UNIVERSIDADE TÉCNICA DE LISBOA

INSTITUTO SUPERIOR TÉCNICO

Long-term Interactions with Empathic Social Robots

Iolanda Margarete dos Santos Carvalho Leite

Supervisor: Doctor Carlos António Roque Martinho

Co-Supervisor: Doctor Ana Maria Severino de Almeida e Paiva

Thesis approved in public session to obtain the PhD degree in

Information Systems and Computer Engineering

Jury final classification: Pass with Distinction

Jury

Chairperson: Chairman of the IST Scientific Board

Members of the Committee:

Doctor Vanessa Evers

Doctor Ana Maria Severino de Almeida e Paiva

Doctor Maria Teresa Caeiro Chambel

Doctor Carlos António Roque Martinho

Doctor Rodrigo Martins de Matos Ventura

2013

INSTITUTO SUPERIOR TÉCNICO

UNIVERSIDADE TÉCNICA DE LISBOA

INSTITUTO SUPERIOR TÉCNICO

Long-term Interactions with Empathic Social Robots

Iolanda Margarete dos Santos Carvalho Leite

Supervisor: Doctor Carlos António Roque Martinho

Co-Supervisor: Doctor Ana Maria Severino de Almeida e Paiva

Thesis approved in public session to obtain the PhD degree in

Information Systems and Computer Engineering

Jury final classification: Pass with Distinction

Jury

Chairperson: Chairman of the IST Scientific Board

Members of the Committee:

Doctor Vanessa Evers, Full Professor at University of Twente, Netherlands

Doctor Ana Maria Severino de Almeida e Paiva, Associate Professor with Habilitation at Instituto Superior Técnico, Universidade Técnica de Lisboa

Doctor Maria Teresa Caeiro Chambel, Assistant Professor at Faculdade de Ciências, Universidade de Lisboa

Doctor Carlos António Roque Martinho, Assistant Professor at Instituto Superior Técnico, Universidade Técnica de Lisboa

Doctor Rodrigo Martins de Matos Ventura, Assistant Professor at Instituto Superior Técnico, Universidade Técnica de Lisboa

Funding Institutions

Fundação para a Ciência e Tecnologia

Grant SFRH / BD / 41358 / 2007

2013

Resumo

Esta tese investiga o papel da empatia em robôs sociais. Em particular, o principal objectivo é avaliar um modelo de empatia em interacções de longo prazo entre utilizadores e robôs sociais. O modelo empático proposto recebe uma estimativa do estado afectivo do utilizador e informação sobre a tarefa que o robot está a realizar com o utilizador, e inclui um mecanismo de decisão que permite ao robô escolher, de uma base de dados de estratégias empáticas e de suporte social, o comportamento empático mais apropriado tendo em conta as interacções passadas com aquele utilizador. Ao adaptar o comportamento às preferências de um utilizador em particular, é esperado que a relação entre o robô e aquele utilizador melhore.

O modelo proposto foi desenvolvido segundo uma abordagem iterativa, tendo sido continuamente avaliado e refinado através de vários estudos com utilizadores. Em geral, os resultados revelam uma relação entre empatia e uma maior percepção de presença social, envolvimento, ajuda e motivação em utilizadores que interagiram com o robô várias vezes. Pode portanto concluir-se que robôs empáticos conseguem captar o interesse dos utilizadores por longos períodos de tempo.

Palavras-chave: Robôs Sociais, Interacção de Longo Prazo, Computação Afectiva, Empatia, Suporte Social, Adaptação.

Abstract

In this dissertation, we investigate the role of empathy in social robotic companions. In particular, our aim is to study the effects of an empathic model in long-term interactions between users and social robots. The inputs of the proposed empathic model are the user's affective state and contextual information of the task that the robot is performing with the user. Considering a database of several empathic and supportive behaviours available to display, the model includes a decision-making mechanism that allows the robot to select the most appropriate empathic behaviours based on the history of interactions with that user. By adapting its behaviour to the preferences of a particular user, we expect to improve the relationship between the robot and the user.

