across all sessions. It is to note that all duration reported in the result were in *msecs*.

Our findings showed that during game phase there was a significant effect of session on the children's facial expressions (smiles) (F(3, 69) = 4.48, p = .007, η_p^2 = .244), and verbal responses (F(3, 69) = 7.40, p = .000, η_p^2 = .163). This suggests that the values for the duration of children's facial expressions and verbal responses significantly changes between sessions. Additionally, there was no effect of session on children's gazes facing the robot (F(3, 69) = 1.45, p = .235, η_p^2 = .059) and children's gestures (F(3, 69) = .146, p = 0.932, η_p^2 = .006). This suggests that the values for the duration of children's gazes facing the robot and gestures remained consistent throughout the four sessions.

To understand how values for the duration of all four DVs changed between sessions, we also conducted Bonferroni test to further examine the significance between sessions for all the DVs. In the case of children's gazes facing the robot, we didn't find significant effect between sessions. The mean values for the duration of children's gaze facing the robot were as follows: Session 1 (M: 46.08, SD: 13.80), Session 2 (M: 53.87, SD: 23.20), Session 3 (M: 48.70, SD: 20.80), and Session 4 (M: 45.25, SD: 24.18). We observed consistent values for the social engagement measured in terms of the duration of childrens' gazes facing the robot from first to the fourth session.

For the children's facial expressions, the first session was statistically significant in comparison with the second ($p = .016$), fourth ($p = .042$) session. The mean values for the duration of children's facial expressions were as follows: Session 1 (M: 79.24, SD: 69.31), Session 2 (M: 43.94, SD: 45.22), Session 3 (M: 62.47, SD: 54.18), and Session 4 (M: 44.48, SD: 34.36). This suggests that we witnessed a decline in the duration of children's facial expressions from first to the second and fourth session.

In the case of the children's verbal responses, we found that the third session was significant in comparison with first ($p < .012$) session and fourth session was significant in comparison with the first ($p < .007$) session.The mean values for the duration of children's verbal responses were as follows: Session 1 (M: 19.42, SD: 21.00), Session 2 (M: 29.43, SD: 24.34), Session 3 (M: 47.57, SD: 34.53), and Session 4 (M: 35.67, SD: 18.35). This suggests that we witnessed an upward trend for the social engagement measured in terms of the duration of childrens' verbal responses from the first to the fourth session.

Lastly, for gestures, we didn't observe significance between sessions. The mean values for the duration of children's gestures were as follows: Session 1 (M: 9.35, SD: 22.61), Session 2 (M: 9.44, SD: 12.18), Session 3 (M: 9.85, SD: 15.82), and Session 4

(M: 7.45, SD: 13.22). This suggested that we witnessed consistent amount of social engagement measured in terms of the duration of gestures from first to the fourth session. Our results based on the four DVs suggested that although the duration of eye-gaze and gestures were consistent throughout all the session but the value for both facial expressions and verbal responses varied between session. Therefore, we suggest that H2 was partially accepted.

5.5.1.2 Vocabulary Learning

We conducted Kolmogorov − Smirnov test to ensure that the generated data was normally distributed before conducting Analysis of Variance (ANOVA). The results showed that the data was normally distributed for the learning outcomes of the children.

To find an answer on the **RQ3** and also to test **H3**, we checked for the immediate retention of words learnt during each gameplay in all the sessions for all the participants. The immediate retention refers to the six words taught during each game. We conducted a repeated measure ANOVA with the *session* as the within-subjects factor with four levels using immediate retention of words learnt per session as a Dependent Variable (DV). Results showed that there was a significant effect of session (F(3, 69) = 8.46, p = .000, η_p^2 = .269) on children's vocabulary learning performance. We executed Bonferroni posthoc to further examine the effect on the immediate retention of the words learned within sessions. This suggests that we witnessed that the amount of words retained during the third session were significantly higher in comparison with the first (p = .002) and fourth session (p = .007). The mean values of the learning outcome for all the sessions are as follows: Session 1 (M: 4.75, SD: 0.98), Session 2 (M: 5.37, SD: 0.76), Session 3 (M: 5.75, SD: 0.60), and Session 4 (M: 5, SD: 0.65). Our learning performance results with respect to the immediate retention of words learned during sessions were encouraging as children were on average able to learn four to five out of six words from the NAO robot capable of generating responses through following our emotion and memory model. As the memory account of the participants was applied from the second session and we did witness a significant increase in the mean value of the retention of words across the third session, therefore, we partially accept our H3.

5.5.2 Effects of Robot's Emotional Feedback

In this section, we report on the effects of different emotional feedback of the robot on children's social engagement and vocabulary learning as specified earlier during the post-test.

5.5.2.1 Social Engagement

To find an answer to **RQ1b** and test **H1b**, We conducted a repeated measure ANOVA with session as the within-subjects factor having $four$ levels and robot's emotional feedback type as the between-subject factor. We used children's gazes facing robot, facial expressions (smiles), verbal responses, and gestures during first, second, third and fourth session as the DVs.

We found a significant effect of the robot's emotional feedback during the post-tests on the duration of children's gazes facing robot (F(2, 21) = 4.77, p = .019, η_p^2 = .313). We didn't find significant effect of robot's emotional feedback on the duration of facial expressions (F(2, 21) = 1.23, p = .312, η_p^2 = .105), verbal responses (F(2, 21) = .516, p = .604, η_p^2 = .047) and gestures (F(2, 21) = .906, p = .419, η_p^2 = .079).

We also find an effect of session (interaction effect) for each emotional feedback type on the duration of children's gazes facing the robot (F(2, 21) = 4.77, p = .016, η_p^2 = .325). We didn't find a significant effect for the duration of facial expressions (F(2, 21) = 4.77, p = .105, η_p^2 = .231), verbal responses (F(2, 21) = .917, p = .500, η_p^2 = .126) and gestures (F(3, 19) = .625, p = .709, η_p^2 = .091). The mean-values are shown in Table 5.5.

We also conducted the Bonferroni post-hocs check to further examine the significance of the duration of children's gazes facing the robot for the robot's emotional feedback during the post-test. We found that the duration of children's gazes facing the robot during the positive emotional feedback given by the robot was significantly higher (p = .018) in comparison to the neutral feedback. We didn't find significant difference of negative feedback in comparison to the neutral feedback. In other words, positive feedback resulted in highest duration followed by the negative and neutral feedback. In summary, this suggests that children were engaged in terms of their gazes facing robot during the positive feedback condition, therefore, our hypothesis H1b was partially accepted in terms of children's eye-gaze facing the robot.

5.5.2.2 Vocabulary Learning

To find an answer to **RQ1**, we conducted a one-way between-subject ANOVA with robot emotional feedback type as the independent variable (IV) and using a total number of words learned during all the sessions as a DV. The purpose of our analysis was to measure the overall effect of the type of emotional feedback on child's overall vocabulary learning. Our results show that there was a significant effect (F(2, 19) = 5.7 $p = 0.011$, $\eta_p^2 = .281$) of the type of robot's emotional feedback on the child's learning outcome. The mean retention rate across conditions were as follows: Positive Condition) **M:** 22.25, **S.D.:** 1.03510, Negative Condition) **M:** 20.25, **S.D.:** 1.16496 and Neutral Condition) **M:** 20.50, **S.D.:** 1.60357. We performed a Bonferroni posthoc check to further examine this significant difference. We found that robot with a positive emotional feedback had the better effect on child's overall learning outcome as compared with negative ($p = 0.016$) and neutral feedback ($p = 0.039$).

To check the effect of the robot's feedback on children's learning of words across sessions, mainly to test **H1b** or to find an answer to **RQ1a**, we conducted a repeated

DV	**Condition**	**Session 1**	**Session 2**	**Session 3**	**Session 4**
Eye Gaze	Positive	M: 26.59, SD: 21.02	M: 38.39, SD: 4.04	M: 28.38, SD: 11.92	M: 31.10, SD: 9.99
	Negative	M: 27.58, SD: 4.54	M: 37.97, SD: 9.83	M: 21.16, SD: 2.90	M: 26.90, SD: 3.00
	Neutral	M: 25.87, SD: 7.23	M: 22.40, SD: 2.99	M: 22.92, SD: 4.2	M: 24.15, SD: 5.04
Facial expressions	Positive	M: 16.43, SD: 11.75	M: 22.11, SD: 19.06	M:6.43, SD: 4.14	M: 5.20, SD: 3.46
	Negative	M: 22.33 SD: 18.93	M: 11.46, SD: 10.05	M: 8.49, SD: 9.17	M: 9.11, SD: 10.14
	Neutral	M: 9.19, SD: 7.72	M: 6.25, SD: 9.71	M: 6.57, SD: 7.93	M: 6.67, SD: 6.66
Verbal Responses	Positive	M: .97 SD: 1.23	M: 1.95, SD: 2.94	M: 2.26, SD: 3.45	M:1.39, SD: 2.30
	Negative	M: 3.14, SD: 3.15	M: 2.09, SD: 2.05	M: 2.47, SD: 2.23	M: 1.09, SD: 1.49
	Neutral	M: 2.91, SD: 3.40	M: 3.01, SD: 2.93	M: 1.82, SD: 1.63	M: 2.02, SD: 1.72
Gestures	Positive	M:1.12, SD: 1.24	M: 2.11, SD: 2.48	M: .95, SD: 1.60	M: 1.09, SD: 1.24
	Negative	M: .88, SD: .98	M: 1.37, SD: 1.85	M: 2.4, SD: 3.46	M: 1.32, SD: 3.11
	Neutral	M: 1.76, SD: 1.61	M: 2.17, SD: 1.92	M: 2.7, SD: 3.86	M: 2.60, SD: 2.31

Table 5.5: Mean and Standard Deviation for the duration of eye-gaze, facial expressions, verbal responses and gestures across first, second, third and fourth sessions for each type of robot's emotional feedback.

Condition	Session 2	Session 3	Session 4
Positive	M: 5.37, SD: .74	M: 5.37, SD: 1.06	M: 5.25, SD: .70
Negative	M: 4.37, SD: .74	M: 4.50, SD: 1.06	M: 4.5, SD: .75
Neutral	M: 4.62, SD: 1.06	M: 5.00, SD: .92	M: 4.5, SD: .92

Table 5.6: Mean and Standard Deviation for the retention of words during Session 1 words across second, third and fourth sessions.

measure ANOVA with the session as the within-subjects factor with *three*, *two* levels and type of emotional feedback as the between-subject factor using the retention of words learned during the first and second session across remaining interaction sessions as the DV. In essence, we conducted three separate ANOVAs. We found a significant effect (F(2,21) = 4.10; p = .031, η_p^2 = .281) of robot's emotional feedback on the retention of words learned during the first session. The Bonferroni posthoc check showed that positive emotional feedback from the robot has a significant effect on child's retention of words as compared with negative feedback ($p = .033$). The mean values are shown in Table 5.6.

We also found a significant effect (F(2,21)= 4.882; $p = 0.018$, $\eta_p^2 = .317$) of robot's emotional feedback on the retention of words remembered during the second session. The Bonferroni posthoc check showed that robot with a positive emotional feedback has the significant effect on child's overall learning outcome ($p = .023$) as compared to the negative feedback. The mean values are shown in Table 5.7.

For the word remembered during the third session, we conducted a one-way between-subject analysis of variance (ANOVA) with robot behaviour type as the IV and using a total number of words retained during the fourth session as DV. We ran a new test as the results for the words learned during the third session were only checked once in the fourth session. Our results showed that there was a significant effect (F(2,21)=5.96, $p = 0.009$, $\eta_p^2 = .317$) of robot's emotional expression on the child's learning outcome.The Bonferroni posthoc check showed that robot with both positive and negative emotional feedback has the significant effect ($p < 0.03$) on child's overall

Condition	Session 3	Session 4
Positive	M: 5.62, SD: .51	M: 5.12, SD: .64
Negative	M: 4.85, SD: 1.12	M: 4.00, SD: .75
Neutral	M: 5.37, SD: .74	M: 3.87, SD: 1.24

Table 5.7: Mean and Standard Deviation for the retention of words during Session 2 words across third and fourth sessions.

Condition	Session 4
Positive	M: 5.50, SD: .53
Negative	M: 5.50, SD: .53
Neutral	M: 4.37, SD: 1.06

Table 5.8: Mean and Standard Deviation for the retention of words during Session 3 words across the fourth sessions.

learning outcome as compared to the neutral feedback. The mean values are shown in Table 5.8.

In summary, our results suggested that overall, a robot reacting positively to the childs vocabulary learning outcome resulted in better children vocabulary retention followed by the negative and neutral feedback across sessions. Therefore, our hypothesis H1b was also accepted.

5.6 General Discussion

In this sections, we discuss on the main and significant results as presented in the results section.

5.6.1 Discussion on Emotion and Memory Model Results

We conjecture that the reason for the consistency of the duration of the children gazes (i.e. no resulting saturation) facing the robot during the game playing sessions may have been due to the utilisation of the emotional memory from the second session. Our emotion and memory model generated the behaviour for the robot through applying the memory of emotional events of each child stored and generated during the past game sessions. In other words, our model enabled the robot to adapt its behaviour to the user characteristics. These attributes may have enabled the robot to sustain the children's social engagement in terms of their gazes facing the robot. Our results also further strengthen observations reported by several researchers that a robot adapting its behaviour can result in the long-term use of robots during educational settings (Mubin et al., 2013). Additionally, an increase in the gazes facing the robot during the second session may have been due to the augmentation of the context of the memory of emotional events in robot's response from the second session. This use of memory as a part of robot's behaviour resulted in novel robot's responses and it may have enhanced the duration of gazes facing the robot. As children may have became accustomed to the robot use of memory in its responses during the third and fourth session, we conjecture that it may have resulted in the consistency in terms of the duration of children gazes facing robot in the rest of the sessions. On the contrary, it can be believed that the consistency of the duration of gazes may be due to indifference of the application of the model however, in our previous study (M. Ahmad, Mubin, & Orlando, 2017a), the gazes facing robot significantly declined from the third session in the condition where the robot didn't adapt its comments to user game events during the same snakes and ladders game.

We speculate that the consistent duration of facial expressions, mainly smiles, may have been due to robot reminiscing about emotional events. However, a decline in the duration of smiles in the second, an increase in the third and then slight decrease in the fourth sessions may be due to several reasons. One of the reasons for the decline in the second session may be again due to the loss of the wow factor as all of these participants had previously interacted with the robot for the first time (Sung et al., 2009). We believe, during the first session, the children may have showed more expressions due to the robot's cuteness (M. Ahmad, Mubin, & Orlando, 2016a; Mubin, Khan, & Obaid, 2016) and wow factor (Komatsubara et al., 2014). During the second session, the excitement level of children may have normalised. However, an increase in the duration of the children's smiles during the third session might have been due to more favourable game outcome and subsequent robot's behaviour. Additionally, the use of memory in robot's behaviour may have also created an element of personalisation and as in one of the previous studies, it was shown that a personalised behaviour of the robot generated more smiles in comparison with a non-personalised behaviour (M. K. Lee et al., 2012). Another reason can be the use of words during each session. During the game play, as the child landed on a snake, a new word was taught to the participant, we also understand that some words may have been more relatable than others (even though words were randomly chosen) as children associate to words while hearing and remembering them (Miller & Gildea, 1987).

We believe that the increase in the duration of verbal response can be due to the feedback provided by the robot; this has been shown in the recent exploratory study that providing an appropriate feedback enhance child-robot tutoring interactions (de Haas et al., 2016). As our model enabled a robot to provide context-aware feedback based on emotional experiences, it might have resulted in the increase in the verbal responses. Additionally, our robot also reacted positively through providing encouragement as well as negatively through acting critically based on emotional events, it may have also resulted in an increase in verbal responses. It is also reported in the literature that emphatic dialogue sustains engagement (Leite et al., 2014). Furthermore, children were extremely involved in the process of learning words, in most cases, children repeated the word after the robot and also in return asked questions to the robot. For instance; The robot said; "the meaning of the word 'Fupama' is music, do you like the music", upon which the child replied; "I don't listen to music, do you?". Many similar patterns were witnessed throughout the sessions. The slight decline in the last session may have been due to the less association with the words used during the interaction.

We infer that the children were also competing against the NAO robot when it comes to playing the game, as we witnessed a number of children showcasing pumping "fist" gestures. However, it was also observed that some children were showing fewer gestures as compared to others. It could have been due to different personalities of the participants, as some children are less expressive than others (Lever, 1976). It may also be because some children reacted strongly to game events that took them close to winning the game than the others. It could be due to the finding that some children are also more competitive than the others (Van Lange et al., 1997; Benenson et al., 2007; McClintock & Nuttin Jr, 1969).

Our emotion and memory model for the NAO robot resulted in partially supporting our hypothesis as children were able to retain most of the taught words during each of the sessions. However, we witnessed a difference in children's retention of words learnt during the third session in comparison to first and fourth sessions. We conjecture that the reason might be due to the higher engagement during the third session because the relationship of engagement and learning outcome have been specified in literature (Carini et al., 2006; Sharan & Tan, 2008) and higher engagement do result in better performance (Carini et al., 2006). This relationship can also be witnessed in our results on children's social engagement as the duration of children verbal responses are higher in comparison to the first and fourth session. Therefore, a slight decline in the last session on the learning performance can also be grounded in the varying levels of children's social engagement during the interaction. It is also reported that children's learning can be affected if there is a fall in their interest level (Kanda et al., 2004). Additionally, due to the WoW factor (as the children interacted with the robot) during the first session, the children may have been more focused on the robot rather than on the task as such findings have also been previously reported (Rosenthal-von der Pütten et al., 2016).

5.6.2 Discussion on Effects of Emotional Feedback Results

We found that in the condition where robot provided positive emotional feedback on children's learning outcome also resulted in highest duration of eye-gazes facing the robot. The preference on the choice of the type of emotional feedback on children's learning outcome may be dependent on the individual differences between the children as also indicated by (Haas et al., 2017). We also believe that positive emotional feedback of the robot may have motivated children and may have also resulted in a positive emotional state in the child. On the other hand, negative feedback may have resulted in activation of negative emotional state in children (Feldman Barrett & Russell, 1998). Due to this activation of positive state, a child might have looked at the robot due to the feeling of being relaxed and calm. While, in the negative emotional state, the child may have been disappointed or sad and as a result shed away from looking at the robot (Pekrun & Linnenbrink-Garcia, 2012). We would also like to highlight that although we tried to make the length of the positive, negative and neutral feedback given by the robot consistent throughout the evaluation however, we understand that this could also be due to the length of the feedback may have effected the overall affects.

We found that the positive emotional feedback provided by the robot on the child's vocabulary learning performance did have a significant effect on the children's retention of vocabulary. Our hypothesis was accepted as children were able to retain the most number of words during robot's positive emotional feedback condition. In addition, the neutral emotional feedback on child's performance was also found to have an influence on child's learning. Moreover, the negative emotions were least regarded in terms of children's long-term learning performance. We conjecture that when the robot positively empathizes with the child during the interaction, it creates a positive affect on the child's development and progress in general. In the past, the role of emphatic robotic behaviour has been appreciated during a playful interaction as it was able to sustain children's interest in a long-term child-robot interaction (Leite et al., 2014). We also speculate that the reason children were able to better retain words during positive emotional feedback is due to the positive emotional state of the children in response to the robot's reaction. It is shown in literature that humans store different kinds of information during different emotional states (Levine & Pizarro, 2004). In addition, positive emotional feedback is conducive to feeling confident and successful in the learning process (Hattie & Timperley, 2007), as it facilitates enhanced learning. Therefore, we believe that a positive reaction of a robot created a positive emotional state of the child. It in return, made the child perform better as compared with the negative or neutral reaction of the robot. One of our findings also showed that from the third to fourth session, the negative feedback was preferred over neutral feedback. We speculate that although positive feedback is highly desirable, but negative feedback is also needed during the learning process for motivation. As it is shown in previous literature that negative feedback in terms of criticism may positively encourage student engagement and attention on learning task (Wentzel, 2002). Therefore, negative feedback or directive critique can be useful in certain situations. In summary, it indicated towards implementing and evaluating robotic tutoring applications where robot's reacts positively on children's learning performance and may also react negatively through providing gentle criticism at times during HRI for the better learning experience. Such all round capabilities have also been indicated by school teachers in one of our past studies (M. Ahmad et al., 2016c).

5.7 Conclusion

In this paper, we presented an emotion and memory model for a social robot capable of creating emotional memories and adapting accordingly. We performed an exploratory study to evaluate the applicability of our model in a vocabulary learning task at a school during a children-game-robot interaction. The preliminary results from the evaluation of the emotion and memory model resulted in positive findings based on child's vocabulary learning and also sustaining social engagement during all the sessions. In particular, we had the following findings:

- Social engagement of the participants measured in terms of the duration of the eye-gazes facing the robot and gestures towards the robot remained consistent throughout all the four sessions.
- We witnessed an upward trend for the social engagement measured in terms of the duration of childrens' verbal responses.
- On the contrary, we witnessed a decline in the duration of children's facial expressions from first to the second and fourth session.

We also explored the effect of three different type of robot's emotional response on children learning performance during long-term cHRI. Our preliminary results showed that in a condition, where robot provided positive/supportive emotional feedback or response had the better effect on the child's learning performance and engagement in terms of the duration of children's eye gazes as compared to the negative and neutral condition. Our preliminary results highlighted towards the exploration of more research on testing intelligent robotic tutoring systems where robot should play positive, emotionally supportive roles.

5.8 Limitations

One may argue on the number of sessions in an experiment before it comes to being categorised as "long-term", however, our selection of a number of sessions is based on the findings of our previous study (M. Ahmad, Mubin, & Orlando, 2017a), where children's novelty factor diminished from the third session. The novelty was measured using the same social engagement metrics identified in this chapter.

We didn't compare our results based on the effects of our emotion and memory model on the social engagement with a control condition, where no memory was incorporated during the interaction. However, the reason for not having a control was based on the results of our previous study (M. Ahmad, Mubin, & Orlando, 2017a), where we didn't have the user modelling and we found that the social engagement declined from the second session during a similar experimental setup. Additionally, we find examples of literature in the past where model was evaluated without a control condition (Leite et al., 2014). Additionally, one of the guidelines of the long-term studies emphasised that the control condition should only be evaluated under extreme conditions as it usually duplicates the effort to analyse the data and also user's experience over time becomes a strong independent variable (Leite, 2013a).

We also had a technical limitation with respect to speech recognition through the wizard of oz (WoZ), in case, the child asked questions that were out of robot's scope, the robot stayed silent.

We also understand that some participants would speak more than one language and/or have a greater level of capacity for learning another language. Our study did not select children on this basis.

Our results of vacabulary learning were only based on the recognition of words and we didn't test for the pronunciation of words.

CHAPTER VI

Automating the behaviour selection for the Emotion and Memory Model

In this chapter[1], we present a mechanism for a social robot to determine the selection of an action from one of the three feedback (positive, negative and neutral) based on the learning performance and social engagement of the users. We augmented our model's feedback as described in Chapter V through creating classification of positive, negative and neutral behaviours on the occurrence of each event during the game. Our rationale for using this mechanism was based on the findings with respect to the effects of three different emotional feedback on children's vocabulary retention performance and maintaining social engagement in a long-term HRI in chapter V.

6.1 Introduction

"Reinforcement learning (RL) is the problem faced by an agent that must learn behaviour through trial-and-error interactions with a dynamic environment" (Kaelbling et al., 1996). In essence, an agent (a robot) learns through interacting within an environment after receiving rewards on its actions.

We witness the applications of RL in the domain of Human Computer Interaction,

[1]A Journal article has been submitted at one of the reputable journals in the area of Human Computer/Robot Interaction

especially in the education technology, particularly to design and implement intelligent tutoring systems. For instance, Sottilare et al. (2012) presented a *Generalized Intelligent Framework for Tutoring* (GIFT) that enables the educational technologies to utilise on RL to learn about the appropriate strategy for each user based on the users state to achieve maximum learning gains. We also find examples of the use of the *GIFT* framework in the design of different educational technologies (Kelsey et al., 2015). Other researchers have also promoted the use of RL in the design of the educational technology. For instance, an adaptive educational system was designed that adapts the teaching style through modelling user's learning style (Dorça, 2015). Similarly, the use of RL in the field of HRI and Social robotics to enable the robots to learn about the selection of appropriate actions has been a growing phenomenon in the field of HRI. In particular, the use of simple algorithms to implement personalisation or adaptation has greatly improved the effectiveness of robot's tutoring in comparison to non-personalised way Leyzberg et al. (2014). Leite (2015) proposed a non-intrusive RL approach in which the robot can learn and adapt to the user based on different user preferences in real time. They also applied this technique in one of their studies where they enabled the robot to choose from a set of supportive behaviours during a playful interaction (Leite et al., 2014). In addition, Castro-González et al. (2011) applied a RL algorithm based on q-learning to enable the Maggie robot to select the right action for every state in order to maximise user motivation. Moreover, Ritschel & André (2017) have proposed to use RL algorithm to inform robot's personality adaptation in real-time based on social signals or cues of the user.

We also find social robot's applications in the education domain, where different RL algorithms have been applied to implement adaptive behaviour selection informing personalisation during Human-Robot Interaction across various domains such as Education. For instance: Gordon et al. (2016) also used an affective RL algorithm based on q-learning to determine social robot's verbal and non-verbal behaviours during language learning game to promote affective personalisation. Additionally, Ramachandran & Scassellati (2015) also applied contextual bandit algorithm that can adaptively control the pace of the interaction based on users performances and affective feedback. Ramachandran & Scassellati (2014) also implemented personalisation in robots using RL based on the learning difficulty levels of the individuals. Moreover, Gao et al. (2017) proposed a RL framework that enabled the robot to select the supportive behaviour to maximise task performance in a game-based learning scenario. Furthermore, we witness an increasing trend of applying different learning based mechanism in the field of HRI (Jones, Bull, & Castellano, 2017; Jones & Castellano, 2018) and similarly, RL models are applied to inform better robotic tutors (Roy et al., 2018). The findings from all of these studies have resulted in influencing positive attitudes in individuals and has also promoted their task performance or improved their learning.

Considering the positive findings in terms of improving learning through the use of RL algorithms to promote personalisation during social HRI, we also attempted to apply a RL algorithm to automate the process of decision making for the selection of social robot's behaviour based on our Emotion and Memory Model. Similarly, also as one of the findings mentioned in Chapter V highlighted the benefits of both positive and negative feedback during the child-robot tutoring interaction in a long-term setup. Therefore, we also conjecture that modelling the robot's feedback based on the user characteristics would result in promoting user learning performance. The rationale for this assumption is also based on the finding as highlighted in education literature that the process of user modelling to inform teaching strategies improve individual's learning (Haider, Sinha, & Chaudhary, 2010; T.-C. Liu, Graf, et al., 2009). Lastly, we understand that implementing the RL based algorithm for behaviour selection in the HRI scenario would also result in other benefits such as autonomy, lifelike, scalability and wider applications (Breazeal, 2003).

6.2 Research Method

Our research further tried to study the effects of the adaptive action selection during our emotion and memory model towards sustaining social engagement and vocabulary retention during a long-term interaction. Keeping this in mind, we tried to answer the following RQs:

RQ1 - What are the effects of a social robot selecting behaviour based on the adaptive strategy implemented through RL on the social engagement?

RQ2 - What are the effects of a social robot selecting behaviour based on the adaptive strategy implemented through RL on the children's immediate vocabulary retention?

RQ3 - What are the effects of a social robot selecting behaviour based on the adaptive strategy implemented through RL on the children's delayed vocabulary retention?

We hypothesize that a robot learning based on a mechanism would sustain social engagement and will also maintain the retention of vocabulary during a long-term interaction. To find answers to this question, in this study, we applied the RL algorithm to learn about a behaviour based on the user social engagement during the game and compared it with the findings as presented in the Chapter V, where the robot's behaviour was pre-defined and was based on the type of the event during the game. We didn't choose to have a control group for this study because it seemed logical to rather simply compare the finding against chapter V due to exactly similar nature of the study. The only difference between the two studies lies in the process of the selection of behaviours for the robot. In Chapter V, we choose positive, negative and neutral behaviour based on the type of the event in a pre-defined mechanism where as in the present study, we inform behaviour selection through RL algorithm.

6.2.1 System Description

We tried to refine our emotion and memory model through automating the behaviour selection mechanism as shown in figure 6.1. The behaviour selection Unit (BSU) described earlier in Chapter V enabled the robot to react positively, negatively and neutrally based on the type of event during the game. However, we updated the behaviour selection unit through using a RL based algorithm to enable the robot to select one of three behaviours (positive, negative, and neutral). We created three aforementioned categories of robot's behaviour during the game. On each event in the game, the robot selected one of the three behaviours from the classification of the behaviours based on the social engagement of the child. The rationale for the classification of the behaviours was due to the effect of the social role of the robot (competitive/cooperative) on the task performance and engagement of a user (Zaga et al., 2015). Similarly, children may have a different preference for the robots during an educational setup. Additionally, Mutlu et al. (2006) has also conducted an exploratory study on the perception of a humanoid robot possessing both Co-operative and competitive characteristics. Their results revealed that people perceived robots as significantly more desirable in a cooperative role than in a competitive role. However, it was further mentioned that the preference on social characteristics may vary according to the task. Moreover, in Chapter V, we found that there is an effect of robot's emotional feedback that has an effect on children retention of vocabulary and also social engagement in terms of eye-gaze (M. Ahmad, Mubin, Shahid, & Orlando, 2017). Furthermore, we also find the positive effects of informing teaching strategy based on user modelling (Graf, Liu, & Kinshuk, 2008; Bajraktarevic, Hall, & Fullick, 2003). Therefore, we also believe, some children may prefer a competitive behaviour as compared to the encouraging behaviour or a neutral response.

The classification of the behaviours were as follow:

- **Positive (Emphatic/Supportive):** The robot behaves positively on the game events through encouraging the child and/or informing the child about its emotional feelings on the game events. For instance, in the case of the snake far away from 100 mark[2], in the first session, the robot said "I am glad to learn you are looking happy and the snake is not worrying you and you are ahead of me. Lets learn a new word". During the other sessions, the robot said "I am glad to learn you are looking happy, in the last session you had three snakes in the beginning but you still won the game. Lets learn a new word."

- **Negative (Competitive/Critical):** The robot behaves competitively on the game events through contesting with the child and/or informing the child about its emotional feelings on the game events. For instance, in case of snake near 100, in the first session, the robot said, "A snake near 100, I happy to see you are feeling Happy as you are ahead of me, lets learn a new word". During the other sessions, the robot said, "A snake near 100, you also had a snake on 99 in the last game Although you won the game but you cant be lucky every time, lets learns a new word".

- **Neutral:** The robot reacts neutrally on the game events and also reacts neutral in terms of emotional feelings. For instance, in case of six near 100, during the first session, the robot said, "you have a six near 100, its ok, keep playing". During the other sessions, the robot said, "you have a six near 100, its ok, I remember you had a snake after the six in the last game".

The details of the behaviours are also specified in detail in the Chapter V.

[2]Snake Far 100 refers to the case, where the position of the user is at least 50 steps away from100 (winning position).

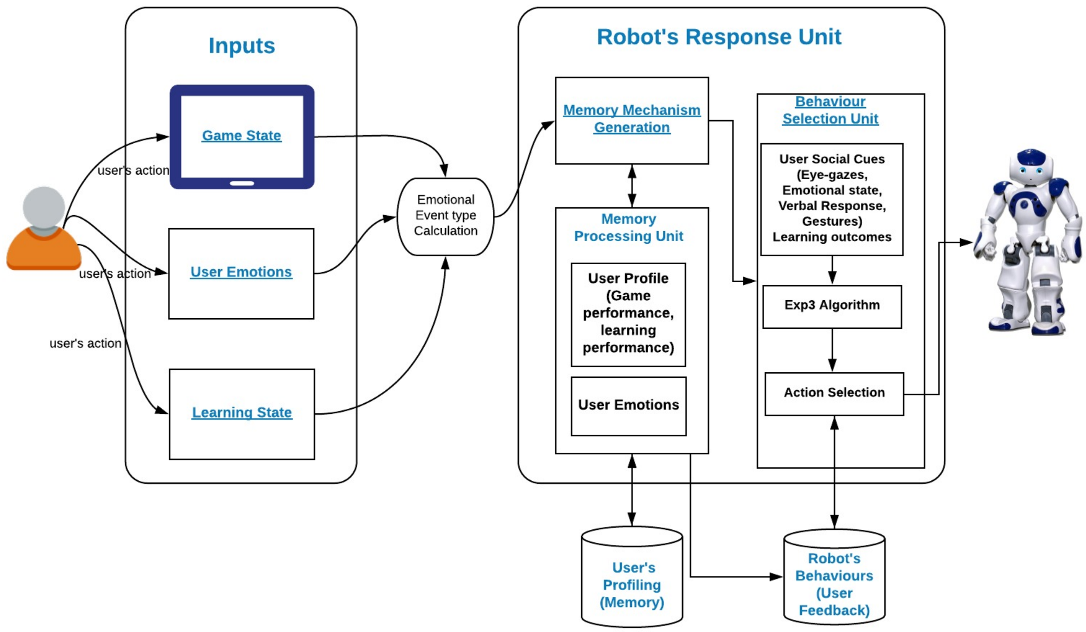

Figure 6.1: Emotion and Memory model (Behaviour Selection Unit Updated).

We used the same snakes and ladders game scenario and also applied the emotion and memory model in the similar way as described in Chapter V during the game and the post-test. In essence, we used the same method to store information during different emotional states of children and applied it in robot's behaviour after the first interaction. In the aforementioned exemplars, the use of the memory varied for each child based on the emotional state of the child during different events.

6.2.1.1 Adaptive Action Selection Strategy

In the study, we designed an adaptive strategy for the interactive scenario and core part of this strategy is modelled by a reinforcement learning framework called Multi-Armed Bandit(MAB) (Mahajan & Teneketzis, 2008). MBA is a standard framework considered when the resources are needed to be distributed to competing actions. In MAB, the ultimate goal is to ask an agent to receive as much accumulated reward as possible in a fixed maximum iteration T. For each iteration t, the agent has K actions and the agent needs to decide what action should it take. After each action, the agent receives an immediate reward R for the action. In our study, there are in total three classifications of actions for the robot to make and the robot needs to decide what is the best action. Therefore, we understand it also becomes a MAB problem. We will describe the Exponential-Weight Algorithm for Exploration and Exploitation (Exp3) algorithm (Bubeck et al., 2012) in detail in the following paragraph.

The algorithm is described in Algorithm 1. In this algorithm, γ is the exploration factor, which decides how much exploration is needed. In our case, we set the γ to be 0.1, which is the same as the default value. For each action i, a weight w_i is assigned. The process was modelled as a Multi-Armed Bandit problem. In order to explain the algorithm in detail, we can consider a process within total K different actions. The Exp3 algorithm with K actions are described in Algorithm 1, where γ is the exploration factor, and w_i is the weight matrix of each action i. $p_i(t)$ indicates the probability of selecting an action i at t iteration, while the capital T indicates the maximum number of iterations. The algorithm starts with the exploration rate of γ. The exploration rate decides the possibility of exploring the unknown actions even when the algorithm knows what is the best action to get the highest reward. The algorithm associates each action with a wight w_i to indicate the importance of each action. The weights will be used later to generate a probability for each action.

After the phase of exploration, the algorithm attempts to iterate in total T times the learning phase in order to optimize the policy. The optimization is done by changing the probabilistic distribution over all the actions. The ultimate goal of the algorithm is to receive as much accumulated reward as possible. For each iteration, the algorithm selects an action i based on the previous distribution P. After executing the action i, the algorithm receives a reward signal $x_{i_t}(t)$ from the environment. The algorithm then generate an estimated reward $\hat{x}_{i_t}(t)$ by taking influence of probability of each action $p_{i_t}(t)$ into account. The estimated reward is defined as $\hat{x}_{i_t}(t) = x_{i_t}(t)/p_{i_t}(t)$. After this step, the algorithm updates the selected action's weight w_i while keeping other actions' weights unchanged. When the algorithm converges, the probabilistic distribution of all the actions is regarded as the best distribution for maximizing the reward.

Algorithm 1 Exp3 Algorithm to select behaviours

procedure INITIALIZATION
    initialize $\gamma \in [0, 1]$
    initialize $w_i(1) = 1, \forall i \in \{1, \ldots, K\}$
    for distribution $\mathcal{P}$,
        set $p_i(t) = (1-\gamma)\frac{w_i(t)}{\sum_{j=1}^{K} w_j(t)} + \frac{\gamma}{K}, \forall i \in \{1, \ldots, K\}$
procedure ITERATION
    repeat
        draw i_t according to $\mathcal{P}$
        observe reward $x_{i_t}(t)$
        define the estimated reward $\hat{x}_{i_t}(t)$ to be $x_{i_t}(t)/p_{i_t}(t)$
        set $w_{i_t}(t+1) = w_{i_t}(t)e^{\gamma \hat{x}_{i_t}(t)/K}$
        set $w_j(t+1) = w_j(t), \forall j \neq i_t \textit{ and } j \in \{1, \ldots, K\}$
        update $\mathcal{P}$:
            $p_i(t) = (1-\gamma)\frac{w_i(t)}{\sum_{j=1}^{K} w_j(t)} + \frac{\gamma}{K}, \forall i \in \{1, \ldots, K\}$
    until T times

We used the Exp3 algorithm (Bubeck et al., 2012) for the social robot to select on of the three behaviours on a game event during the game-phase. It is important to note that the game events on which the robot uttered the response were similar to the ones as discussed in the Chapter V such as "Snake near 100", and "Ladder far from 100". To compute the reward $x_{i_t}(t)$, we measured the social engagement of each individual on each event during game-phase. The social engagement of the user was measured based on the four variables (Eye-gaze facing robot face, Facial Expressions, Verbal Responses and the Gestures). The rationale for choosing these four variables was based on the understanding of the measure of social engagement as explained in Chapter IV. On each event, we computed all the four variables. When a robot generated a response on an event, if the child gaze were facing the robot face during the time the robot was displaying a response, we marked the reward for gaze value as 0.25 otherwise 0. Similarly, we computed the emotional state of the user on the utterance of a robot's response on each event, as one of the goals of applying the algorithm was to enhance social engagement, and happiness or smiling are regarded as some of the positive indicators for social engagement. Therefore, if the child emotional state was happy or smiling on the occurrence of the robot's response on an event, we marked the facial expressions reward as 0.25 otherwise 0. For the verbal response, if the child reacted to an event verbally after the robot's response on an event, we marked the verbal response as 0.25 and lastly, if the child displayed a gesture in the form of a fist or waving hand after the robot has uttered a response on an event, we also marked gestures as 0.25 otherwise 0. We were limited to the number of gestures due to technical limitation. However, the choice of the gestures were based on the observations that fist and wave were the most commonly displayed gestures in the chapter IV. We computed the reward $x_{i_t}(t)$ on each event through adding all the values of the aforementioned four factors. Therefore, the maximum reward on an event was $x_{i_t}(t) = 1$.