The proposed model was developed using an iterative approach, being continuously evaluated and refined through several user studies. Overall, the results raise the possibility of a link between empathy and higher perceived social presence, engagement, help, and self-validation in users who interact with social robots for extended periods of time. This means that empathic robots will be more likely to engage users during long-term interactions.

Keywords: Social Robots, Long-term Interaction, Affective Computing, Empathy, Social Support, Adaptation.

Acknowledgements

First of all, I want to express my gratitude to all the enthusiastic children who took part in the user studies. They definitely made me want to work harder and harder. Since they were so many, I leave here only the names of the schools and chess clubs where they belong: *Clube de Xadrez da Mata de Benfica*, *Clube de Xadrez de Sintra*, *Escola EB1 da Mina* and *Associação Escola 31 de Janeiro*.

I thank my supervisors, Prof. Carlos Martinho and Prof. Ana Paiva, for their extraordinary guidance and continuous encouragement during these years. I also owe sincere gratefulness to my colleagues from GAIPS. Joana Campos, Lara Leite and Sandra Sá deserve a special mention for their support in the user studies and data analysis, Samuel Mascarenhas for the fruitful discussions on empathy, and Francisco Melo for sharing his knowledge on Reinforcement Learning.

I would also like to thank all the members of the LIREC project, in particular Ginevra Castellano and Hamit Soyel for their restless work on the Affect Recognition framework, and Sibylle Enz for her helpful comments on the design of the experiments.

Finally, my warmest thanks go to all my family. To my parents, for their unconditional support, and to my sister, for always giving me her honest and straight to the point opinion on everything. And to my boyfriend André, for being so many things beyond that, and for being able to comfort me even in the most stressful moments.

This work was supported by a scholarship (SFRHBD/41358/2007) granted by *Fundação para a Ciência e Tecnologia* and by EU 7^{th} Framework Program project LIREC (Living with Robots and Interactive Companions) under grant agreement no. 215554.

Contents

1 Introduction **1**

1.1 Motivation . . . 1

1.2 Research Goals . . . 3

1.3 Expected Contributions . . . 4

1.4 Outline . . . 6

2 Theoretical background **7**

2.1 Human-Human Relationships . . . 7

2.2 Human-Pet Relationships . . . 9

2.3 Human-Machine Relationships . . . 11

2.4 Empathy and Social Support . . . 12

2.5 Concluding Remarks . . . 14

3 Related Work **15**

3.1 Long-term Interaction Studies . . . 15

3.1.1 Health Care and Therapy . . . 15

3.1.2 Educational Settings . . . 22

3.1.3 Work Environments and Public Spaces . . . 26

3.1.4 At Home . . . 31

3.1.5 Summary of Results . . . 36

3.2 Empathy in Virtual Agents and Social Robots . . . 36

3.2.1 Computational Models of Empathy . . . 37

3.2.2 Studies with Empathic Agents . . . 40
3.2.3 Discussion . . . 44
3.3 Concluding Remarks . . . 45

4 Towards the Design of an Empathic Robot Companion 47
4.1 Scenario . . . 47
4.1.1 Interaction Cycle . . . 48
4.1.2 Game Engine . . . 49
4.1.3 Animating the iCat . . . 50
4.2 *As Time Goes By*: Exploratory Long-term Study . . . 51
4.2.1 Procedure . . . 52
4.2.2 Measures . . . 53
4.2.3 Results . . . 54
4.2.4 Discussion . . . 58
4.3 *iCat Watching Two Players*: Studying Empathic Behaviours . . . 59
4.3.1 Modelling Empathy: a First Approach . . . 60
4.3.2 Procedure . . . 64
4.3.3 Conditions . . . 65
4.3.4 Measures . . . 67
4.3.5 Quantitative Results . . . 68
4.3.6 Qualitative Results . . . 71
4.3.7 Discussion . . . 73
4.4 Concluding Remarks . . . 74

5 An Empathic Model for Long-term Interaction 77
5.1 Scope, Assumptions and Inputs . . . 78
5.2 Empathic Appraisal . . . 80
5.3 Empathic Action Selection and Adaptation . . . 83
5.3.1 Supportive Strategies . . . 83
5.3.2 User Model . . . 84