To measure the eye-gaze facing the robot, we used *openCV* (Bradski & Kaehler, 2000) to detect eye-gazes of the users. We captured the human-face through an external camera. We detected the gazes in the captured image, in case, the gazes were found in the image, we considered it as eye-gazes facing robot. To ensure that the gazes were directed to the robot, we used the robot's camera installed on the head of the NAO robot. We kept checking user's eye-gaze facing the robot for the complete duration of the robot response. We logged each occurrence of the eye-gaze facing robot in the database. When the robot's response was completed, we counted the number of occurrences of the gazes, in case the gazes was greater than 0, we also calculated reward based on eye-gaze as 0.25 otherwise 0. The rationale for this choice was to make sure that in case the child didn't look at the robot initially, we should not calculate 0 as the reward value.

To measure facial expressions, we detected the emotional state based on the human face. We used the same process as mentioned in Chapter IV where we used Indico (Indico, 2016) API to measure the emotional state on the child's face. We took the average of the six-value pairs as returned by the Indico API throughout the duration of the robot response. At the end of the response, we computed the emotional state through taking an average of the emotional responses. If the detected state remained happy, we gave the reward based on emotional state as 0.25 otherwise 0.

To check for the verbal response, we used Google speech recognition API. On each event at the end of the robot's response, we recorded the user verbal response, in case, the speech API detected a verbal response, we regarded it as child's verbal response toward the robot and gave the reward as 0.25 otherwise 0.

Lastly, we detected gestures through using OpenCV (Bradski & Kaehler, 2000). We captured the frame using an external camera and calculated the number of pointed finger. To calculate the number of fingers, we initially created a bounding rectangular frame around the hand. In the bounding rectangular frame, we checked for the fingertips and finger webbing. We then calculated the number of fingertips pointed towards the camera, a fist was regarded as 0 fingertips and an waving hand was regarded as 5 fingertips. The gestures were checked throughout the robot's response and were also simultaneously logged in the database. In case, at the end, we found a "0" or "5" value, we gave gesture specific reward as 0.25 otherwise 0.

We also used the Exp3 during the post-test phase to choose one of the three feedback (positive, negative, and neutral). To compute the reward function $x_{i_t}(t)$ during the post-test, we considered the learning outcome of the individual. If the child remembered the word, we gave a reward $x_{i_t}(t) = 1$ otherwise, $x_{i_t}(t) = 0$. As our goal was to enhance the learning outcome (retention of vocabulary), and we found in our previous study (M. Ahmad, Mubin, Shahid, & Orlando, 2017), also presented in chapter V that both positive and negative feedback can be critical, therefore, our reward function was based on learning outcome.

6.2.2 Study

6.2.2.1 Interaction Scenario

The interaction scenario was the same as the study presented in Chapter V and also comprised of four phases: 1) Introduction, 2) Pre-test, 2) Snakes and Ladders gameplay and 4) Post-test.

The introduction and pre-test remained the same as explained in Chapter V. In the gameplay and post-test, the robot choose behaviour based on the adaptive strategy as compared to the selection of robot behavior in chapter V. The robot selects behaviour based on a pre-defined mechanism identified in the Chapter V. However, the dice outcomes were also fixed (pre-defined) in this study similar to the previous chapter.

6.2.2.2 Participants

We conducted our between-subject study with 24 children (12-males, 12 females) aged between 10-12 at a school. The mean (M) and standard deviation (SD) of the ages were M: 10.69 and SD: 0.47 respectively. All of the participants were bilinguals. None of the participants had previously interacted with a robot. All the participants characteristics were similar to that of the study presented in the chapter V.

6.2.2.3 Procedure

Our study was setup as a long-term between-subject evaluation that spanned for four school weeks. The study was conducted individually with one child at a time. Each child played the snakes and ladders game with the NAO robot 4 times for 4 days (one session per day) over the course of four school weeks with a gap of six days

Figure 6.2: A Child playing snakes and ladders with NAO.

during sessions, for a total of 96 sessions (24 child * 4 sessions). We conducted the study for four consecutive weeks where eight children participated on their assigned days, either on Wednesday, Thursday or Friday during each week for four weeks.

Each session lasted for approximately 24 minutes and had five steps similar to the procedure presented in Chapter V: 1) a 2-minute introduction, 2) a 4-minute pre-test, 3) a 10-minute game playing session, 4) a 4-minute break, and 5) a 4-minute post-test.

6.2.2.4 Setup and Materials

The setup and materials was similar to the study presented in Chapter V as shown in figure 6.2. The only difference was that in the study presented in chapter V, the NAO robot sat in the crouching posture while in the study presented in the chapter V, NAO was in the sitting posture. The choice of posture of the robot was different due to the limitation of the logistical support i.e. due of difference in the height of the tables against the chairs used in both studies. The height was an important issue as we needed to record the eye-gaze facing robot through the robot's face camera. We also used the same 24 vocabulary words from the Robot Interaction Language (ROILA) similar to Chapter V.

6.2.2.5 Measurements

To measure the effects of our model towards promoting vocabulary learning and sustaining engagement during the game phase, we looked the following Dependent variables (DVs):

1. Children's immediate retention of new words during the session (total number of words remembered in the post-test of every session).

2. Children's Social Engagement across sessions.

To measure the impact of feedback generated during the Post-test on the retention of vocabulary we measured the following DV:

1. Retention of old words across sessions (total number of words remembered during the pre-test taught in the previous sessions).

To measure social engagement, we conducted a video analysis to code following DVs as identified in chapter V. We coded videos for following dependent variables: Gaze facing robot, Verbal responses, Facial Expressions, and Gestures. One researcher was involved in video coding process. It is important to note that this researcher was also involved in the coding process of the Chapter V.

6.3 Results

6.3.1 Social Engagement During Game Phase

To find answers to **RQ1**, we conducted a repeated measure ANOVA with the session as the within-subjects factor with four levels on the game play phase for the following Dependent Variables (DVs): the duration of children's gaze facing the robot, facial expressions (smiles), verbal response, and gestures. It is also important to note there was no difference in the quantity of the interaction during the game phase, therefore, we didn't reqiure a need to normalize our data. The quantity of interaction refers to the total duration of interaction during the game across all sessions. All the duration were recorded in *msecs* respectively.

Our findings showed that during game phase there was a significant effect of session on the children's gazes facing the robot (F(3, 69) = 14.58, p = .000, η_p^2 = .388), children's facial expressions (smiles) (F(3, 69) = 6.95, p = .000, η_p^2 = .232), and verbal responses (F(3, 69) = 4.39, p = .007, η_p^2 = .160). Additionally, there was no effect of session on and children's gestures (F(3, 69) = 1.34, $p = 0.266$, $\eta_p^2 = .055$).

We also conducted Bonferroni test to further examine the significance between sessions for all the DVs. In the case of children's gazes facing the robot, we didn't find significant effect during the first three sessions. However, we witnessed an significant increase in the duration of children's gazes facing the robot during the fourth session. Therefore, the fourth session was statistically significant in comparison with the first ($p = .004$), second ($p = .002$), and third ($p = .000$) session. The mean values for the duration of children's gaze facing the robot were as follows: Session 1 (M: 171.80, SD: 75.18), Session 2 (M: 171.34, SD: 74.64), Session 3 (M: 173.40, SD: 77.00), and Session 4 (M: 258.01, SD: 135.72).

For the children's facial expressions, the third session was statistically significant in comparison with the second ($p = .02$), fourth ($p = .00.29$) session. We witnessed a little decline in the duration of children's facial expressions from first to the second and fourth session. The mean values for the duration of children's facial expressions were as follows: Session 1 (M: 39.96, SD: 25.61), Session 2 (M: 30.41, SD: 26.13), Session 3 (M: 53.22, SD: 40.31), and Session 4 (M: 30.98, SD: 26.18).

In the case of the children's verbal responses, we found that the fourth session was significant in comparison with first (p = .043), second (p = .046) session. We witnessed an upward trend in the duration of children's verbal responses from the first to the fourth session. The mean values for the duration of children's verbal responses were as follows: Session 1 (M: 13.93, SD: 9.95), Session 2 (M: 15.94, SD: 8.07), Session 3 (M: 17.33, SD: 8.23), and Session 4 (M: 20.30, SD: 5.94).

Lastly, for gestures, we didn't observe significance between sessions. We witnessed consistent duration of gestures from first to the fourth session. The mean values for the duration of children's gestures were as follows: Session 1 (M: 6.70, SD: 5.65), Session 2 (M: 5.46, SD: 7.20), Session 3 (M: 8.10, SD: 12.53), and Session 4 (M: 4.63, SD: 4.38).

6.3.2 Vocabulary Learning During Game Phase

To find an answer on the **RQ2**, we checked the immediate retention of words learnt during each gameplay in all the sessions for all the participants. The immediate retention refers to the six words taught during each game.

We conducted a repeated measure ANOVA with the *session* as the within-subjects factor with four levels using immediate retention of words learnt per session as a Dependent Variable (DV). Results showed that there was a significant effect of session ($F(3, 69) = 13.46$, $p = .000$, $\eta_p^2 = .369$) on children's vocabulary learning performance.

We executed Bonferroni posthoc to further examine the effect on the immediate retention of the words learned within sessions. We witnessed that the amount of words retained during the third session were significant in comparison with the first ($p = .000$) and fourth session ($p = .000$). Similarly, the words retained during the second session were significant in comparison to the first ($p = .004$) and fourth ($p = .001$) session. The mean values of the learning outcome for all the sessions were as follows: Session 1 (M: 4.62, SD: 0.96), Session 2 (M: 5.58, SD: 0.77), Session 3 (M: 5.70, SD: 0.55), and Session 4 (M: 4.62, SD: 0.87).

6.3.3 Vocabulary Learning During Post-test

To find an answer on the **RQ3**, we checked for the delayed retention of words tested accross all sessions for all the participants. We conducted a repeated measure ANOVA with the session as the within-subjects factor with *three*, *two* levels using the retention of words learned during the first and second session across remaining interaction sessions as the DV. We didn't find a significant effect of interaction (session) (F(2,46) = .870; p = .426, η_p^2 = .036) on the retention of words learnt during the first session across the second, third and fourth sessions. The mean values of the retention of session 1 words across session was as follow: Session 2 (M: 4.50, SD:1.06), Session 3 (M: 4.70, SD: 1.26), and Session 4 (M: 4.83, SD: 1.16).

We found a significant effect (F(1,23)= 9.12; p = 0.006, η_p^2 = .284) of session on the delayed retention of words learnt during the second session across third and fourth session. The mean values of the retention of session 2 words across session was as follow: Session 3 (M: 5.20, SD:0.58), and Session 4 (M: 4.58, SD: 1.17).

Lastly, the mean values of the retention of session 3 words during the fourth session was as follow: Session 4 (M: 5.29, SD: 0.69).

6.4 Discussion

Our results showed that children's social engagement sustained during all the sessions. We understand that the adaptive strategy to select robot's behaviour based on the four social cues (gazes facing the robot, facial expressions, verbal responses, and gestures) enabled the robot to select the appropriate behaviour according to child's preference on the classification of behaviour. Our findings revealed that the gazes facing the robot and verbal response significantly improved during the fourth session, we conjecture that the Exp3 algorithm received a reward on each action and during the fourth session it was able to generate the behaviour that generated the maximum amount of gazes facing the robot. We also speculate that in the case of the fifth and sixth session, we will witness similar trends. The reason for our conjecture is due to two reasons. Firstly, the robot is adapting and generating novel behaviours incrementally. Secondly, the algorithm enables the robot to not only exploit the existing preferred behaviour but also explore new behaviours at the set rate.

In terms of facial expression, we understand that due to different traits of the participants (Cole et al., 1996), it would have indeed resulted in these findings. Expressions are related to the level of expressiveness of each child and we believe that some children would have become less expressive over time during the fourth session. We also speculate that the algorithm may have converged to generate the most appropriate behaviour for each user and it may have resulted in the maximization of the reward.

In comparison with the findings presented in Chapter V as shown in table 6.1, we believe our learning mechanism was able to generate an increase in the duration of eye-gaze facing the robot and verbal response across the sessions. We understand that this highlights the benefits of applying a learning mechanism for the robot as it was able to personalize its behaviour according to individual's preference. Additionally, we didn't witness any difference between the trends in terms of facial expressions and gestures among both studies presented in this chapter and chapter V, we conjecture that the trends for the duration of gestures and expressions in both studies were similar may be due to similar division of expressiveness and un-expressiveness participants in both studies. In other words, we believe any of the effects may have been balanced out.

In terms of retention of vocabulary, we also witnessed similar trends in comparison to the previous study. We believe that the reason for a decline in vocabulary retention during the last session could be due the reason that children were notified that this was the last interaction session and we understand that they might not have participated in the post-test with the similar enthusiasm or interest as they knew that they won't be tested for the similar words again.

Dependent Variables	Chapter V V.S. Chapter VI
Eye-Gaze across sessions	Higher duration were witnessed in Chapter VI
Facial Expressions across sessions	Similar trends were observed across sessions
Verbal Response across sessions	Higher duration were witnessed in Chapter VI
Gestures across sessions	Similar trends were observed across sessions
Vocabulary retention across sessions	Similar trends were observed across sessions

Table 6.1: Results trends for Eye-gaze, facial expressions, verbal responses and gestures during Chapter V and VI studies

In terms of delayed retention of vocabulary on the words tested across session during the post-test, we understand that the learning mechanism enable the robot to learn the feedback based on the children learning outcome and it reflect on their learning performance across sessions. A decline for the vocabulary retention during the fourth session may have been a result of lack of interest in the task. As mentioned earlier, children may have done it quickly due to the last interaction session. Similar observations have been reported by other researchers (Rosenthal-von der Pütten et al., 2016).

6.5 Conclusion

In this chapter, we automated the process of behaviour selection for our emotion and memory model for a social robot capable of creating emotional memories and adapting accordingly. We applied the MAB algorithm Exp3 to select the behaviour based on the social cues of the child reflecting child's social engagement. We re-evaluated the model in a vocabulary learning task at a school during a children-game-robot interaction. In general, the model also resulted in positive findings based on child's vocabulary learning and also sustaining social engagement during all the sessions.

6.6 Limitations

We didn't compare our results based on the effects of our emotion and memory model on the social engagement with the previous study due to two reasons: Firstly, there was a difference of robot's posture during both studies presented in chapter V and in this chapter. Secondly, this study was performed for four weeks instead of two weeks. We were limited due to logistical issues. The school was not able to provide us with two straight weeks similar to the study in Chapter V because of school events across different weeks. However, we tried to reflect on the findings of both studies in our discussion of results.

CHAPTER VII

Emotion and Memory Model to Promote Mathematics Learning - An Exploratory Long-term Study

In this chapter[1], we present another study in which we applied our emotion and memory model in a mathematics learning scenario. The scenario involved calculating area and perimeter of the regular and irregular shapes. The motivation behind conducting this study was to explore the outcomes of our model in a real-life education scenario. More importantly, the learning scenario implemented in this study was based on the real curriculum and closely represents and resembles the real usage of adaptive robots in education. More particularly, it was not a simulated study or simulated content. Lastly, we also wanted to further ascertain the value of our model in the educational settings.

7.1 Research Method

We conducted an exploratory long-term study, where the robot used the adaptation mechanism based on our emotion and memory model. The aim was to understand

[1] A Conference paper based on the results of this Chapter will be submitted to a peer-reviewed conference on Human-Agent Interaction

the effects of the feedback generated based on our model on children's learning and social engagement during a Mathematics learning task. We measured the effects of our model by comparing it with a control condition, where the robot didn't have any adaptation mechanism. More specifically, we tried to find answers to the following RQs:

RQ1 - What are the effects of our emotion and memory model on the children's social engagement and learning performance, when applied to the robot in a mathematics learning task during a long-term interaction?

RQ1a - Does Children interacting with the robot utilising our emotion and memory model show better learning performance as compared to the control condition, where no model is applied during the mathematics learning task?

RQ1b - Does Children interacting with the robot utilising our emotion and memory model shows the highest level of social engagement as compared to the control condition, where no model is applied during a mathematics learning task?

We hypothesize that in the condition, where the robot applied our emotion and memory model would result in the highest level of social engagement **(H1)** and it will result in better learning outcome **(H2)** during a long-term interaction. Our hypothesis is based on the positive findings of our model as presented in the chapter V and VI. Additionally, also due to the positive findings in terms of improving user's learning due to the inclusion of a learning algorithm to inform robot's behaviour during HRI (Leyzberg et al., 2014; Kennedy et al., 2016).

7.1.1 System Description

We created a mathematics learning test based on calculating areas and perimeter of different shapes as shown in figure 7.1. These shapes included: square, rectangle, triangle, parallelogram and trapezium. We consulted teachers to select these shapes for this study. We were interested in creating the learning task based on the knowledge of the children. In particular, we tried to make sure that the children have limited or no background knowledge about the content of the scenario. Therefore, all the content was created based on the feedback from the teachers. We created the content and received approval from the teachers before incorporating it into our system.

7.1.1.1 Applying Emotion and Memory Model during the maths Test

Keeping the same criteria as explained in Chapter V, we used the learning outcome, child's emotional state as the inputs for the model. We later generated the behaviour for the robot through following the similar mechanism described in chapter V. We used the same set of emotional states as described in the Chapter V. We here explain the information stored by the robot during different events and emotional states. For instance; if a child calculated the area or perimeter of a shape correctly and the emotional state of the child is happy, we created the memory about the outcome of the question about calculating the area for the given shape and its session. Similarly, when the child did not calculate the area or perimeter of the given shape correctly, we created memory depending on the detected emotional state. In the case of sadness, we created memory about the number of attempts taken to calculate the area or perimeter of the given shape correctly. We also created memory about the type of mistake and/or learning outcome for the given shape on the given test. In the case of an angry emotional state, we created memory about not knowing the formula to calculate the area of the given shape. We understand that the child may get angry about not being able to remember how to calculate the area of the given shape. Lastly, in the case of fear, we created memory about the child's outcome on the previous test(s) as the child may feel threatened to match the total score of the last test. Similarly, we also created memory about the type of mistake as the child may fear to repeat the same mistake.

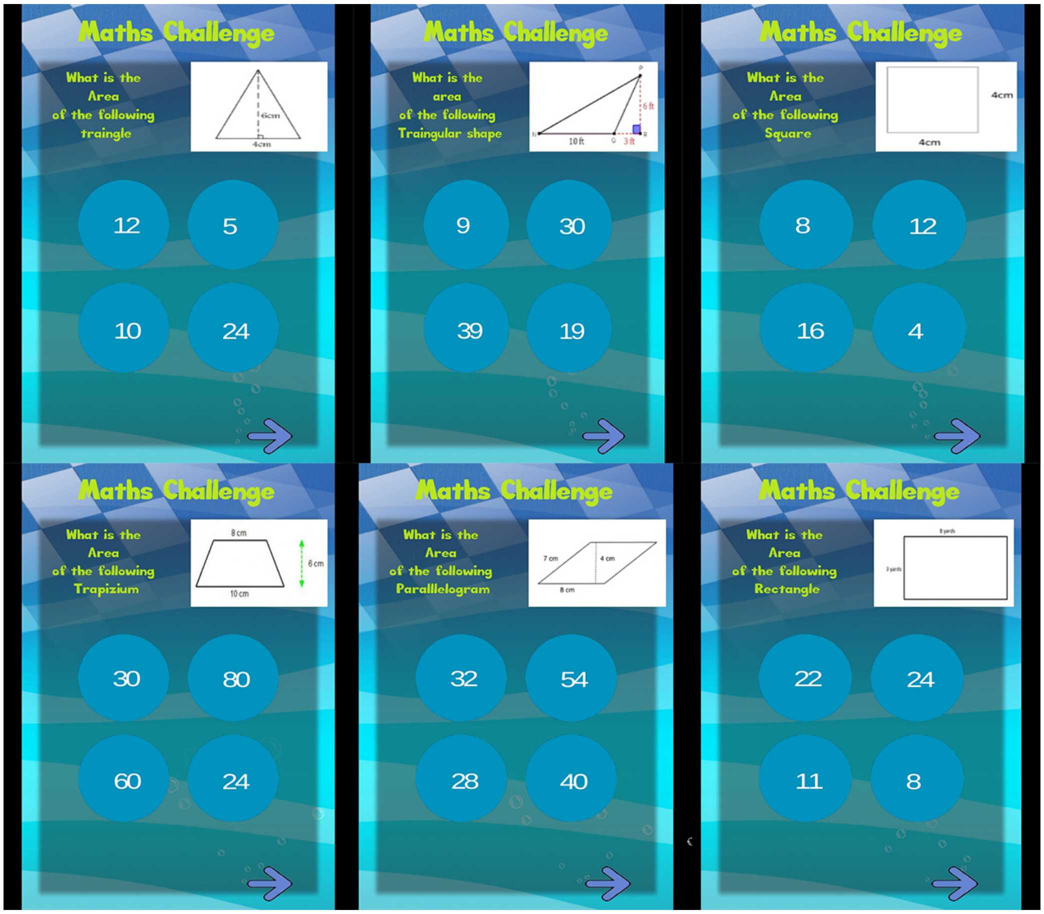

Figure 7.1: Maths Test

Based on this created memory, the robot applied this stored memory in its dialogue and generated one of the three behaviours (positive, negative or neutral) on children's learning outcome. We briefly present an example of the robot's behaviour in table 7.1. At the end of each behaviour presented in the table 7.1, the robot also concatenates the method to calculate the area or perimeter of the given shape along with the correct answer as a part of the feedback. On the occurrence of an event, the robot uses the adaptive strategy described below to select one of the actions from the classification of three different behaviours.

Event	Info	NAO's Behaviour
Learning Outcome: Correct Outcome **Emotional State:** Happy **Event Type:** Positive	outcome of the question to calculate the area of the Triangle	First Session: **Positive Condition:** 1) I am **delighted** to tell that you got this correct. — **Gesture:** Happy or Joy **Negative Condition:** I was not expecting but you got this correct. **Gesture:** Surprised **Neutral Condition:** It is **fine** that you got this correct. — **Gesture:** Neutral Other Sessions: **Positive Condition:** I am **delighted** to remind you that you calculated the area of the triangle correctly in the <SESSION NO> test, and today this is correct — **Gesture:** Joy, Happy. **Negative Condition:** I am **Surprised** to remind that, you calculated the area of the triangle correctly in the <SESSION NO> test , and today this is correct — **Gesture:** Surprised. **Neutral Condition:** This is fine that you calculated the area of the triangle correctly in the <SESSION NO> test and today this is correct — **Gesture:** Neutral.
Learning Outcome: Incorrect Outcome **Emotional State:** Sad **Event Type:** Negative	User's emotional state, no of attempts to calculate the area of the triangle, type of mistake	First Session: **Positive Condition:** 1) It is **alright**, but this is incorrect. — **Gesture:** Emphatic **Negative Condition:** It is **sad**, that this is incorrect. — **Gesture:** Sad **Neutral Condition:** It is **fine**.that this is incorrect. — **Gesture:** Neutral Other Sessions: **Positive Condition:** It is **alright** that in the <SESSION NO> session, you got it wrong and this time your answer is also incorrect. you are not <TYPE OF MISTAKE>. — **Gesture:** Emphatic **Negative Condition:** It is **sad** to remind you that you did not calculate the area of the triangle correctly in the <SESSION NO> test, and today this is incorrect, you are not <TYPE OF MISTAKE> — **Gesture:** Sad. **Neutral Condition:** It is **fine** that you did not calculate the area of the triangle correctly in the <SESSION NO> test, and today this is incorrect, you are not <TYPE OF MISTAKE> — **Gesture:** Neutral

Table 7.1: Taxonomy for the exemplar Robot's Behaviour during maths-test based on Emotion and Memory Model

7.1.1.2 Adaptive Strategy for Behaviour Selection

We used the same adaptive strategy mechanism as described in Chapter VI. The Exp 3 algorithm was used to enable the robot to choose one of the three behaviours (positive, negative and neutral) on the child's learning outcome. However, in this study, the reward function was based on the user task performance. If the outcome of the question was correct, the algorithm receive $x_{i_t}(t) = 1$ otherwise $x_{i_t}(t) = 0$. The rationale for the choice of reward was based on the goal of our system. The goal was to enhance children learning and as our previous findings showed that different types of feedback do impact children learning (M. Ahmad, Mubin, Shahid, & Orlando, 2017).

7.1.2 Study

The entire study and associated protocol was approved by the host universitysethics office (approval number H11429).

The purpose of our study was to evaluate the effects of our model on children engagement and learning performance during a long-term HRI. Our study was a between-subject evaluation. We had two conditions of the robot's behaviour. In condition 1, we enabled the robot to generate behaviour based on our emotion and memory model. In condition 2, the robot didn't implement any adaptation strategy as it was the control condition.

7.1.2.1 Interaction Scenarios

We programmed the NAO robot to autonomously test the mathematics skills of children. NAO was capable of autonomously performing actions during the test.

The Interaction Scenarios was divided into two sections: 1) Introduction, 2) Maths Test. The NAO robot began the introduction through one-to-one interaction with a child and asked the child a set of pre-defined questions similar to the ones in the studies mentioned in the previous chapters.

During the maths test, we had two different conditions. The robot was programmed to ask children ten different questions ranging from calculating area and perimeter of the five shapes: Triangle, Square, Rectangle, Parallelogram, Trapezium. Each test has two questions each on calculating the area and perimeter of each shape. The robot reacted on the outcome of each question in two different ways. In condition 1, the robot applied emotion and memory model. In condition 2 (Control condition), the robot provided feedback on the outcome of the question. For instance, the robot mentioned "This is correct" or "This is incorrect" and concatenated it with the method to calculate the area or perimeter of the given shape along with the correct answer as a part of its feedback.

In both conditions, upon finishing the test, the robot mentioned the score on the test and also greeted the child through saying "Bye, I will see you next time". However, during the last session, the robot said: "Good Bye, hope to see you again".

7.1.2.2 Participants

We conducted our between-subject evaluation at a primary school with 20 participants (10 girls and 10 boys). The study took place at a school during school timings with 3rd-grade children as we wanted to make sure that they didn't have prior knowledge about calculating area and perimeter of the shape. Therefore, the choice of the 3rd grade was based on teacher's advise. The ages of participants were between 8-9 years. The mean and standard deviation for the ages of the participants were M: 8.4, SD: 0.50 respectively. None of the children had interacted with NAO or any social robot before this study.

7.1.2.3 Procedure

Our study was a between-subject evaluation (2 conditions of robot behaviour) spanning a period of two weeks. During the first school week, the first group interacted

Condition	**Session**	**(M, SD) in minutes**
Model	1	M:10.19 ,SD:3.20
Model	2	M:10.34 ,SD:4.25
Model	3	M:9.64 ,SD:3.97
Control	1	M:6.47 ,SD:1.94
Control	2	M:5.08 ,SD:1.27
Control	3	M:4.51 ,SD:1.25

Table 7.2: Duration of the session during the mathematics test

with the robot for the first three consecutive days followed by the second group in the second school week also for the first three consecutive days.

The study was conducted in two steps. In the first step, each group of participants attended a lesson with their maths teacher for the duration of 30 minutes. The contents of the lesson included how to calculate the area and perimeter of square, rectangle, triangle, parallelogram and trapezium. The rationale for conducting a lesson was two fold; firstly, we wanted to give children an introduction to the topic before interacting with the robot, secondly, we were interested in utilising our emotion and memory model on a social robot in a real-life (school-like) situation to further investigate its use at schools, particularly in the real classroom. After the lesson, participants went to the room to participate in the interaction study with the robot.

In the second step, the evaluation was conducted individually with one child at a time in both conditions (control and model). Each child performed the maths test with the NAO robot 3 times on three different days (one session per day), for a total of 60 sessions (20 children * 3 sessions). We conducted our sessions with each group on the 1st, 2nd and 3rd days of the school week respectively. The lesson happened only on the first day in both of the evaluation conditions during both weeks. From the second day, children only interacted with the robot to participate in the mathematics test. The length of each session varied but on average each session lasted for approximately 11 minutes comprising an on-average 10-minute of the mathematics test and 1 minute of introduction at the beginning with the NAO robot.

Figure 7.2: Children participating in the lesson before the interaction with the robot.

Figure 7.3: Setup: A Child performing the maths test

It is important to note that the length of the interaction during the mathematics test was not consistent for all the participants. The Mean (M) and Standard Deviation (SD) values for the duration of the maths test across sessions per condition is given in table 7.2.

7.1.2.4 Setup and Materials

The lesson took place in the school library as shown in figure 7.2 whereas the robot was placed in one of the rooms inside the school library as shown in figure 7.3.

We used NAO robot designed and developed by Aldebaran Robotics. The list of questions on the test during all the session was made after receiving advice from the teachers at the school. The children were also provided with paper-sheets to perform calculations required to complete the test as can be seen in Figure 7.3.

7.1.2.5 Measurements

To measure the effects of our model towards promoting mathematics learning and sustaining social engagement during the mathematics testing phase, we looked the following DVs similar to the previous studies:

1. Children's learning performance on the mathematics test across all the sessions (total number of correct answers on the test during each session).
2. Children's Social Engagement across sessions.

To measure social engagement, we conducted a video analysis to code the following DVs as identified in chapter V. However, we only report results based on the Gazes facing robot because we didn't observe the significant number of Verbal responses, Facial Expressions, and Gestures during the interaction. One researcher was involved in video coding process. It is important to note that this researcher was also involved in the coding process of the Chapter VI.

7.2 Results

To find answers to **RQ1a** and **RQ1b**, we conducted a repeated measure Analysis of Variance (ANOVA) with the session as the within-subjects factor with three levels and type of interaction (model and control) as the between-subject factor using the following set of Dependent Variables (DVs) 1) Gazes, 2) Learning outcome. In this section, we present the result for the duration of the gazes and the learning performance results. It is also important to note there was a difference in the quantity of the interaction during the maths phase, therefore, we normalized our data. The quantity of interaction refers to the total duration of interaction during the game across all sessions. All the duration were recorded in *msecs* respectively. To normalise our data, we computed the percentage of the duration of the eye-gaze facing the robot in each session for all the participants.

7.2.1 Social Engagement Results

Our social engagement results based on the children's gazes showed that we didn't find a significant effect of the type of robot behaviour (i.e. model group vs control group) (F(1, 18) = 1.036, p = .322, η_p^2 = .054) on the children's eye gaze facing the robot during the mathematics test.

On the other hand, there was a significant effect of session on the type of robot's behaviour (i.e. model group vs control group) on children's gazes during the mathematics test (F(2, 36) = 4.14, p = .024, η_p^2 = .695). In other words, robot's behaviour based on emotion and memory model was preferred over the control condition in terms of their eye gazes across sessions. In essence, children were socially more engaged in the model's condition across sessions. The mean and SD values are shown in Table 7.3.

Condition	**Session 1**	**Session 2**	**Session 3**
Model	M: 13.39, SD: 4.21	M: 19.18, SD: 9.71	M: 24.25, SD: 11.55
Control	M: 15.00, SD: 5.61	M: 15.48, SD: 7.08	M: 17.07, SD: 6.52

Table 7.3: Mean and Standard Deviation for the children's eye gazes facing robot during first, second, and third sessions (in percentage).

7.2.2 Maths Learning Results

Our learning results based on the outcome of the children on the test showed that there was a significant effect of the type of robot behaviour (F(1, 18) = 1.036, p = .012, η_p^2 = .301) on the learning performance of the children during the mathematics test.

In essence. results showed that the condition, where robot generated behaviour based on emotion and memory model, resulted in high scores on the test ($p = .012$) as compared to the the control condition.

We also found a significant effect of session on the type of robot's behaviour (i.e. model group vs control group) on children's learning performance during the mathematics test ($F(2, 36) = 6.34$, $p = .004$, $\eta_p^2 = .54$). In other words, robot's behaviour based on emotion and memory model resulted in higher test scores in comparison with the control condition. In other words, children learning performance was better in the condition where robot's behaviour was based on emotion and memory model. The mean and SD values are shown in Table 7.4.

Condition	**Session 1**	**Session 2**	**Session 3**
Model	M: 4.2, SD: 1.14	M: 6.70, SD: 1.39	M: 7.8, SD: 1.22
Control	M: 4.4, SD: 1.50	M: 4.60, SD: 1.77	M: 5.7, SD: 1.41

Table 7.4: Mean and Standard Deviation for the children's learning performance during first, second, and third sessions (the scores are out of 10).

7.3 Discussion

Our results showed that children's level of social engagement was not affected by the type of robot's behaviour. Therefore, our hypothesis (H1) was not accepted. It may have happened because we observed that the children gaze faced the robot each time the robot informed them about the outcome of the question on the test. On the other hand, we found a significant effect of the type of robot's behaviour across sessions as children showed the highest level of engagement when the behaviour was generated through our emotion and memory model in comparison with the control condition across different interaction sessions. We speculate that due to the augmentation of memory from the second session, it would have encouraged children to direct their gazes towards the robot for the longer duration as compared to the control condition.

The reason for this conjecture is based on the results observed in our previous studies (M. Ahmad, Mubin, Shahid, & Orlando, 2017; M. Ahmad, Mubin, & Orlando, 2017a) and also in the study presented in the Chapter VI. Additionally, the use of memory from the second session may have resulted in creating an element of personalisation as also shown in the Chapter V. Moreover, in the third session, we believe the algorithm may have converged and was able generated user-specific feedback behaviour and it would have enabled children to show the highest level of children's social engagement. Lastly, our findings further highlight the role of the adaptive behaviour of the robots in the educational interactions towards impacting their level of engagement.

We also found that the type of robot's feedback based on our model also effected the learning performance of the children. Therefore, our hypothesis (H2) was accepted. We believe that the emotional feedback generated through applying our model would have motivated children to perform better in the mathematics learning task. In particular, the robot's identifying the children specific type of mistakes (missing to add one of the sides in case of calculating the perimeter of the given shape and not dividing the multiple of base and height by 2 in the case of the area of the triangle) and also reminding them about their past performance as a part of its feedback may have also resulted in encouraging children to perform better in the subsequent sessions (M. Ahmad et al., 2016c). Another reason can be the level of interest and attention of the group of individuals during the control condition. We observed that children were quick to give an answer during the control condition as the time taken to answer the questions on the test also declined during the second and third session. On the other hand, children in the group where we applied our model spent consistent time in all the sessions as shown in the table 7.2. Additionally, we provided paper sheets in both conditions, however, children from the second session in the control condition didn't use the sheets. We believe that children lost their interest in interacting with the robot because in the control condition the robot was repetitive and non-incremental in its behaviour as also discussed in previous studies (Kanda et al., 2004). Our results also showed that the learning performance of all the individuals significantly improved during each session. Particularly, children showed the highest level of performance in the last session. We understand that the algorithm may have converged and was able generated user-specific feedback behaviour. As it can be found in the literature that learning is improved when user-specific teaching style is applied in the learning process (Haider et al., 2010; T.-C. Liu et al., 2009), therefore, we also believe that our model had enabled the robot to adapt the feedback style according to the preference of the user.

It is also worth mentioning that we found similar trends in our results based on children's social engagement and learning performance in this study and also in the study presented in Chapter VI. The finding from both studies highlighted the significance of implementing adaptivity in robots in the educational settings to select user-specific feedback. Both of the studies further highlight the positive effects of incorporating our model under different educational scenarios as we believe the feedback generated based on our model motivated children to improve their learning performance during both scenarios. In general, it reflects on implementing and evaluating different types of adaptation mechanism during educational settings to promote children's learning outcome in different learning tasks.

7.4 Conclusion and Future Work

In this chapter, we presented the results of our study that applied our emotion and memory model in the wild in a mathematics learning scenario. The rationale to conduct this study was to explore the outcomes of our model in a real-life scenario and also to further emphasise the value of our model. We conducted an exploratory long-term study, where the robot used the adaptation mechanism based on our emotion and memory model. The purpose of the study was to understand the effects of our model on children's learning and social engagement during a Mathematics concept learning task. Our results showed that in a condition, where our model was implemented on the social robot, children showed the highest level of social engagement across sessions as compared to the condition, where no model was implemented. Similarly, their learning performance based on calculating area and perimeter of regular and irregular shapes was also better in the model condition.

In the future, it would be interesting to evaluate the effects of our model towards promoting children's learning performance during various real-life educational scenarios.

CHAPTER VIII

Conclusions

In this thesis, we addressed the role of adaptivity within a social robot in long-term HRI. More specifically, we investigated the effects of a social robotic companion capable of adapting its behaviour based on user emotions and memory. An Emotion and Memory model was designed by following an iterative process based on the type of information retained by humans during various emotional states. The model was capable of informing future behaviours based on the created account of the archive of user emotional situations. The model was also evaluated to measure its effects on user's social engagement and learning outcome in a long-term HRI educational setup. We evaluated the model in two different educational setups comprising of learning vocabulary and learning how to calculate area and perimeter of both regular and irregular shapes during long-term HRI.

We began with a qualitative exploration of the perception of robots in education. We conducted studies with both teachers and children to understand their perspective on different adaptations by the robot in the education domain. The main distinction between our study and those from other researchers Serholt et al. (2014); Serholt & Barendregt (2014) was that we attempted to acquire perceptions of users after providing them with an experience of interaction with the robot. We believe that results of the previous studies may have been affected by the assumed Knowledge of teachers and children on robots. In both of our studies, our findings highlighted the need for implementing both emotion and memory based adaptations in robots during educational interactions. Teachers were of the opinion that memory based adaptations would result in motivating children to improve their learning and would also improve user experience with the robotic technology.