5.3.3 Adaptive Decision-Making 85

5.3.4 Recalling Past Interactions 88

5.4 Applying the Empathic Model to the iCat 88

5.4.1 Empathic Appraisal Implementation 90

5.4.2 Adaptive Action Selection Implementation 95

5.5 Concluding Remarks 96

6 Short-Term Evaluation 99

6.1 Method 99

6.1.1 Conditions 99

6.1.2 Procedure 100

6.1.3 Measures 102

6.2 Results 104

6.2.1 Questionnaire 104

6.2.2 Open-Ended Interviews 108

6.3 Discussion 117

7 Long-term User Evaluation 121

7.1 Scenario Refinements 121

7.1.1 Supportive Behaviours 122

7.1.2 Empathic Action Selection and Adaptation 123

7.2 Method 126

7.2.1 Participants 127

7.2.2 Procedure 127

7.2.3 Measures 128

7.3 Results 132

7.3.1 Questionnaires 133

7.3.2 Video Analysis 138

7.3.3 Interviews 142

7.3.4 Interaction Logs 149

7.4 Discussion 150

8 Conclusions **155**

8.1 Design Guidelines 157

8.1.1 Building Artificial Companions 157

8.1.2 Long-term Interaction Studies 160

8.2 Future Work 162

8.3 Selected Publications 163

A Final Long-term Study Questionnaire **167**

A.1 Questionnaire Items 168

A.2 The Questionnaire 170

B Personality Questionnaire **177**

B.1 Questionnaire Items 177

B.2 The Questionnaire 180

Bibliography **189**

List of Figures

3.1 Robots used in the Health Care and Therapy related long-term studies (images used with permission of the authors). 16

3.2 Some of the robots used in the long-term studies in the Education domain (images used with permission of the authors). 25

3.3 Robots used in the long-term studies in Work Environments and Public Spaces (images used with permission of the authors). 27

3.4 Robots used in the long-term studies in home environments (images used with permission of the authors). 33

4.1 User playing with the iCat. 48

4.2 Overview of the interaction cycle. 49

4.3 Degrees of freedom of the iCat robot. 51

4.4 Total percentage of time that each user spent looking at the iCat in the first, second and fifth weeks of interaction (U5 did not attend in the fifth week). 56

4.5 Average percentage of time that users looked at the iCat in the first, second and fifth weeks, broken down by the phases of the game. Since all the exercises had different durations, the results are presented as percentages of the total duration of the exercise. 57

4.6 Overview of the empathic model in this scenario. 60

4.7 Snapshots of the different empathic facial expressions of the iCat. 62

4.8 Users interacting with the iCat. 65
4.9 Mean values of the friendship functions (except intimacy) in the *empathic* and *neutral* conditions. 69
4.10 Mean values of the intimacy assertions in the *empathic* and *neutral* conditions. 70

5.1 Generic empathic framework for long-term interaction. 78
5.2 Overview of the proposed Empathic Model. 79
5.3 Partial view of the model with a timeline that exemplifies how the reward values are updated during the interaction. 85
5.4 Instantiation of the Empathic Model in the iCat scenario. 89
5.5 Mood filter function. 92
5.6 Two different body parametrizations for the iCat's mood (positive mood on the left and negative on the right). 94

6.1 Children playing with the iCat during the study. 102
6.2 Boxplot of the Engagement ratings in the three conditions. 105
6.3 Boxplots of the game result. 106
6.4 Boxplot of the Help ratings in the three conditions. 107
6.5 Boxplot of the Self-Validation ratings in the three conditions. 108
6.6 Children's answers to the question "Do you believe that the iCat recognised how you felt during the game?" in the three groups of users. 110
6.7 Children's answers to the question "Do you believe that the iCat felt emotions similar to the ones you felt?" in the three groups of users. . . . 112

7.1 Cards containing different supportive behaviours displayed by the iCat. . 131
7.2 5-point and 4-point Likert scales with *Smileyometer* [118] used in the questionnaire (translated to English). 133
7.3 Means and 95% confidence intervals for the ratings of the Social Presence dimensions in the 1^{st} and 5^{th} weeks of interaction. 135