Keeping the opinions of teachers and children in mind, we conducted a long-term HRI study with three groups of children who played a snakes and ladders game with the NAO robot. The NAO performed 1) game-based adaptations (control group), 2) emotion-based adaptations (the child adapted its behaviour on user emotions), and 3) memory-based adaptation (the child adapted its behaviour on the memory of prior game events). Our goal was to understand the effects of different types of robot's adaptations towards sustaining social engagement in the long-term interaction. Our results showed that emotion-based adaptations were found out to be most effective in terms of social engagement, followed by memory-based adaptations. Game adaptation didn't result in sustaining social engagement during a long-term interaction.

Based on these findings and also understanding the need of a model for the social robot to adapts its behaviour based on the emotions and particularly on memory, we created an emotion and memory model for the social robot that is based on the theory of how humans create memory under various emotional events. The model allowed the robot to create a memory account of a child's emotional events and then adapted its behaviour based on the developed memory. The model was applied to the NAO robot to teach vocabulary to children while playing the game 'Snakes and Ladders'. We conducted an evaluation of our model with 24 children at a primary school for two weeks to verify its impact on children's long-term social engagement and overall vocabulary learning. Our results showed that the behaviour generated based on our model was able to sustain social engagement. Additionally, it also helped children to improve their vocabulary. We also evaluated the impact of the positive, negative, and neutral emotional feedback of the NAO robot on childrens vocabulary learning. Three groups of children (8 per group) interacted with the robot on four separate occasions over a period of two weeks. Our results showed that the condition where the robot displayed positive emotional feedback had a significantly positive effect on the childs vocabulary learning performance as compared to the two other conditions: negative feedback and neutral feedback.

These aforementioned findings on the effects of our emotion and memory model and also on the effects of different emotional feedback towards sustaining social engagement and also promoting children vocabulary retention motivated us to further refine the behaviour selection mechanism in our model. In the previous study, the robot reacted positively, negatively and neutral based on the type of events within the interaction. We also understand that different users may have different preferences on the feedback behaviour of the robot on an event. Similarly, we also understand that different users may prefer the robot to react positively or negatively or neutrally differently under different situations. Therefore, we implemented a mechanism for the social robot to determine about the selection of an action from one of the three classifications of the robot's feedback (positive, negative and neutral) based on the learning performance and social engagement of the individuals. The social engagement of the users was measured through computing eye-gazes facing robot, emotions, verbal responses and gestures of the users during the interaction. Our learning mechanism enabled the robot to choose to learn about an appropriate behaviour for the child based on the social engagement of the child. We also conducted a long-term evaluation of the refined model with a group of 24 children. Our results revealed that the social engagement measured in terms of eye-gaze not only sustained but in fact improved from the first to the last session. Similarly, in comparison to the results of our last study, we also found that children eye-gazes were significantly higher. We believe that the learning mechanism learnt to generate user-centric behaviours during the interaction. In general, our findings highlighted the positive value of our model.

Lastly, in order to further reflect on the value of our emotion and memory model, we applied our model during a real-life scenario based on mathematics learning at a school and also evaluated it in a long-term interaction. Our results showed a positive effect of our model towards promoting children's learning and sustaining social engagement. Based on all the studies and their positive findings, we conclude that our model based on user emotions and memory resulted in creating a social relationship between the robot and children and it was also reflected in terms of their level of interests and learning outcomes.

8.1 Theoretical Refection of the Results

In general, we found empirical evidence showing the positive effects of the use of the robot capable of adapting its behaviour based on user emotions and memory towards maintaining engagement during a social HRI. Our measure for social engagement was based on the amount of verbal and non-verbal interaction between the child and the robot during the interaction. In essence, our research focused on creating or maintaining a social relationship between children and the companion robot by measuring the level of social interaction between them.

In reference to the theoretical elaboration of our results, we understand that Levinger's model of human-human relationship development explains our findings (Levinger, 1983). Levinger (1983) presented a model highlighting five stages of human relationships: 1) (1) acquaintance, (2) buildup, (3) continuation, (4) deterioration and (5) termination. We are particularly interested in the first three stages to describe the theoretical relevance of our findings. There exist a number of factors that involves acquainting with someone (human) such as first impressions, physical appearance, behaviour, attitude and personality (Feingold, 1992). According to one of the attitude similarity theories, the similarity of attitudes, individual preferences, previous relational history is among the reinforcing factors towards creating an element of attraction between the two individuals (Byrne, 1997). Other factors include common circumstance between the two individuals (Orbuch & Sprecher., 2006).

The second and third stage of Levinger's model deals with the maintenance of the human relationship. We understand that a number of actions/behaviours are performed by humans to maintain a relationship. These actions/behaviours have been categorized into two types (routine and strategic behaviours) (L. Stafford et al., 2000). Routine behaviours are defined as "those behaviours where people engage in for other reasons which serve to maintain a relationship as a side effect (such as performing daily tasks together (Bickmore & Picard, 2005)". On the other hand, strategic behaviours are those "which individuals enact with the conscious intent of preserving or improving the relationship" (L. Stafford et al., 2000). Particularly, we are interested in the strategic behaviours such as: have a social dialogue, recalling past events, providing support, giving advise or increasing trust (Duck, 2007).

Keeping the theoretical perspective of human-human relationship in mind, researchers in HRI have also highlighted various human-robot relationship maintenance strategies. Researchers believe that different strategic behaviours could be applied to the robots during long-term interaction in different social settings (Bickmore & Picard, 2005). These strategic behaviours include using humour during the dialogue, recalling user' past events, understanding and reacting to user emotions, and several similar behaviours (Fong et al., 2003). Similarly, as we also implemented similar strategic behaviours in robots and evaluated them during long-term interactions with children (humans) in educational settings, therefore, there exists a relevance between our findings and the aforementioned human-human relationship theories. In essence, it can be inferred from our findings that humans create relationship with robots in the similar fashion. When a robot adapts its behaviour through understanding human emotions or through recalling past events, it creates an element of attraction and it, as a result, generates the higher amount of both verbal and non-verbal behaviours during the HRI.

8.2 Guidelines for Designing Adaptations in Future Robots

Based on the state of the art and the work presented in this thesis, we develop a set of directions/guidelines for designing future adaptive social robot capabilities for the long-term interaction. Our aim is to direct researchers in the field of HRI and social robotics to design different adaptations capabilities in social robots to promote their long-term applicability across various social domains. Additionally, we would also present a set of open challenges that need to addressed to facilitate future research on designing and evaluating future adaptive social robots.

1. *User and Adaptations:* The personal characteristics of a user (level of expertise on a given task, age, or gender, personality) need to be taken into consideration when designing adaptive robots. Most of the current adaptation mechanisms have focused on user performance or sentiment on a certain task and user profile. Our systematic review presented in Chapter II highlighted limited research reported on adaptive systems that have been designed to adapt according to user characteristics other than task performance and similar variables. Some studies have reported on the effect of user's gender, age and skill level during the interaction. For instance, female users have been reported to be more social as compared to male (de Greeff & Belpaeme, 2015). In addition, the gender, age of and skill of the user have also affected the social engagement and interest of the user during the interaction (Torrey, Powers, Marge, Fussell, & Kiesler, 2006; Cameron et al., 2015). Therefore, we need to design and evaluate robots that can adapt based on user characteristics in real-time and study their effect on perception, engagement and task performance of the user.

2. *User Emotions and Adaptation:* Emotions are one of the basic principles of social interaction (Andersen & Guerrero, 1998) and the significance of understanding emotions and adapting to them during robotic interactions has been consistently reported in various studies based on short-term interactions (Belpaeme et al., 2012; M. Ahmad, Mubin, & Orlando, 2017a). Most of the past research has utilised algorithms that give information about user's affective state via a facial scan in real-time. However, results of these studies have shown a positive effect on user experience but we intend to direct researchers to develop novel methods to understand the emotional state of the users. For instance; it may be achieved through measuring the varying pitch of the voice as also identified in the study present in Chapter III during an interaction, or through understanding the common patterns during the interaction to recognise the emotions. For instance; Cuadrado et al. (2016) presented a model to recognise user emotions by analysing keyboard and mouse movements in relation to their interactions with robots.

3. *Robot's Memory and Adaptation:* Our thesis findings highlighted the significance of the role of memory towards promoting learning and sustaining social engagement in a long-term interaction in an educational scenario. Additionally, other researchers have emphasised on implementing robots that can simulate having memory. It has been predicted the future of social human-robot interaction resides in the past (Baxter, 2016; Leite et al., 2017; Kennedy et al., 2016) as it can help mitigate many HRI existing challenges (Tapus et al., 2007b). One of the key elements of memory adaptations lies into exploring the possibilities in which the robot should be able to forget about the past interactions as also identified in the recent past (Stanton, 2017). We direct researchers to use reinforcement learning based algorithms or ranking algorithms to understand forgetfulness in HRI. For instance: we can enable the robot to forget about

events based on the varying social engagement during an interaction. A robot can choose the type of memory based on the perceived interest of the user. To implement this process, we can either apply reinforcement learning or use one of the ranking algorithms to rank memory events. We also propose a similar kind of technique that is followed at Facebook (Constine, 2016), where a user does not see the news feed of friends whom he has not visited or commented on the feeds in a long time.

4. *Personality and Adaptation:* In general, the personality of the user is categorised as extrovert or introvert. It is also known that the mood of a user can influence user's personality depending on various events that may happen during the day (Smith & Petty, 1995). Hence, the personality of the user does not remain static and may vary depending on the mood of the user. Our literature review presented in Chapter II showed that limited work has been reported on social robots that can modify their behaviour based on the personality of the users in the real-time (De Smedt, 2015). Most of the work on such Adaptive social robots is based on the adaptation mechanism where user personality is understood through a standard pre-test questionnaire. In other words, most prior work has reported a limited set of personalities or adaptive behaviour. We envision that future research should be performed on understanding user's personality in real-time. One of the methods can be based on a joint approach to machine learning that can enable robot's to modify its behaviour in real-time. For instance; Ritschel & André (2017) proposed a method to design real-time personality adaptation based on reinforcement learning through understanding social signals of the users.

5. *Robot's Voice Adaptation:* Another aspect that needs attention deals with implementing voice adaptation in social robots and also studying its effect on

the user preferences in various domains. Recently, Lubold et al. (2016) showed that a social voice-adaptive dialogue had a significant effect on social presence as compared to a simple social dialogue. Therefore, we direct researchers to implement various ways of robot's voice adaptation and evaluate their effect on children's learning and engagement. Voice-based adaptations can be implemented through analysing the speech of the user. For instance: understanding user's motivation and engagement through measuring the expressiveness in their speech. Similarly, the social robot may adapt their conversational style based on the user's level of engagement on the task.

6. *Culture and Adaptation:* In the prior literature reported in the field of HRI, we did not find examples of robots adapting based on the user's culture or demographic background. It has been shown in the past that children belonging to the Pakistani culture were found to be more expressive as compared to children from the Dutch culture in one of the HRI studies (Shahid et al., 2008). Similarly, the perception of robots also varies across different cultures (Haring et al., 2014). Additionally, every social environment has its own culture and it can have a different effect on the users during a social interaction (Rau et al., 2009). For instance, the use of robots in the home settings may require robots to adapt their behaviour differently depending on the culture followed at that particular home. This has also been specified in the Chapter II. Consequently, we emphasise on the need of integrating culture while implementing Adaptive Social Robots. We believe that understanding and applying culture-specific adaptation would also promote personalisation during long-term HRI.

8.2.1 Open Challenges and Future Work

Social Robotics and implementing adaptivity in the robots is relatively a new area of research. We need to find answers to a range of research questions in the near future.

In this section, we will be presenting general issues that revolve around designing, implementing and evaluating Adaptive Social Robots.

1. *Understanding the Context of Adaptation:* It is important to understand which individual adaptive behaviours lead to a positive influence on users. In other words, which of these behaviours result in intimidation or confusion in various contexts or environments. We believe that depending on the user characteristic and the social environment, a particular set of adaptive robot behaviours need to be implemented. The set of adaptive behaviours may depend on the environment a certain robot is operating in. We believe, for instance, that user personality based adaptation may not be needed in a public space, simply because of the complexity of multi-user interaction. We also compared the effect of different adaptations based on user emotions, memory and game events towards maintaining social engagement during a long-term interaction at a school with children (M. Ahmad et al., 2016b). Our results showed that emotion-based adaptations were found out to be most effective, followed by memory-based adaptations. Game adaptation didnt result in sustaining long-term social engagement. We also highlighted on implementing the varying level of robot's adaptations across different domain areas in chapter II. We understand that a deeper investigation is needed towards understanding the impact of a certain adaptation on users engagement or task performance during a certain scenario.

2. *Evaluation Metrics:* The evaluation metrics needed to evaluate adaptive systems should be investigated deeply. We, unfortunately, do not find a common protocol to evaluate the effects of an adaptive social robot on different factors in different domains. There is a need for designing a protocol as it would help researchers to discuss their results with previous findings in a systematic manner. As adaptive system adapts and changes their behaviour based on user behaviours, therefore, these systems should be evaluated for long-term inter-

actions to confirm their potential. Similarly, It would also be interesting to find whether existing questionnaires, measurements, protocol for one-off (single) interactions would also apply long-term interactions? Most of the results reported in our review in Chapter II are based on the video analysis conducted for the interactive sessions. Unfortunately, there is also no protocol to analyse these videos for a set of measurements for different domains. We, in this thesis, have given the scheme to measure social engagement during HRI. In summary, a set of guidelines are required which define evaluation methodologies for ASR as well as analysis of data emerging from the evaluation sessions.

3. *Ethical Concerns:* An adaptive robot needs to store information about the patterns of interaction with the user. Therefore, the privacy of data is one of the issues that need to be taken into consideration. We need to define guidelines that can maintain ethical considerations and give directions on what kind of data that needs to be stored and that would potentially be used, especially in the cases when the user group is based on children. Another issue is about understanding the acceptable adaptations, as a certain group of users might be intimidated with the adaptation of the robot, or develop a sense of discomfort with the robot's unpredictability. Similarly, we also found such observations made by both teachers and children in our previous studies (M. Ahmad et al., 2016c). Therefore, we also need to research the issue of user fears while interacting with the robot. Other researchers have also suggested considering ethical and privacy consideration in terms of data collections (Hudlicka, 2017; Leite, Martinho, & Paiva, 2013; M. Ahmad, Mubin, & Orlando, 2017b).

We understand that there are several research directions that can be taken in the future as a result of the work presented in this thesis. Firstly, as all of our studies were conducted in primary schools, it would be interesting to investigate the effect of our method on the social engagement and learning outcome during a long-term interaction on the older children belonging to high schools. Secondly, It would be interesting to apply our model on the other types of social agents such as avatars or virtual characters in different educational settings. It would also be interesting to compare the results of our studies with the one conducted with other social agents in order to further signify the value of our model in terms of its modular design. Lastly, our model was grounded in the process of creation and retrieval of the memory about the external emotional event as described by Ledoux (J. LeDoux, 2007). It is well-recognized that humans create memories of both positive and negative emotional experiences. It has been shown that different types of information are remembered under various emotional states (Levine & Pizarro, 2004). We utilised this procedure to enable a robot to create a memory of user emotional situations. We then directed the social robot use this memory in its dialogue and behave accordingly. However, one of the limitations of our and most past models has been not including culture as a factor during the understanding of emotions based on facial expressions and information processing to generate a social behaviour for the robot (Russell, 2017).

Culture can play an important role in terms of understanding of emotions that can be used to inform robots behaviour (Russell, 2017). As our emotion and memory model was based on creating a memory of the users emotional events, we believe that the emotional events may be influenced due to their cultural context. A person's culture may reflect on ones thinking and may also influence ones judgment about a certain emotional event (Russell, 2017). For instance; an event type may be considered as negative in one culture but not in another. Consequently, the cultural differences may change the definition of an event. Additionally, in our emotion and memory model, we coded the types of event based on the definition of positive and negative events and also through the facial expressions coded in one of our previous studies (M. Ahmad, Mubin, & Orlando, 2017a). We didn't take culture into consideration. Attribution is another aspect of the understanding of an event. It is the process of linking causes and effects. As people differ in their respective cultures, therefore, attribution of an event as negative for an Australian may not align with the Japanese (Weiner, 1985). Furthermore, as it is important to realize that display rules and feelings are socially shared and individuals, belonging to different cultures may differ in the way they appraise an event. Therefore, it is important to understand that does assessment of the same event varies across culture? Lastly, we recognized basic emotions of the user based on their facial expressions. The emotion recognition was primarily based on the Basic Emotion Theory that emphasizes on the universality of recognition (Ekman, 1980). As it has been reported that the matching scores on the emotion recognition vary based on the culture and language (Nelson & Russell, 2016). Therefore, we would want to reflect on this to inform the future research.

8.3 Selected Publications

The work presented in this dissertation appeared in a number of publications. We provide below a list of most relevant publications that contributed to the distribution of this work.

1. Ahmad, M., Mubin, O., Escudero, P. (2015, October). Using adaptive mobile agents in games based scenarios to facilitate foreign language word learning. In Proceedings of the 3rd International Conference on Human-Agent Interaction (pp. 255-257). ACM.

2. Ahmad, M. I., Mubin, O., Orlando, J. (2016, October). Understanding behaviours and roles for social and adaptive robots in education: teacher's perspective. In Proceedings of the Fourth International Conference on Human Agent Interaction (pp. 297-304). ACM.

3. Ahmad, M. I., Mubin, O., Orlando, J. (2016, November). Children views' on social robot's adaptations in education. In Proceedings of the 28th Australian

Conference on Computer-Human Interaction (pp. 145-149). ACM.

4. Ahmad, M. I., Mubin, O., Orlando, J. (2016, November). Effect of different adaptations by a robot on children's long-term engagement: an exploratory study. In Proceedings of the 13th International Conference on Advances in Computer Entertainment Technology (p. 31). ACM.

5. Ahmad, M., Mubin, O., Orlando, J. (2017). A systematic review of adaptivity in human-robot interaction. Multimodal Technologies and Interaction, 1(3), 14.

6. Ahmad, M. I., Mubin, O., Orlando, J. (2017). Adaptive social robot for sustaining social engagement during long-term childrenrobot interaction. International Journal of HumanComputer Interaction, 33(12), 943-962.

7. Ahmad, M. I., Mubin, O., Shahid, S., Orlando, J. (2017). Emotion and memory model for a robotic tutor in a learning environment. In Proceedings of the Seventh ISCA workshop on Speech and Language Technology in Education 2017, August 25-26, 2017, Djur, Stockholm, Sweden (pp. 26-32).