7.4 Means and 95% confidence intervals of the Engagement ratings in the 1st and 5th weeks of interaction. 136
7.5 Means and 95% confidence intervals of the Help and Self-Validation ratings in the 1st and 5th weeks of interaction. 137
7.6 Percentage of time that each child (u1,...,u16) spent *looking at the iCat* in every interaction with the robot (s1,...,s5). 139
7.7 Average percentage of time that users spent looking at the iCat after playing their moves, after playing the iCat's moves and during the remaining parts of the game. 140
7.8 Means and 95% confidence intervals of the percentage that children spent looking at the iCat in each interaction session. 142
7.9 Children's answer to the interview question *Do you believe iCat is more similar to an opponent, a teacher, a friend or something else?* in the 1st and 5th weeks. 145
7.10 Overall rankings of the preferred support behaviour categories. 146
7.11 Children's answers to the question "Did you enjoy more the first games, the last ones or was it always the same?" 148

List of Tables

3.1 Summary of the long-term studies in the Health Care and Therapy domains. 17

3.2 Summary of the long-term studies in the domain of Education. 21

3.3 Summary of the long-term studies in Work Environments and Public Spaces. 26

3.4 Summary of the long-term studies in Home environments. 35

4.1 Means of the Social Presence Questionnaire items in the first and fifth week, with the corresponding minimum and maximum values between brackets. The last three columns contain, for each question, the number of subjects who increased, maintained and decreased their ratings from the first to the fifth week. 54

4.2 Description of the possible affective sensations experienced by the iCat, with examples of the verbal comments of the robot for both conditions. In the neutral utterances, the comments only vary depending on "reward" and "punishment" sensations. 64

4.3 Friendship Questionnaire items used in this study (translated to English) with the corresponding means and standard deviations for each item. . . . 66

5.1 Sensations elicited by the *emotivector* model [101]. 81

5.2 Visual and contextual features for the discrimination of user's valence. . . 90

5.3 Mapping between the *emotivector* sensations and animations in the iCat's library. The animations marked with * were slightly modified to recreate the desired state. 93

6.1 Questionnaire items used in this study for each measure, translated to English. 103
6.2 Structure of the interview. 104
6.3 Frequency of the reported affective and motivational states in the three conditions. 111

7.1 Examples of supportive behaviours implemented in the iCat based on the framework of Cutrona et al. [38]. 123
7.2 Summary of the methods and measures of the study. 132
7.3 Means and Standard Deviations of the Social Support questionnaire items obtained in our study (2nd column) and the baseline values (3rd column) obtained by Gordon [60]. The scale ranges from 0 (Strongly Disagree) to 3 (Strongly Agree). 138
7.4 Results of the personality questionnaire. The bold values indicate the most salient personality trait(s) of each child. 139
7.5 Total number of victories, draws and losses of children when playing against the iCat for each interaction session. 142
7.6 Interview questions divided by themes. The questions marked with * were only asked in the final interview. 143
7.7 Correlation coefficient between the supportive behaviour category rankings in the interview and the final values obtained from the adaptation algorithm for each user. The values marked with * are positively correlated. The last column contains children's Extroversion scores obtained in the Personality Questionnaire. 150

Chapter 1

Introduction

> "It is the obvious which is so difficult to see most of the time. People say 'it's as plain as the nose on your face.' But how much of the nose on your face can you see, unless someone holds a mirror up to you?"
>
> Isaac Asimov, *I, Robot*

1.1 Motivation

The idea of artificial companions, either social robots or virtual agents, capable of engaging users for extended periods of time is receiving increasingly more attention in the recent years. Demographic trends, with a significant part of the population living alone, the decreasing costs in technology and the emergence of commercial robots for domestic settings are some of the reasons why research in this area is becoming progressively more relevant. The application domains where these companions can assist people are innumerous, from health-care assistants for the elderly to learning companions for children [24]. Wilks [154] envisions artificial companions as follows:

> "Intelligent 'cognitive' agents, implemented in software or a physical embodiment such as a robot. They can stay with their 'owner' for long periods of time, learning to 'know' their owner's preferences, habits and wishes. An

> artificial companion could enter a close relationship with its owner by chatting to, advising, informing, entertaining, comforting, assisting with tasks and otherwise supporting her or him."

Some aspects of this definition overlap with other similar concepts existing in the literature, such as *sociable robots* [23], *relational agents* [19] or *relational artifacts* [146]. Recently, in Europe, several projects have received funding to explore the potential of future companion technology in everyday life (for example, LIREC[1], COMPANIONS[2] and ALIZ-E[3]). The main motivation behind these projects is that current robots and virtual agents lack social capabilities that allows them to engage users in the long-term, especially when it comes to affective interactions [26]. In the field of social robotics, several studies show that the novelty effect quickly wears out and, after that, people lose interest and change their attitudes towards the robots [78, 51].

Many researchers agree that for robots to become part of our lives, they should be able to communicate with people in similar ways people interact with each other [24]. As social beings, relationships are very important to humans. It is often argued that empathy facilitates the creation and development of social relationships [3], as it increases similarity, fondness and affiliation [42]. Empathy can be broadly defined as "an affective response more appropriate to someone else's situation than to one's own" [70, p. 4]. Not only affective processes are influenced by empathy, but also cognitive and prosocial behaviours, such as actions taken to reduce the other's distress [116].

There are still many open challenges when developing robots and virtual agents for long-term social interaction. In particular, the role of affective interactions is still unclear. Previous work showed that empathic agents are, for example, perceived as more caring, likeable, and trustworthy than agents without empathic capabilities [22] and also that empathic agents can foster empathic feelings on users [112]. In the field of Human-Robot Interaction, there has also been a growing interest in studying different forms of empathy on user's attitudes towards robots [36, 123]. However, these findings

[1] http://lirec.eu/
[2] http://www.companions-project.org/
[3] http://www.aliz-e.org/

were obtained in studies where subjects interacted with these agents for a short period of time. Further research is needed to ensure that these results still apply in long-term interaction, as users' opinion is likely to change [82].

1.2 Research Goals

In this thesis, we propose to investigate the role of empathy in social robotic companions for long-term interaction. We argue that artificial companions capable of behaving in an empathic manner will be more successful at establishing and maintaining a positive relationship with users. To behave empathically, social robots need to understand some of the user's affective states and respond appropriately. Thus, our aim is to study the effects of a computational model of empathy in the long-term relationship established between the robot and the user. By doing so, we expect to find answers to the following research questions:

- Does the presence of empathic behaviour in a social robot affect the way people perceive the robot?
- Which perception and expression mechanisms are necessary for a social robot to behave in an empathic manner?
- How to select the appropriate empathic behaviour based on the affective state of the user?
- How to choose personalised empathic strategies for a particular user?
- What is the effect of the robot's empathic strategies in the long-term relationship established with the user?

To address these questions, we started by performing some preliminary user studies that motivated and supported the solution presented in this dissertation. We propose a general computational model of empathy for social robots that interact with users in real-world environments. This model includes the perception of naturalistic affective states from the user through several modalities, the implementation (and validation) of generic empathic behaviours in the robot, and the development of a decision-making mechanism that will allow the robot to select the most appropriate empathic behaviour considering the context of the interaction.

Most of the existing examples of social robots or virtual agents that recognise and respond to user's affect are based either on pre-scripted rules (inspired by cognitive or psychological theories) [45] or on machine learning techniques that determine the "optimal" intervention for each case [32]. This means that the agent displays the same behaviour whenever the user is experiencing a certain affective state, without knowing if such response is actually effective for that particular user, or if it is just making the user feel more frustrated. However, as suggested by Rich [122], "often individual users vary so much that a model of a canonical user is insufficient", especially if the user will spend a lot of time interacting with the system. As our goal is to improve the relationship between the robot and the user after several interactions, the robot should be capable of adapting its affective behaviour to the preferences of a particular user. The proposed model will allow us to study not only the long-term effects of empathy in social robots, but also to evaluate the influence of adapting the robot's empathic behaviour towards a particular user over time. The model was implemented and evaluated in a scenario where a social robotic companion with empathic capabilities plays chess with children.