APPENDICES

APPENDIX A

Tables for the p-values of the Frequencies and Duration Results of Chapter IV

Interval	Dependent Variable-with-frequency	F(1 , 20)	p-value
introduction-greetings	gaze	13.53	< 0.001
	facial-expression	35.28	< 0.001
	verbal-response	26.58	0.001
	gesture	2.96	0.101
game-play	gaze	0.581	< 0.001
	facial-expression	0.732	< 0.001
	verbal-response	0.434	0.001
	gesture	5.505	0.0036
Interval	**Variable-with-duration**		
introduction-greetings	gaze	59.414	< 0.001
	facial-expression	17.018	0.001
	verbal-response	11.11	0.003
	gesture	2.024	0.17
game-play	gaze	0.0005	0.942
	facial-expression	1.175	0.199
	verbal-response	0.324	0.576
	gesture	1.226	0.058

Table A.1: Frequency and duration Results of the effect of session during introduction-greetings and game-play on the DVs

Interval	Variable-with-frequency	F(2, 20)	p
introduction-greetings	gaze	0.478	0.627
	facial-expression	16.807	< 0.001
	verbal-response	0.145	0.86
	gesture	1.96	0.169
game-play	gaze	4.604	0.023
	facial-expression	7.132	0.005
	verbal-response	3.997	0.035
	gesture	5.217	0.015
Interval	**Variable-with-duration**		
introduction-greetings	gaze	2.442	0.11
	facial-expression	4.527	0.024
	verbal-response	0.917	0.416
	gesture	1.818	0.190
game-play	gaze	5.711	0.011
	facial-expression	3.6	0.04
	verbal-response	4.13	0.031
	gesture	3.01	0.07

Table A.2: Frequency and duration results of the effect of session * adaptation type during introduction-greetings and game-play on the DVs

Interval	**Variable-with-frequency**	**F(2, 20)**	**p**
introduction-greetings	gaze	0.706	0.506
	facial-expression	0.036	0.965
	verbal-response	3.171	0.06
	gesture	1.946	0.169
game-play	gaze	1.355	0.281
	facial-expression	0.990	0.389
	verbal-response	0.545	0.588
	gesture	0.629	0.543
Interval	**Variable-with-duration**		
introduction-greetings	gaze	5.379	0.01
	facial-expression	1.613	0.224
	verbal-response	1.577	0.231
	gesture	1.808	0.19
game-play	gaze	0.552	0.584
	facial-expression	0.781	0.471
	verbal-response	1.276	0.301
	gesture	1.014	0.381

Table A.3: Frequency and duration results of the effect of adaptation type during introduction-greetings and game-play on the DVs

Interval	Variable-with-frequency	F(1 , 19)	p
end-greetings	gaze	8.121	0.01
	facial-expression	18.170	< 0.0011
	verbal-response	0.016	0.899
	gesture	8.728	0.008
complete-session	gaze	0.623	0.44
	facial-expression	0.136	0.71
	verbal-response	1.69	0.20
	gesture	5.795	0.02
Interval	**Variable-with-duration**		
end-greetings	gaze	37.359	< 0.001
	facial-expression	15.536	0.001
	verbal-response	19.77	0.001
	gesture	6.133	0.02
complete-session	gaze	0.416	0.527
	facial-expression	2.233	0.15
	verbal-response	3.433	0.07
	gesture	1.405	0.250

Table A.4: Frequency and duration Results of the effect of session during end-greetings and complete-session on the DVs

Interval	Variable-with-frequency	F(2 , 19)	p
end-greetings	gaze	1.186	0.327
	facial-expression	4.04	0.035
	verbal-response	6.601	0.047
	gesture	0.663	0.527
complete-session	gaze	4.083	0.03
	facial-expression	9.56	0.001
	verbal-response	4.6	0.02
	gesture	5.866	0.01
Interval	**Variable-with-duration**		
end-greetings	gaze	10.756	0.001
	facial-expression	6.729	0.006
	verbal-response	4.976	0.018
	gesture	1.172	0.331
complete-session	gaze	4.12	0.03
	facial-expression	4.41	0.02
	verbal-response	3.89	0.03
	gesture	3.521	0.05

Table A.5: Frequency and duration results of the effect of session * adaptation type during end-greetings and complete-session on the DVs

Interval	**Variable-with-frequency**	**F(2 , 19)**	**p**
end-greetings	gaze	0.605	0.327
	facial-expression	4.215	0.031
	verbal-response	0.726	0.497
	gesture	2.454	0.113
complete-session	gaze	0.698	0.510
	facial-expression	1.094	0.355
	verbal-response	0.261	0.773
	gesture	1.535	0.241
Interval	**Variable-with-frequency**		
end-greetings	gaze	42.756	< 0.001
	facial-expression	2.479	0.111
	verbal-response	4.976	0.018
	gesture	3.161	0.06
complete-session	gaze	5.5	0.01
	facial-expression	1.3	0.28
	verbal-response	0.57	0.57
	gesture	1.63	0.22

Table A.6: Frequency and duration results of the effect of adaptation type during end-greetings and complete-session on the DVs

APPENDIX B

List of ROILA Words

No.	ROILA	ENGLISH
1	jabami	hi
2	make bama	good bye
3	kanek	go
4	botama	turn
5	babalu	stop
6	koloke	forward
7	webufo	left
8	besati	right
9	nole	back
10	jinolu	ball
11	lakowo	cat
12	fipuko	dog
13	belutu	boy
14	batuno	girl
15	piwaja	flower
16	wikute	fruit
17	fupuma	music
18	lobo	robot
19	lujusi	box
20	bubas	house
21	bokubo	school
22	panope	garden
23	kepete	shop
24	wapisi	bucket

Table B.1: Words from ROILA Language Omar Mubin (2015).

APPENDIX C

Language Background Questionnaire

Please encircle the languages you speak at home and school

- English
- Arabic
- Urdu
- Hindi
- German
- Dutch
- Others:

BIBLIOGRAPHY

BIBLIOGRAPHY

Abe, K., Hieida, C., Attamimi, M., Nagai, T., Shimotomai, T., Omori, T., & Oka, N. (2014). Toward playmate robots that can play with children considering personality. In *Proceedings of the second international conference on human-agent interaction* (pp. 165–168).

Adcock, A., & Van Eck, R. (2012). Adaptive game-based learning. In *Encyclopedia of the sciences of learning* (pp. 106–110). Springer.

Ahmad, M., Mubin, O., & Orlando, J. (2016a). Children views' on social robot's adaptations in education. In *Proceedings of the 28th australian conference on computer-human interaction* (pp. 145–149).

Ahmad, M., Mubin, O., & Orlando, J. (2016b). Effect of different adaptations by a robot on children's long-term engagement: An exploratory study. In *Proceedings of the 13th international conference on advances in computer entertainment technology* (p. 31).

Ahmad, M., Mubin, O., & Orlando, J. (2016c). Understanding behaviours and roles for social and adaptive robots in education: teacher's perspective. In *Proceedings of the fourth international conference on human agent interaction* (pp. 297–304).

Ahmad, M., Mubin, O., & Orlando, J. (2017a). Adaptive social robot for sustaining social engagement during long-term children–robot interaction. *International Journal of Human–Computer Interaction*, 1–20.

Ahmad, M., Mubin, O., & Orlando, J. (2017b). A systematic review of adaptivity in human-robot interaction. *Multimodal Technologies and Interaction*, *1*(3), 14.

Ahmad, M., Mubin, O., Shahid, S., & Orlando, J. (2017). Emotion and memory model for a robotic tutor in a learning environment. In *Proceedings of the seventh isca workshop on speech and language technology in education 2017, august 25-26, 2017, djur, stockholm, sweden* (pp. 26–32).

Ahmad, M. I., & Shahid, S. (2015). Design and evaluation of mobile learning applications for autistic children in pakistan. In *Human-computer interaction* (pp. 436–444).

Aly, A., & Tapus, A. (2013). A model for synthesizing a combined verbal and nonverbal behavior based on personality traits in human-robot interaction. In *Proceedings* *of the 8th acm/ieee international conference on human-robot interaction* (pp. 325–332).

Andersen, P. A., & Guerrero, L. K. (1998). Principles of communication and emotion in social interaction. *Handbook of communication and emotion: Research, theory, applications, and contexts*, 49–96.

Argyle, M., & Dean, J. (1965). Eye-contact, distance and affiliation. *Sociometry*, 289–304.

Asif, M., Sabeel, M., & Mujeeb-ur Rahman, K. Z. (2015). Waiter robot–solution to restaurant automation. In *Proceedings of the 1st student multi disciplinary research conference (mdsrc). university of wah, pakistan.*

Bajraktarevic, N., Hall, W., & Fullick, P. (2003). Incorporating learning styles in hypermedia environment: Empirical evaluation. In *Proceedings of the workshop on adaptive hypermedia and adaptive web-based systems* (pp. 41–52).

Bartneck, C. (2003). Interacting with an embodied emotional character. In *Proceedings of the 2003 international conference on designing pleasurable products and interfaces* (pp. 55–60).

Bartneck, C., Kanda, T., Mubin, O., & Al Mahmud, A. (2007). The perception of animacy and intelligence based on a robot's embodiment. In *2007 7th ieee-ras international conference on humanoid robots* (pp. 300–305).

Bartneck, C., Kulić, D., Croft, E., & Zoghbi, S. (2009). Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International journal of social robotics*, *1*(1), 71–81.

Bates, J., et al. (1994). The role of emotion in believable agents. *Communications of the ACM*, *37*(7), 122–125.

Baxter, P. (2016). Memory-centred cognitive architectures for robots interacting socially with humans. *arXiv preprint arXiv:1602.05638*.

Baxter, P., Ashurst, E., Read, R., Kennedy, J., & Belpaeme, T. (2017). Robot education peers in a situated primary school study: Personalisation promotes child learning. *PLoS One*, *12*(5), e0178126.

Baxter, P., & Belpaeme, T. (2014). Pervasive memory: the future of long-term social hri lies in the past. In *Third international symposium on new frontiers in human-robot interaction at aisb 2014.*

Becker, W. (1977). Teaching reading and language to the disadvantagedwhat we have learned from field research. *Harvard Educational Review*, *47*(4), 518–543.

Beer, J., Fisk, A. D., & Rogers, W. A. (2014). Toward a framework for levels of robot autonomy in human-robot interaction. *Journal of Human-Robot Interaction*, *3*(2), 74.

Beer, J., Liles, K. R., Wu, X., & Pakala, S. (2017). Affective human–robot interaction. In *Emotions and affect in human factors and human-computer interaction* (pp. 359–381). Elsevier.

Belpaeme, T., Baxter, P. E., Read, R., Wood, R., Cuayáhuitl, H., Kiefer, B., ... others (2012). Multimodal child-robot interaction: Building social bonds. *Journal of Human-Robot Interaction*, *1*(2), 33–53.

Benenson, J. F., Pascoe, J., & Radmore, N. (2007). Children's altruistic behavior in the dictator game. *Evolution and Human Behavior*, *28*(3), 168–175.

Benitti, F. B. V. (2012). Exploring the educational potential of robotics in schools: A systematic review. *Computers and Education*, *58*(3), 978–988.

Bickmore, T. W., & Picard, R. W. (2005). Establishing and maintaining long-term human-computer relationships. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *12*(2), 293–327.

Boccanfuso, L., & OKane, J. M. (2011). Charlie: An adaptive robot design with hand and face tracking for use in autism therapy. *International journal of social robotics*, *3*(4), 337–347.

Bower, G. H. (1992). How might emotions affect learning. *The handbook of emotion and memory: Research and theory*, *3*, 31.

Bradski, G., & Kaehler, A. (2000). Opencv. *Dr. Dobbs journal of software tools*, *3*.

Breazeal, C. (2003). Toward sociable robots. *Robotics and autonomous systems*, *42*(3-4), 167–175.

Breazeal, C. (2009). Role of expressive behaviour for robots that learn from people. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *364*(1535), 3527–3538.

Broadbent, E., Stafford, R., & MacDonald, B. (2009). Acceptance of healthcare robots for the older population: review and future directions. *International Journal of Social Robotics*, *1*(4), 319–330.

Brown, L., Kerwin, R., & Howard, A. M. (2013). Applying behavioral strategies for student engagement using a robotic educational agent. In *Systems, man, and cybernetics (smc), 2013 ieee international conference on* (pp. 4360–4365).

Bruce, A., Nourbakhsh, I., & Simmons, R. (2002). The role of expressiveness and attention in human-robot interaction. In *Robotics and automation, 2002. proceedings. icra'02. ieee international conference on* (Vol. 4, pp. 4138–4142).

Bubeck, S., Cesa-Bianchi, N., et al. (2012). Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *Foundations and Trends® in Machine Learning*, *5*(1), 1–122.

Buche, C., Querrec, R., De Loor, P., & Chevaillier, P. (2003). Mascaret: pedagogical multi-agents systems for virtual environment for training. In *Cyberworlds, 2003. proceedings. 2003 international conference on* (pp. 423–430).

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of educational research*, *65*(3), 245–281.

Byrne, D. (1997). An overview (and underview) of research and theory within the attraction paradigm. *Journal of Social and Personal Relationships*, *14*(3), 417–431.

Cahill, L., Haier, R. J., Fallon, J., Alkire, M. T., Tang, C., Keator, D., ... Mcgaugh, J. L. (1996). Amygdala activity at encoding correlated with long-term, free recall of emotional information. *Proceedings of the National Academy of Sciences*, *93*(15), 8016–8021.

Cameron, D., Fernando, S., Collins, E., Millings, A., Moore, R., Sharkey, A., ... Prescott, T. (2015). Presence of life-like robot expressions influences children's enjoyment of human-robot interactions in the field. In *Proceedings of the aisb convention 2015.*

Canli, T., Zhao, Z., Brewer, J., Gabrieli, J. D., & Cahill, L. (2000). Event-related activation in the human amygdala associates with later memory for individual emotional experience. *Journal of Neuroscience*, *20*(19), RC99–RC99.

Carini, R. M., Kuh, G. D., & Klein, S. P. (2006). Student engagement and student learning: Testing the linkages. *Research in higher education*, *47*(1), 1–32.

Castellano, G., Pereira, A., Leite, I., Paiva, A., & McOwan, P. W. (2009). Detecting user engagement with a robot companion using task and social interaction-based features. In *Proceedings of the 2009 international conference on multimodal interfaces* (pp. 119–126).

Castro-González, Á., Malfaz, M., & Salichs, M. A. (2011). Learning the selection of actions for an autonomous social robot by reinforcement learning based on motivations. *International Journal of Social Robotics*, *3*(4), 427–441.

Cesta, A., Cortellessa, G., Giuliani, M. V., Pecora, F., Scopelliti, M., & Tiberio, L. (2007). Psychological implications of domestic assistive technology for the elderly. *PsychNology Journal*, *5*(3).

Chang, C.-W., Lee, J.-H., Chao, P.-Y., Wang, C.-Y., Chen, G.-D., et al. (2010). Exploring the possibility of using humanoid robots as instructional tools for teaching a second language in primary school. *Educational Technology & Society*, *13*(2), 13–24.

Cheng, L.-P., Liang, H.-S., Wu, C.-Y., & Chen, M. Y. (2013). igrasp: grasp-based adaptive keyboard for mobile devices. In *Proceedings of chi conference* (pp. 3037–3046).

Chidambaram, V., Chiang, Y.-H., & Mutlu, B. (2012). Designing persuasive robots: how robots might persuade people using vocal and nonverbal cues. In *Proceedings of the seventh annual acm/ieee international conference on human-robot interaction* (pp. 293–300).

Christianson, S.-A. (2014). *The handbook of emotion and memory: Research and theory.* Psychology Press.

Cole, P. M., Zahn-Waxler, C., Fox, N. A., Usher, B. A., & Welsh, J. D. (1996). Individual differences in emotion regulation and behavior problems in preschool children. *Journal of Abnormal Psychology*, *105*(4), 518.

Comblain, C., D'Argembeau, A., & Van der Linden, M. (2005). Phenomenal characteristics of autobiographical memories for emotional and neutral events in older and younger adults. *Experimental Aging Research*, *31*(2), 173–189.

Conati, C. (2002). Probabilistic assessment of user's emotions in educational games. *Applied Artificial Intelligence*, *16*(7-8), 555–575.

Coninx, A., Baxter, P., Oleari, E., Bellini, S., Bierman, B., Henkemans, O. B., ... others (2016). Towards long-term social child-robot interaction: using multi-activity switching to engage young users. *Journal of Human-Robot Interaction*, *5*(1), 32–67.

Constine, J. (2016). *How facebook news feed works.* Retrieved from https://techcrunch.com/2016/09/06/ultimate-guide-to-the-news-feed/

Cooney, M., Kanda, T., Alissandrakis, A., & Ishiguro, H. (2014). Designing enjoyable motion-based play interactions with a small humanoid robot. *International Journal of Social Robotics*, *6*(2), 173–193.

Cuadrado, L.-E. I., Riesco, Á. M., & López, F. D. L. P. (2016). Artie: An integrated environment for the development of affective robot tutors. *Frontiers in Computational Neuroscience*, *10*.

Dang, T.-H.-H., & Tapus, A. (2015). Stress game: The role of motivational robotic assistance in reducing users task stress. *International Journal of Social Robotics*, *7*(2), 227–240.

D'Argembeau, A., Comblain, C., & Van der Linden, M. (2003). Phenomenal characteristics of autobiographical memories for positive, negative, and neutral events. *Applied Cognitive Psychology*, *17*(3), 281–294.

Dautenhahn, K. (1998). The art of designing socially intelligent agents: Science, fiction, and the human in the loop. *Applied artificial intelligence*, *12*(7-8), 573–617.

Dautenhahn, K. (2014). Social robots as companions: Challenges and opportunities. *Sociable Robots and the Future of Social Relations: Proceedings of Robo-Philosophy 2014*, *273*, 9.

Dautenhahn, K., & Werry, I. (2002). A quantitative technique for analysing robot-human interactions. In *Intelligent robots and systems, 2002. ieee/rsj international conference on* (Vol. 2, pp. 1132–1138).

Dautenhahn, K., Woods, S., Kaouri, C., Walters, M. L., Koay, K. L., & Werry, I. (2005). What is a robot companion-friend, assistant or butler? In *Intelligent robots and systems, 2005.(iros 2005). 2005 ieee/rsj international conference on* (pp. 1192–1197).

de Greeff, J., & Belpaeme, T. (2015). Why robots should be social: Enhancing machine learning through social human-robot interaction. *PLoS one*, *10*(9), e0138061.

de Haas, M., Vogt, P., & Krahmer, E. (2016). Enhancing child-robot tutoring interactions with appropriate feedback. In *Proceedings of first workshop on long-term child-robot interaction. ieee ro-man,(new york, ny).*

De Smedt, J. (2015). *Natural language processing research with nao robots* (Unpublished doctoral dissertation). Universiteit Antwerpen.

Dorça, F. (2015). Implementation and use of simulated students for test and validation of new adaptive educational systems: A practical insight. *International Journal of Artificial Intelligence in Education*, *25*(3), 319–345.

dos Santos, C. T., & Osório, F. S. (2004). An intelligent and adaptive virtual environment and its application in distance learning. In *Proceedings of the working conference on advanced visual interfaces* (pp. 362–365).

Downe-Wamboldt, B. (1992). Content analysis: method, applications, and issues. *Health care for women international*, *13*(3), 313–321.

Druin, A. (2002). The role of children in the design of new technology. *Behaviour and information technology*, *21*(1), 1–25.

Druin, A., Bederson, B., Boltman, A., Miura, A., Knotts-Callahan, D., & Platt, M. (1998). Children as our technology design partners+.

Duck, S. (2007). *Human relationships*. Sage.

Ekman, P. (1980). *The face of man: Expressions of universal emotions in a new guinea village*. Scholarly Title.

Emmeche, C. (2014). Robot friendship: Can a robot be a friend? *International Journal of Signs and Semiotic Systems (IJSSS)*, *3*(2), 26–42.