1.3 Expected Contributions

By addressing the research questions presented above, in this thesis we expect to enrich the areas of social robotics and computational empathy by making the following contributions:

Contribution 1: Development of a generic computational model of empathy for social robots that interact with users in real-world environments. The existing computational models of empathy are more appropriate for agent-agent interactions or virtual environments where the user is represented as an avatar. They usually consider affective states that are impossible to recognise in real environments, considering the current technological limitations in perceptual systems. As our goal is to investigate empathy in a real-world scenario, we also need to investigate which perception and expression mechanisms are necessary for a social robot to express empathy towards the user, in order to build a realistic model that will allow social robots to respond appropriately to some of the user's affective states.

Contribution 2: Study the long-term effects of empathy in social robots. The studies conducted so far to investigate the effects of empathic social robots have only considered short-term interactions. Therefore, research is still needed to investigate whether the same results are also true when users interact with a social robot for extended periods of time. But how often should a robot interact with the same user so that the interaction can be considered as "long-term"? Some authors argue that two months is the answer [79, 141], while in Human-Computer Interaction there are reports of longitudinal studies lasting five weeks [82]. We believe that more than providing an exact number of weeks or months, it is more important to look at the actual number of interaction sessions with the robot and to the length of each session. Moreover, other factors should be considered, such as the number of users interacting with the robot at the same time and the repertoire of different behaviours of the robot. Taking this into account, we consider that long-term interaction happens when the user becomes familiarised with the robot to a point that her perception of such robot is not biased by the novelty effect anymore.

Contribution 3: Evaluate the influence of adapting the robot's empathic behaviour to the preferences of users. Research on interpersonal relationships states that similarity is, among other factors, a major determinant for the establishment of a social relationship [25, 18]. Since empathy can be seen as the capacity of experiencing similar affective states with another entity, we believe that by endowing social robots with the capacity of adapting their behaviour to the preferences of the user, similarity may increase. Another reason for evaluating the impact of adaptation is the fact that some affective responses, or actions resulting from the empathic processes, may have different impact in different types of users.

1.4 Outline

This thesis is organised as follows. In the next chapter we introduce some theoretical background on social relationships and empathy. Chapter 3 presents related work that motivated our research, more specifically in the fields of long-term interaction and empathy in virtual agents and social robots. Chapter 4 contains the foundation studies that contributed to the proposed empathic model. We start by presenting the scenario that will serve as the testbed for this work and, after that, the two preliminary studies that have been carried out. In chapter 5, we present our approach for modelling empathy in social robots, and provide some details on how the proposed model was implemented in our case study scenario. Chapters 6 and 7 contain, respectively, one short-term and one long-term user study performed to evaluate the proposed model. Finally, in chapter 6, we draw some final conclusions, design guidelines and future work directions.

Chapter 2

Theoretical background

This chapter presents a brief literature review on theories of human-human and human-pet relationships, and some work that explored the notion of human-machine relationships. Since empathy is an important social mechanism for maintaining social relationships and will be addressed in this thesis, we also introduce empathy and other related concepts.

2.1 Human-Human Relationships

Human relationships emerge from social interaction between two or more individuals. People establish many types of relationships with different degrees of closeness over their lifespan. Some examples of stereotypical interpersonal relationships include friendship, family relations (e.g., parent-child relationships) and intimate relationships such as marriage.

A relationship can be defined as "a persistent construct, incrementally built and maintained over a series of interactions that can potentially span a lifetime" [19]. Such interactions can be divided in two categories: generic patterns of interaction associated with stereotypical relationships in our society (e.g., friends are supposed to help each other) and specific interaction patterns for a particular dyad [13]. This latter category makes it difficult to define what a relationship is, when it starts or ends.

Research on relationship development states that there are many benefits of establishing relationships, as they play an important role on "human happiness, physical and mental health" [12]. Among other benefits, relationships can provide emotional support, appraisal support (e.g., advice, feedback and guidance) and other forms of nurture [13], and even instrumental support (e.g., material assistance or returning favours).

Relationships change and evolve over time, and therefore they go through different stages. Levinger [97] proposes a model of relationship development that consists of five different stages: (1) acquaintance, (2) buildup, (3) continuation, (4) deterioration and (5) termination. In the remaining of this section we will focus on the first three stages.

The process of getting acquainted with someone can be influenced by a variety of factors. For example, first impressions are extremely important for two persons to become attracted to each other. In this context, attraction is related not only to physical appearance [50], but also to similarity in terms of attitudes and personality. One of the most common references to understand our initial attraction to strangers is Byrne's attitude similarity theory [25], which conceptualises attraction as a function of reinforcement (i.e., something positive, attractive and desirable). Byrne argues that attitude similarity is reinforcing, and usually we find it attractive. Other predictors of attraction include individual characteristics (e.g., one's self-esteem) and previous relational history [110]. Although theories such as Byrne's attraction paradigm are reasonable ways of explaining how people start becoming acquainted with each other, there are other factors that shape our relationships, such as physical proximity or familiarity. Duck [46] argues that "most relationships are constructed by circumstances over which partners have little emotional control", as we are likely to know better people we meet more often. For example, consider the relationships we build in our neighbourhood, workplace and other environmental contexts.

Levinger's second and third stages (buildup and continuation) are related to relationship maintenance. To maintain a relationship, people perform a series of activities and behaviours in the attempt of keeping the relationship in a satisfactory condition. Several researchers divide these behaviours between routine and strategic behaviours.

While routine behaviours are those who "people engage in for other reasons which serve to maintain a relationship as a side effect" (such as performing daily tasks together) [19], strategic behaviours are "those which individuals enact with the conscious intent of preserving or improving the relationship" [137]. Examples of strategic behaviours include relational communication (e.g. social dialogue, talking about the relationship or recalling past activities together), performing planned activities together and self-disclosure (revealing personal information as a mechanism to give/receive advice, increase trust and intimacy) [46]. Based on the literature review of human-human relationship maintenance, Bickmore [18] identified several relationship maintenance strategies that could be employed by agents or robots to engage people in long-term interactions. These strategies include performing new activities together, meta-relational communication (i.e., talk about the relationship), empathy, reciprocal self-disclosure, the use of humour, talking about past and future events, continuity behaviours such as greetings, and emphasise commonalities and de-emphasise differences, to increase solidarity and rapport.

In essence, people communicate to establish and maintain relationships with each other, both verbally and non-verbally. Duck [46] states that "talking is fundamental to relationships", and Goldsmith and Baxter [59] found that everyday relating is dominated by six types of talk events: "gossip, making plans, joking around, catching up, small talk and recapping the day's events". But not only language is important, non-verbal communication also serves very important social functions in relationships, for example, to convey information about ourselves, to regulate the interaction (e.g. turn-taking and proximity) and express intimacy and emotional closeness [113]. Non-verbal behaviours are also used to express emotions. According to Bickmore [18], emotions and processes that help to manage emotions, such as empathy, play a major role in close relationships.

2.2 Human-Pet Relationships

Domestic animals came into people's lives many years ago for functional purposes (for example, cats were kept around to kill rats), but nowadays people have pets for other reasons. A study by Endenburg and colleagues [49] shows that people keep domestic animals mainly for social reasons. Within this category, companionship is in the first place, followed by child rearing considerations (e.g., children become more responsible if they have to take care of a living being), tactile interaction, attachment purposes and finally because they like to take care of the animal. As the authors of this study explain, "social interaction is one of the basic needs of human beings and even social interaction with animals can enlarge the well-being of humans". Moreover, some pet-owners establish such a strong bond with their pets that they see it almost like a human member of the family [4].

According to a recent survey from the American Pet Products Association, dogs are the most popular species to keep as pets, together with cats, fishes and birds[1]. Certain characteristics make dogs a very special type of animal companion for humans. They are loyal, devoted, affectionate (e.g., they exhibit "happy" behaviours when they see their owners), empathic, and can more easily adapt their behaviour to the owner's preferences than other pets [63]. Being able to play with it and touching it are other factors that seem to influence people's attachment to their dogs, and they can also work as social mediators, enabling social interactions between their owners and other people [49].