Endsley, M. R. (1999). Level of automation effects on performance, situation awareness and workload in a dynamic control task. *Ergonomics*, *42*(3), 462–492.

EVP, U. (2015). *Emotion detection algorithm based on opencv.* https://github.com/liy9393/emotion-detection. (Accessed: 2016-05-08)

Fagin, B. S., & Merkle, L. (2002). Quantitative analysis of the effects of robots on introductory computer science education. *Journal on Educational Resources in Computing (JERIC)*, *2*(4), 2.

Feingold, A. (1992). Good-looking people are not what we think. *Psychological bulletin*, *111*(2), 304.

Feldman Barrett, L., & Russell, J. A. (1998). Independence and bipolarity in the structure of current affect. *Journal of personality and social psychology*, *74*(4), 967.

Fink, J., Mubin, O., Kaplan, F., & Dillenbourg, P. (2011). Roomba is not a robot; aibo is still alive! anthropomorphic language in online forums. In *3rd international conference on social robotics, icsr 2011.*

Fischer, G. (2001). User modeling in human–computer interaction. *User modeling and user-adapted interaction*, *11*(1-2), 65–86.

Fong, T., Nourbakhsh, I., & Dautenhahn, K. (2003). A survey of socially interactive robots. *Robotics and autonomous systems*, *42*(3), 143–166.

Forlizzi, J. (2007). How robotic products become social products: an ethnographic study of cleaning in the home. In *Proceedings of the acm/ieee international conference on human-robot interaction* (pp. 129–136).

Forlizzi, J., & DiSalvo, C. (2006). Service robots in the domestic environment: a study of the roomba vacuum in the home. In *Proceedings of the 1st acm sigchi/sigart conference on human-robot interaction* (pp. 258–265).

Fortunati, L., Esposito, A., & Lugano, G. (2015). *Introduction to the special issue beyond industrial robotics: social robots entering public and domestic spheres.* Taylor & Francis.

François, D., Polani, D., & Dautenhahn, K. (2008). Towards socially adaptive robots: A novel method for real time recognition of human-robot interaction styles. In *Humanoid robots, 2008. humanoids 2008. 8th ieee-ras international conference on* (pp. 353–359).

Franklin, D. J. (2013). Adaptive interaction design for online mathematics education: The way of the game.

Gao, A. Y., Barendregt, W., & Castellano, G. (2017). Personalised human-robot co-adaptation in instructional settings using reinforcement learning. In *Persuasive embodied agents for behavior change (peach2017) workshop at the international conference on intelligent virtual agents (iva2017).*

Gifford, T., SRINIVASAN, S., KAUR, M., DOTOV, D., WANAMAKER, C., DRESSLER, G., ... BHAT, A. (2011). Using robots to teach musical rhythms to typically developing children and children with autism. *University of Connecticut*.

Goldfield, B. A., & Reznick, J. S. (1990). Early lexical acquisition: Rate, content, and the vocabulary spurt. *Journal of child language*, *17*(01), 171–183.

Gonzalez-Pacheco, V., Ramey, A., Alonso-Martín, F., Castro-Gonzalez, A., & Salichs, M. A. (2011). Maggie: A social robot as a gaming platform. *International Journal of Social Robotics*, *3*(4), 371–381.

Gordon, G., Spaulding, S., Westlund, J. K., Lee, J. J., Plummer, L., Martinez, M., ... Breazeal, C. (2016). Affective personalization of a social robot tutor for children's second language skills. In *Aaai* (pp. 3951–3957).

Görer, B., Salah, A. A., & Akın, H. L. (2013). A robotic fitness coach for the elderly. In *Ambient intelligence* (pp. 124–139). Springer.

Graf, S., Liu, T.-C., & Kinshuk, K. (2008). Interactions between students learning styles, achievement and behaviour in mismatched courses. In *Proceedings of the international conference on cognition and exploratory learning in digital age (celda 2008). iadis international conference* (pp. 223–230).

Gross, H.-M., Mueller, S., Schroeter, C., Volkhardt, M., Scheidig, A., Debes, K., ... Doering, N. (2015). Robot companion for domestic health assistance: Implementation, test and case study under everyday conditions in private apartments. In *Intelligent robots and systems (iros), 2015 ieee/rsj international conference on* (pp. 5992–5999).

Gross, H.-M., Schroeter, C., Mueller, S., Volkhardt, M., Einhorn, E., Bley, A., ... Merten, M. (2011). Progress in developing a socially assistive mobile home robot companion for the elderly with mild cognitive impairment. In *Intelligent robots and systems (iros), 2011 ieee/rsj international conference on* (pp. 2430–2437).

Haas, M. d., Baxter, P., de Jong, C., Krahmer, E., & Vogt, P. (2017). Exploring different types of feedback in preschooler and robot interaction. In *Proceedings of the companion of the 2017 acm/ieee international conference on human-robot interaction* (pp. 127–128).

Haider, M., Sinha, A., & Chaudhary, B. (2010). An investigation of relationship between learning styles and performance of learners. *International Journal of Engineering Science and Technology*, *2*(7), 2813–2819.

Haring, K. S., Mougenot, C., Ono, F., & Watanabe, K. (2014). Cultural differences in perception and attitude towards robots. *International Journal of Affective Engineering*, *13*(3), 149–157.

Häring, M., Bee, N., & André, E. (2011). Creation and evaluation of emotion expression with body movement, sound and eye color for humanoid robots. In *2011 ro-man* (pp. 204–209).

Hassani, K., Nahvi, A., & Ahmadi, A. (2016). Design and implementation of an intelligent virtual environment for improving speaking and listening skills. *Interactive Learning Environments*, *24*(1), 252–271.

Hastie, H., Lim, M. Y., Janarthanam, S., Deshmukh, A., Aylett, R., Foster, M. E., & Hall, L. (2016). I remember you!: Interaction with memory for an empathic virtual robotic tutor. In *Proceedings of the 2016 international conference on autonomous agents & multiagent systems* (pp. 931–939).

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research*, *77*(1), 81–112.

Hayes, C. J., & Riek, L. D. (2014). Establishing human personality metrics for adaptable robots during learning tasks. In *Aaai fall symposium on artificial intelligence and human robot interaction., https://www. aaai. org/ocs/index. php/fss/fss14/rt/capturecite/9213/0.*

Hays, R. T. (2005). *The effectiveness of instructional games: A literature review and discussion* (Tech. Rep.). DTIC Document.

Heerink, M., Krose, B., Evers, V., & Wielinga, B. (2007). Observing conversational expressiveness of elderly users interacting with a robot and screen agent. In *Rehabilitation robotics, 2007. icorr 2007. ieee 10th international conference on* (pp. 751–756).

Heerink, M., Kröse, B., Evers, V., & Wielinga, B. (2010). Assessing acceptance of assistive social agent technology by older adults: the almere model. *International journal of social robotics*, *2*(4), 361–375.

Ho, W. C., Dautenhahn, K., Lim, M. Y., Vargas, P. A., Aylett, R., & Enz, S. (2009). An initial memory model for virtual and robot companions supporting migration and long-term interaction. In *Robot and human interactive communication, 2009. ro-man 2009. the 18th ieee international symposium on* (pp. 277–284).

Hoffman, G., & Breazeal, C. (2007). Effects of anticipatory action on human-robot teamwork efficiency, fluency, and perception of team. In *Proceedings of the acm/ieee international conference on human-robot interaction* (pp. 1–8).

Huber, A., Lammer, L., Weiss, A., & Vincze, M. (2014). Designing adaptive roles for socially assistive robots: a new method to reduce technological determinism and role stereotypes. *Journal of Human-Robot Interaction*, *3*(2), 100–115.

Hudlicka, E. (2017). Computational modeling of cognition–emotion interactions: Theoretical and practical relevance for behavioral healthcare. In *Emotions and affect in human factors and human-computer interaction* (pp. 383–436). Elsevier.

Huschilt, J., & Clune, L. (2012). The use of socially assistive robots for dementia care. *Journal of gerontological nursing*, *38*(10), 15–19.

Indico. (2016). *Facial emotion recognition.* https://indico.io/docs. (Accessed: 2016-06-08)

Isbister, K., & Nass, C. (2000). Consistency of personality in interactive characters: verbal cues, non-verbal cues, and user characteristics. *International journal of human-computer studies*, *53*(2), 251–267.

Janssen, D. (1997). Instructional design models for well-structured and ill-structured problem solving learning outcomes. *Educational Technology Research & Development*, *45*(1), 45–94.

Janssen, J. B., van der Wal, C. C., Neerincx, M. A., & Looije, R. (2011). Motivating children to learn arithmetic with an adaptive robot game. In *International conference on social robotics* (pp. 153–162).

Jeon, M. (2017). Emotions and affect in human factors and human–computer interaction: Taxonomy, theories, approaches, and methods. In *Emotions and affect in human factors and human-computer interaction* (pp. 3–26). Elsevier.

Jimenez, F., Yoshikawa, T., Furuhashi, T., & Kanoh, M. (2015). An emotional expression model for educational-support robots. *Journal of Artificial Intelligence and Soft Computing Research*, *5*(1), 51–57.

Jones, A., Bull, S., & Castellano, G. (2017). i know that now, im going to learn this next promoting self-regulated learning with a robotic tutor. *International Journal of Social Robotics*, 1–16.

Jones, A., & Castellano, G. (2018). Adaptive robotic tutors that support self-regulated learning: A longer-term investigation with primary school children. *International Journal of Social Robotics*, 1–14.

Kachouie, R., Sedighadeli, S., Khosla, R., & Chu, M.-T. (2014). Socially assistive robots in elderly care: a mixed-method systematic literature review. *International Journal of Human-Computer Interaction*, *30*(5), 369–393.

Kaelbling, L. P., Littman, M. L., & Moore, A. W. (1996). Reinforcement learning: A survey. *Journal of artificial intelligence research*, *4*, 237–285.

Kanda, T., Hirano, T., Eaton, D., & Ishiguro, H. (2004). Interactive robots as social partners and peer tutors for children: A field trial. *Human-computer interaction*, *19*(1), 61–84.

Kanda, T., Sato, R., Saiwaki, N., & Ishiguro, H. (2007). A two-month field trial in an elementary school for long-term human–robot interaction. *Robotics, IEEE Transactions on*, *23*(5), 962–971.

Kanda, T., Shiomi, M., Miyashita, Z., Ishiguro, H., & Hagita, N. (2009). An affective guide robot in a shopping mall. In *Human-robot interaction (hri), 2009 4th acm/ieee international conference on* (pp. 173–180).

Kanda, T., Shiomi, M., Miyashita, Z., Ishiguro, H., & Hagita, N. (2010). A communication robot in a shopping mall. *IEEE Transactions on Robotics*, *26*(5), 897–913.

Karapanos, E. (2013). User experience over time. In *Modeling users' experiences with interactive systems* (pp. 57–83). Springer.

Kasap, Z., & Magnenat-Thalmann, N. (2010). Towards episodic memory-based long-term affective interaction with a human-like robot. In *Ro-man, 2010 ieee* (pp. 452–457).

Kato, Y., Kanda, T., & Ishiguro, H. (2015). May i help you?: Design of human-like polite approaching behavior. In *Proceedings of the tenth annual acm/ieee international conference on human-robot interaction* (pp. 35–42).

Kay, J. S., & Moss, J. G. (2012). Using robots to teach programming to k-12 teachers. In *Frontiers in education conference (fie), 2012* (pp. 1–6).

Keizer, S., Ellen Foster, M., Wang, Z., & Lemon, O. (2014). Machine learning for social multiparty human–robot interaction. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, *4*(3), 14.

Kelsey, E., Ray, F., Brown, D., & Robson, R. (2015). Design of a domain-independent, interactive, dialogue-based tutor for use within the gift framework. In *Generalized intelligent framework for tutoring (gift) users symposium (giftsym3)* (p. 161).

Kennedy, J., Baxter, P., & Belpaeme, T. (2015). The robot who tried too hard: social behaviour of a robot tutor can negatively affect child learning. In *Proceedings of the tenth annual acm/ieee international conference on human-robot interaction* (pp. 67–74).

Kennedy, J., Baxter, P., Senft, E., & Belpaeme, T. (2016). Social robot tutoring for child second language learning. In *Human-robot interaction (hri), 2016 11th acm/ieee international conference on* (pp. 231–238).

Khosla, R., Nguyen, K., & Chu, M.-T. (2016). Human robot engagement and acceptability in residential aged care. *International Journal of Human-Computer Interaction*(just-accepted).

Kidd, C. D., & Breazeal, C. (2008). Robots at home: Understanding long-term human-robot interaction. In *Intelligent robots and systems, 2008. iros 2008. ieee/rsj international conference on* (pp. 3230–3235).

King, M. A., & Yuille, J. C. (1987). *Suggestibility and the child witness*. Springer.

Klinkenberg, S., Straatemeier, M., & Van der Maas, H. (2011). Computer adaptive practice of maths ability using a new item response model for on the fly ability and difficulty estimation. *Computers & Education*, *57*(2), 1813–1824.

Komatsubara, T., Shiomi, M., Kanda, T., Ishiguro, H., & Hagita, N. (2014). Can a social robot help children's understanding of science in classrooms? In *Proceedings of the second international conference on human-agent interaction* (pp. 83–90).

Kory, J., & Breazeal, C. (2014). Storytelling with robots: Learning companions for preschool children's language development. In *Robot and human interactive communication, 2014 ro-man: The 23rd ieee international symposium on* (pp. 643–648).

Kory Westlund, J., Gordon, G., Spaulding, S., Lee, J. J., Plummer, L., Martinez, M., ... Breazeal, C. (2016). Lessons from teachers on performing hri studies with young children in schools. In *The eleventh acm/ieee international conference on human robot interation* (pp. 383–390).

Kozima, H., Michalowski, M. P., & Nakagawa, C. (2009). Keepon. *International Journal of Social Robotics*, *1*(1), 3–18.

Kühnlenz, B., Sosnowski, S., Buß, M., Wollherr, D., Kühnlenz, K., & Buss, M. (2013). Increasing helpfulness towards a robot by emotional adaption to the user. *International Journal of Social Robotics*, *5*(4), 457–476.

Kuo, I. H., Rabindran, J. M., Broadbent, E., Lee, Y. I., Kerse, N., Stafford, R., & MacDonald, B. A. (2009). Age and gender factors in user acceptance of healthcare robots. In *Robot and human interactive communication, 2009. ro-man 2009. the 18th ieee international symposium on* (pp. 214–219).

Kwon, O.-H., Koo, S.-Y., Kim, Y.-G., & Kwon, D.-S. (2010). Telepresence robot system for english tutoring. In *2010 ieee workshop on advanced robotics and its social impacts* (pp. 152–155).

Lawhead, P. B., Duncan, M. E., Bland, C. G., Goldweber, M., Schep, M., Barnes, D. J., & Hollingsworth, R. G. (2002). A road map for teaching introductory programming using lego© mindstorms robots. In *Acm sigcse bulletin* (Vol. 35 - 2, pp. 191–201).

LeDoux, J. (2007). *Emotional memory. scholarpedia 2 (7): 1806.*

LeDoux, J. E. (1993). Emotional memory systems in the brain. *Behavioural brain research*, *58*(1-2), 69–79.

Lee, Kiesler, S., Forlizzi, J., & Rybski, P. (2012). Ripple effects of an embedded social agent: a field study of a social robot in the workplace. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 695–704).

Lee, J.-H., Yakushin, D., Renteria, F., Nakata, K., Sugano, A., Morita, K., & Yamazoe, H. (2015). A competitive and cooperative humanoid software developing scheme loe and its improvements. In *Humanoid robots (humanoids), 2015 ieee-ras 15th international conference on* (pp. 631–636).

Lee, M. K., & Forlizzi, J. (2009). Designing adaptive robotic services. *Proc. of IASDR09*.

Lee, M. K., Forlizzi, J., Kiesler, S., Rybski, P., Antanitis, J., & Savetsila, S. (2012). Personalization in hri: A longitudinal field experiment. In *Human-robot interaction (hri), 2012 7th acm/ieee international conference on* (pp. 319–326).

Lee, S., Noh, H., Lee, J., Lee, K., Lee, G. G., Sagong, S., & Kim, M. (2011). On the effectiveness of robot-assisted language learning. *ReCALL*, *23*(1), 25–58.

Leemkuil, H., & De Jong, T. (2012). Adaptive advice in learning with a computer-based knowledge management simulation game. *Academy of management learning & education*, *11*(4), 653–665.

Leite, I. (2013a). *Long-term interactions with empathic social robots* (Unpublished doctoral dissertation). Ph.D. UNIVERSIDADE TECNICA DE LISBOA INSTITUTO SUPERIOR TECNICO.

Leite, I. (2013b). *Long-term interactions with empathic social robots* (Unpublished doctoral dissertation). Instituto Superior Tcnico, Universidade de Lisboa.

Leite, I. (2015). Long-term interactions with empathic social robots. *AI Matters*, *1*(3), 13–15.

Leite, I., Castellano, G., Pereira, A., Martinho, C., & Paiva, A. (2014). Empathic robots for long-term interaction. *International Journal of Social Robotics*, *6*(3), 329–341.

Leite, I., Martinho, C., & Paiva, A. (2013). Social robots for long-term interaction: a survey. *International Journal of Social Robotics*, *5*(2), 291–308.

Leite, I., Pereira, A., & Lehman, J. F. (2017). Persistent memory in repeated child-robot conversations. In *Proceedings of the 2017 conference on interaction design and children* (pp. 238–247).

Lever, J. (1976). Sex differences in the games children play. *Social problems*, *23*(4), 478–487.

Levine, L. J., & Burgess, S. L. (1997). Beyond general arousal: Effects of specific emotions on memory. *Social Cognition*, *15*(3), 157.

Levine, L. J., & Pizarro, D. A. (2004). Emotion and memory research: A grumpy overview. *Social cognition*, *22*(5: Special issue), 530–554.

Levinger, G. (1983). Development and change. *Close relationships*, 315–359.

Leyzberg, D., Spaulding, S., & Scassellati, B. (2014). Personalizing robot tutors to individuals' learning differences. In *Proceedings of the 2014 acm/ieee international conference on human-robot interaction* (pp. 423–430).

Liu, J., Wong, C. K., & Hui, K. K. (2003). An adaptive user interface based on personalized learning. *IEEE Intelligent Systems*, *18*(2), 52–57.

Liu, P., Glas, D. F., Kanda, T., Ishiguro, H., & Hagita, N. (2016). A model for generating socially-appropriate deictic behaviors towards people. *International Journal of Social Robotics*, 1–17.

Liu, T.-C., Graf, S., et al. (2009). Coping with mismatched courses: students behaviour and performance in courses mismatched to their learning styles. *Educational Technology Research and Development*, *57*(6), 739.

Lubold, N., Walker, E., & Pon-Barry, H. (2016). Effects of voice-adaptation and social dialogue on perceptions of a robotic learning companion. In *The eleventh acm/ieee international conference on human robot interaction* (pp. 255–262).

Lyra, O., Karapanos, E., Gouveia, R., Nisi, V., & Nunes, N. J. (2013). Engaging children in longitudinal behavioral studies through playful technologies. In *Proceedings of the 12th international conference on interaction design and children* (pp. 396–399).

Ma, X., Yang, X., Zhao, S., Fu, C.-W., Lan, Z., & Pu, Y. (2014). Using social media platforms for human-robot interaction in domestic environment. *International Journal of Human-Computer Interaction*, *30*(8), 627–642.

Mahajan, A., & Teneketzis, D. (2008). Multi-armed bandit problems. In *Foundations and applications of sensor management* (pp. 121–151). Springer.

Mahani, M., & Eklundh, K. S. (2009). A survey of the relation of the task assistance of a robot to its social role. *Communication KCSa Royal Institute of Technology: Stockholm, Sweden*.

Malone, T. W., & Lepper, M. R. (1987). Making learning fun: A taxonomy of intrinsic motivations for learning. *Aptitude, learning, and instruction*, *3*(1987), 223–253.

Matarić, M. (2014). Socially assistive robotics: human-robot interaction methods for creating robots that care. In *Proceedings of the 2014 acm/ieee international conference on human-robot interaction* (pp. 333–333).

McClintock, C. G., & Nuttin Jr, J. M. (1969). Development of competitive game behavior in children across two cultures. *Journal of Experimental Social Psychology*, *5*(2), 203–218.

McColl, D., & Nejat, G. (2013). Meal-time with a socially assistive robot and older adults at a long-term care facility. *Journal of Human-Robot Interaction*, *2*(1), 152–171.

McCullough, J. A., Cunningham, L. A., & Wilson, R. H. (1992). Auditory/visual word-identification test materials: Computer application with children. *Journal of the American Academy of Audiology*, *3*(3), 208–214.

Meho, L. I., & Yang, K. (2007). Impact of data sources on citation counts and rankings of lis faculty: Web of science versus scopus and google scholar. *Journal of the Association for Information Science and Technology*, *58*(13), 2105–2125.

Mileounis, A., Cuijpers, R. H., & Barakova, E. I. (2015). Creating robots with personality: The effect of personality on social intelligence. In *Artificial computation in biology and medicine* (pp. 119–132). Springer.

Miller, G. A., & Gildea, P. M. (1987). How children learn words. *Scientific American*, *257*(3), 94–99.

Mitrović, A., & Djordjević-Kajan, S. (1995). Interactive reconstructive student modeling: A machine-learning approach. *International Journal of Human-Computer Interaction*, *7*(4), 385–401.

Miyamoto, E., Lee, M., Fujii, H., & Okada, M. (2005). How can robots facilitate social interaction of children with autism?: Possible implications for educational environments.

Mordvintsev, A., & K, A. (2013). *Opencv-python tutorials.* http://opencv-python-tutroals.readthedocs.io. (Accessed: 2016-10-24)

Moskovsky, C., & Alrabai, F. (2009). Intrinsic motivation in saudi learners of english as a foreign language. *The Open Applied Linguistics Journal*, *2*(1).

Mubin, O. (2011). *Roila: Robot interaction language* (Unpublished doctoral dissertation). Technische Universiteit Eindhoven.

Mubin, O., Khan, A., & Obaid, M. (2016). # naorobot: exploring nao discourse on twitter. In *Proceedings of the 28th australian conference on computer-human interaction* (pp. 155–159).

Mubin, O., Stevens, C. J., Shahid, S., Al Mahmud, A., & Dong, J.-J. (2013). A review of the applicability of robots in education. *Journal of Technology in Education and Learning*, *1*, 209–0015.

Mutlu, B., Osman, S., Forlizzi, J., Hodgins, J., & Kiesler, S. (2006). Perceptions of asimo: an exploration on co-operation and competition with humans and humanoid robots. In *Proceedings of the 1st acm sigchi/sigart conference on human-robot interaction* (pp. 351–352).

Nelson, N. L., & Russell, J. A. (2016). A facial expression of pax: Assessing childrens recognition of emotion from faces. *Journal of experimental child psychology*, *141*, 49–64.

Nwana, H. S. (1996). Software agents: An overview. *The knowledge engineering review*, *11*(03), 205–244.

Obaid, M., Barendregt, W., Alves-Oliveira, P., Paiva, A., & Fjeld, M. (2015). Designing robotic teaching assistants: Interaction design students and childrens views. In *International conference on social robotics* (pp. 502–511).

Okita, S. Y., Ng-Thow-Hing, V., & Sarvadevabhatla, R. K. (2011). Multimodal approach to affective human-robot interaction design with children. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, *1*(1), 5.

Omar Mubin, C. B. (2015). *Vocabulary.* http://roila.org/language-guide/vocabulary/. (Accessed 23-September-2015)

Orbuch, T., & Sprecher., S. (2006). *Attraction and interpersonal relationships. in john delamater, editor, handbook of social psychology, handbooks of sociology and social research.* Springer.

Orlando, J. (2014). Teachers changing practices with information and communication technologies: an up-close, longitudinal analysis. *Research in learning technology*, *22*.

Outtagarts, A. (2009). Mobile agent-based applications: A survey. *International Journal of Computer Science and Network Security*, *9*(11), 331–339.

Pasch, M. (2010). Improving children's self-report in user-centered evaluations. In *Proceedings of the 9th international conference on interaction design and children* (pp. 331–334).

Pekrun, R., & Linnenbrink-Garcia, L. (2012). Academic emotions and student engagement. In *Handbook of research on student engagement* (pp. 259–282). Springer.

Pittermann, J., Pittermann, A., & Minker, W. (2010). Emotion recognition and adaptation in spoken dialogue systems. *International Journal of Speech Technology*, *13*(1), 49–60.

Prain, V., Cox, P., Deed, C., Dorman, J., Edwards, D., Farrelly, C., . . . others (2013). Personalised learning: Lessons to be learnt. *British Educational Research Journal*, *39*(4), 654–676.

Prensky, M., & Berry, B. D. (2001). Do they really think differently. *On the horizon*, *9*(6), 1–9.

Rajagopalan, S. S., Murthy, O. R., Goecke, R., & Rozga, A. (2015). Play with me-measuring a child's engagement in a social interaction. In *Automatic face and gesture recognition (fg), 2015 11th ieee international conference and workshops on* (Vol. 1, pp. 1–8).

Ramachandran, A., & Scassellati, B. (2014). Adapting difficulty levels in personalized robot-child tutoring interactions. In *Proceedings of the workshops at the 28th aaai conference on artificial intelligence.*

Ramachandran, A., & Scassellati, B. (2015). Fostering learning gains through personalized robot-child tutoring interactions. In *Proceedings of the tenth annual acm/ieee international conference on human-robot interaction extended abstracts* (pp. 193–194).

Rau, P. P., Li, Y., & Li, D. (2009). Effects of communication style and culture on ability to accept recommendations from robots. *Computers in Human Behavior*, *25*(2), 587–595.

Rieber, L. P. (1996). Seriously considering play: Designing interactive learning environments based on the blending of microworlds, simulations, and games. *Educational technology research and development*, *44*(2), 43–58.

Rinaldi, D., & Ispita, A. (2012). Teaching vocabulary by using capsule vocabularystrategy at fourth grade of elementary school. *e-Jurnal Mahasiswa Prodi Pend Bahasa Inggris 2012*, *1*(4).

Ritschel, H., & André, E. (2017). Real-time robot personality adaptation based on reinforcement learning and social signals. In *Proceedings of the companion of the 2017 acm/ieee international conference on human-robot interaction* (pp. 265–266).

Robins, B., Dautenhahn, K., Te Boekhorst, R., & Nehaniv, C. L. (2008). Behaviour delay and robot expressiveness in child-robot interactions: a user study on interaction kinesics. In *Proceedings of the 3rd acm/ieee international conference on human robot interaction* (pp. 17–24).

Robinson, H., MacDonald, B., & Broadbent, E. (2014). The role of healthcare robots for older people at home: a review. *International Journal of Social Robotics*, *6*(4), 575–591.

Ros, R., Baroni, I., & Demiris, Y. (2014). Adaptive human–robot interaction in sensorimotor task instruction: From human to robot dance tutors. *Robotics and Autonomous Systems*, *62*(6), 707–720.

Rosenthal-von der Pütten, A. M., Straßmann, C., & Krämer, N. C. (2016). Robots or agents–neither helps you more or less during second language acquisition. In *International conference on intelligent virtual agents* (pp. 256–268).

Rousseau, V., Ferland, F., Létourneau, D., & Michaud, F. (2013). Sorry to interrupt, but may i have your attention? preliminary design and evaluation of autonomous engagement in hri. *Journal of Human-Robot Interaction*, *2*(3), 41–61.

Roy, S., Kieson, E., Abramson, C., & Crick, C. (2018). Using human reinforcement learning models to improve robots as teachers. In *Companion of the 2018 acm/ieee international conference on human-robot interaction* (pp. 225–226).

Russell, J. A. (2017). Cross-cultural similarities and differences in affective processing and expression. In *Emotions and affect in human factors and human-computer interaction* (pp. 123–141). Elsevier.

Saerbeck, M., Schut, T., Bartneck, C., & Janse, M. D. (2010). Expressive robots in education: varying the degree of social supportive behavior of a robotic tutor. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 1613–1622).

Salter, T., Dautenhahn, K., & Bockhorst, R. (2004). Robots moving out of the laboratory-detecting interaction levels and human contact in noisy school environments. In *Robot and human interactive communication, 2004. roman 2004. 13th ieee international workshop on* (pp. 563–568).

Salter, T., Michaud, F., Létourneau, D., Lee, D., & Werry, I. P. (2007). Using proprioceptive sensors for categorizing human-robot interactions. In *Human-robot interaction (hri), 2007 2nd acm/ieee international conference on* (pp. 105–112).

Sampayo-Vargas, S., Cope, C. J., He, Z., & Byrne, G. J. (2013). The effectiveness of adaptive difficulty adjustments on students' motivation and learning in an educational computer game. *Computers & Education*, *69*, 452–462.

Scholar, G. (2017, apr). *Top publications - robotics.* https://scholar.google.es/citations.

Sekmen, A., & Challa, P. (2013). Assessment of adaptive human–robot interactions. *Knowledge-Based Systems*, *42*, 49–59.

Serholt, S., & Barendregt, W. (2014). Students attitudes towards the possible future of social robots in education..

Serholt, S., & Barendregt, W. (2016). Robots tutoring children: Longitudinal evaluation of social engagement in child-robot interaction. In *Proceedings of the 9th nordic conference on human-computer interaction* (p. 64).

Serholt, S., Barendregt, W., Leite, I., Hastie, H., Jones, A., Paiva, A., ... Castellano, G. (2014). Teachers' views on the use of empathic robotic tutors in the classroom. In *Robot and human interactive communication, 2014 ro-man: The 23rd ieee international symposium on* (pp. 955–960).

Shahid, S., Krahmer, E., & Swerts, M. (2008). Alone or together: Exploring the effect of physical co-presence on the emotional expressions of game playing children across cultures. In *Fun and games* (pp. 94–105). Springer.

Shahid, S., Krahmer, E., & Swerts, M. (2014). Child–robot interaction across cultures: How does playing a game with a social robot compare to playing a game alone or with a friend? *Computers in Human Behavior*, *40*, 86–100.

Sharan, S., & Tan, I. G. C. (2008). Student engagement in learning. *Organizing schools for productive learning*, 41–45.

Shiomi, M., Iio, T., Kamei, K., Sharma, C., & Hagita, N. (2015). Effectiveness of social behaviors for autonomous wheelchair robot to support elderly people in japan. *PloS one*, *10*(5), e0128031.

Shiomi, M., Shinozawa, K., Nakagawa, Y., Miyashita, T., Sakamoto, T., Terakubo, T., ... Hagita, N. (2013). Recommendation effects of a social robot for advertisement-use context in a shopping mall. *International Journal of Social Robotics*, *5*(2), 251–262.

Sidner, C., Kidd, C., Lee, C., & Lesh, N. (2004). Where to look: A study of human-robot interaction. In *Intelligent user interfaces conference* (pp. 78–84).

Sidner, C. L., & Dzikovska, M. (2002). Human-robot interaction: Engagement between humans and robots for hosting activities. In *Proceedings of the 4th ieee international conference on multimodal interfaces* (p. 123).

Sidner, C. L., Lee, C., Kidd, C. D., Lesh, N., & Rich, C. (2005). Explorations in engagement for humans and robots. *Artificial Intelligence*, *166*(1), 140–164.

Sisbot, E. A., Alami, R., Siméon, T., Dautenhahn, K., Walters, M., & Woods, S. (2005). Navigation in the presence of humans. In *Humanoid robots, 2005 5th ieee-ras international conference on* (pp. 181–188).

Smith, S. M., & Petty, R. E. (1995). Personality moderators of mood congruency effects on cognition: The role of self-esteem and negative mood regulation. *Journal of personality and social psychology*, *68*(6), 1092.

Sottilare, R. A., Brawner, K. W., Goldberg, B. S., & Holden, H. K. (2012). *The generalized intelligent framework for tutoring (gift).* Orlando, FL: US Army Research Laboratory–Human Research & Engineering Directorate (ARL-HRED).

Srinivasa, S. S., Ferguson, D., Helfrich, C. J., Berenson, D., Collet, A., Diankov, R., ... Weghe, M. V. (2010). Herb: a home exploring robotic butler. *Autonomous Robots*, *28*(1), 5.

Stafford, L., Dainton, M., & Haas, S. (2000). Measuring routine and strategic relational maintenance: Scale revision, sex versus gender roles, and the prediction of relational characteristics. *Communications Monographs*, *67*(3), 306–323.

Stafford, R. Q., MacDonald, B. A., Jayawardena, C., Wegner, D. M., & Broadbent, E. (2014). Does the robot have a mind? mind perception and attitudes towards robots predict use of an eldercare robot. *International journal of social robotics*, *6*(1), 17–32.

Stanton, C. (2017). *We all need to forget, even robots.* Retrieved from https://theconversation.com/we-all-need-to-forget-even-robots-81387

Sung, J., Christensen, H. I., & Grinter, R. E. (2009). Robots in the wild: understanding long-term use. In *Proceedings of the 4th acm/ieee international conference on human robot interaction* (pp. 45–52).

Svenstrup, M., Bak, T., Maler, O., Andersen, H. J., & Jensen, O. B. (2008). Pilot study of person robot interaction in a public transit space. In *International conference on research and education in robotics* (pp. 96–106).

Szafir, D., & Mutlu, B. (2012). Pay attention!: designing adaptive agents that monitor and improve user engagement. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 11–20).

Tanaka, F., & Kimura, T. (2009). The use of robots in early education: A scenario based on ethical consideration. In *Ro-man* (pp. 558–560).

Tapus, A. (2009). Improving the quality of life of people with dementia through the use of socially assistive robots. In *Advanced technologies for enhanced quality of life, 2009. at-equal'09.* (pp. 81–86).

Tapus, A., Mataric, M. J., & Scassellati, B. (2007a). Socially assistive robotics. *IEEE Robotics and Automation Magazine*, *14*(1), 35.

Tapus, A., Mataric, M. J., & Scassellati, B. (2007b). Socially assistive robotics [grand challenges of robotics]. *IEEE Robotics & Automation Magazine*, *14*(1), 35–42.

Tapus, A., Ţăpuş, C., & Matarić, M. J. (2008). Userrobot personality matching and assistive robot behavior adaptation for post-stroke rehabilitation therapy. *Intelligent Service Robotics*, *1*(2), 169.

Tielman, M., Neerincx, M., Meyer, J.-J., & Looije, R. (2014). Adaptive emotional expression in robot-child interaction. In *Proceedings of the 2014 acm/ieee international conference on human-robot interaction* (pp. 407–414).

Torrey, C., Powers, A., Fussell, S. R., & Kiesler, S. (2007). Exploring adaptive dialogue based on a robot's awareness of human gaze and task progress. In *Proceedings of the acm/ieee international conference on human-robot interaction* (pp. 247–254).

Torrey, C., Powers, A., Marge, M., Fussell, S. R., & Kiesler, S. (2006). Effects of adaptive robot dialogue on information exchange and social relations. In *Proceedings of the 1st acm sigchi/sigart conference on human-robot interaction* (pp. 126–133).

Trafton, G., Hiatt, L., Harrison, A., Tamborello, F., Khemlani, S., & Schultz, A. (2013). Act-r/e: An embodied cognitive architecture for human-robot interaction. *Journal of Human-Robot Interaction*, *2*(1), 30–55.

Tulving, E., et al. (1972). Episodic and semantic memory. *Organization of memory*, *1*, 381–403.

Uluer, P., Akalın, N., & Köse, H. (2015). A new robotic platform for sign language tutoring. *International Journal of Social Robotics*, *7*(5), 571–585.

Van Lange, P. A., De Bruin, E., Otten, W., & Joireman, J. A. (1997). Development of prosocial, individualistic, and competitive orientations: theory and preliminary evidence. *Journal of personality and social psychology*, *73*(4), 733.

Vollmer, C., Pötsch, F., & Randler, C. (2013). Morningness is associated with better gradings and higher attention in class. *Learning and Individual Differences*, *27*, 167–173.

Vygotsky, L. S. (1980). *Mind in society: The development of higher psychological processes*. Harvard university press.

Wainer, J., Dautenhahn, K., Robins, B., & Amirabdollahian, F. (2014). A pilot study with a novel setup for collaborative play of the humanoid robot kaspar with children with autism. *International journal of social robotics*, *6*(1), 45–65.

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological review*, *92*(4), 548.

Wentzel, K. R. (2002). Are effective teachers like good parents? teaching styles and student adjustment in early adolescence. *Child development*, *73*(1), 287–301.

Westlund, J. M. K., Dickens, L., Jeong, S., Harris, P. L., DeSteno, D., & Breazeal, C. L. (2017). Children use non-verbal cues to learn new words from robots as well as people. *International Journal of Child-Computer Interaction*, *13*, 1–9.

Williams, A. B. (2003). The qualitative impact of using lego mindstorms robots to teach computer engineering. *Education, IEEE Transactions on*, *46*(1), 206.

WILSON, A., & Revkin, S. (n.d.). Cohen, l. ydehaene, s.(2006) an open trial assessment of the number race, an adaptive computer game for remediation of dyscalculia. *Behavioral and Brain Functions*, *2*(20), 1–16.

Wittrock, M. C. (1974). Learning as a generative process 1. *Educational psychologist*, *11*(2), 87–95.

WU, W.-y. (2010, apr 22). *Adaptive user interface.* (US Patent 20,100,097,331)

Wuttke, M. (2014). Pro-active and socially capable pedagogical agents in computer- and web-based trainings. *International Summerschool Computer Science 2014 Proceedings of Summerschool*, *7*, 34.

Yamamoto, A. (2014). *Softbanks pepper robot makes emotional debut in japan.*

Youssef, K., Asano, T., De Silva, P. R. S., & Okada, M. (2016). Sociable dining table: Meaning acquisition exploration in knock-based proto-communication. *International Journal of Social Robotics*, *8*(1), 67–84.

Zaga, C., Lohse, M., Truong, K. P., & Evers, V. (2015). The effect of a robot's social character on children task engagement: peer vs tutor. In *Proceedings of the icsr2015 conference.*

Zhang, A. (2016). *Speech recognition (version 3.4) [software].* https://github.com/Uberi/speech_recognition. (Accessed: 2016-05-08)

Zheng, M., & Meng, M. Q.-H. (2012). Designing gestures with semantic meanings for humanoid robot. In *Robotics and biomimetics (robio), 2012 ieee international conference on* (pp. 287–292).