Another related area of research in human-pet relationships is animal therapy, defined by Levinson as "a process that introduces a companion animal into the life of a person to enhance his emotional well being" [98]. Usually, the ultimate aim of this type of therapy is the development of an attachment relationship between a patient and the animal. This relationship can evolve into a need for animal companionship, which may eventually induce the capacity for human companionship or other benefits in the patient. Most therapies are based on physical contact with the animal (e.g., touch), and they can range from scheduled activities regulated by a human therapist to long-term arrangements where patients adopt a pet. Several studies suggest beneficial effects of animal assisted therapy (for a comprehensive review see [83]). For example, Barker and Dawson [9] reported a significant reduction in anxiety levels on patients suffering from mood and

[1] http://www.americanpetproducts.org/press_industrytrends.asp

psychotic disorders after dog-therapy sessions.

2.3 Human-Machine Relationships

The question of whether humans can establish relationships with machines has received special interest since Reeves and Nass published their book "The Media Equation" [120], in which through a series of experiments they show that people treat computers and other media in a very similar way as they treat other humans, and they do so unconsciously. If such responses occur towards machines with little life-like characteristics employed in those studies (e.g., text-based interfaces), then what happens when people interact with artificial pets, social robots or embodied conversational agents?

In a study that investigated people's social responses to an AIBO robotic pet through the analysis of posts in online AIBO discussion forums, Friedman and colleagues [55] found that even though AIBO owners frequently refer to the robot as an inanimate artifact (e.g., mentioning its technical capabilities), very often they also attribute life-like and social capabilities, and even mental states to AIBO. Nevertheless, only a few forum posters attributed moral standing to the robot (e.g., consider that it deserves respect or had rights similar to the ones real pets do). Turkle et al. [147] reported several case studies in which AIBOs, My Real Baby robotic dolls and Paro robots were introduced in schools and elder care facilities on a weekly basis for approximately two years. In the presence of such robots, children expressed a desire to nurture and be nurtured by them, and the same happened with the elderly group. According to the authors, "we care about what we nurture", and if people are willing to care about machines, they may also become attached to them. In another study, Chesney and Lawson [31] investigated the interactions of players with Nintendogs[2], a very popular Nintendo DS game where players need to take care of a screen based virtual pet. The results suggest that Nintendogs pets provided companionship to players, even though the levels of companionship were significantly lower that the ones provided by real pets.

[2]http:www.nintendogs.com

2.4 Empathy and Social Support

According to Hoffman, empathy is "an affective response more appropriate to someone else's situation than to one's own" [70, p. 4] that results from one (or more) of the following empathic arousal mechanisms: *motor mimicry*, *classical conditioning*, *direct association* of cues from the person we are empathizing with, *mediated association* (association of the victim's cues through semantic processing) and *perspective-taking*. The first three are considered preverbal, automatic and involuntary, and the last two are higher-order cognitive modes. From these five empathic arousal mechanisms, there is empirical evidence suggesting that perspective-taking, defined as "putting oneself in the other's place and imagining how he or she feels" [70, p. 52], is more "empathy-arousing" than focusing directly on the victims verbal and non-verbal cues [139]. Thus, perspective-taking has been considered a very important component of empathy [52]. In our work, we follow Hoffman's definition of empathy, for being more focused on the response rather than in the process that generates empathy.

Empathy has been considered one of the major determinants of prosocial behaviour [52, 47, 41], especially when observing someone in pain, danger or other types of distress [70]. This happens not only because usually empathy causes similar feelings in both the observer and the victim, but also because observers feel better after helping. Prosocial behaviour can be defined as "voluntary behaviour intended to benefit another, such as helping, donating, sharing and comforting" [48]. Prosocial behaviours can also be designated as socially supportive behaviours, and they can take several forms. For example, as suggested by Cutrona and colleagues [38], they can be separated in the following categories: *information support* (advice or guidance), *tangible assistance* (concrete assistance, for example by providing goods or services), *esteem support* (reinforcing the other's sense of competence), *emotional support* (expressions of caring or attachment) and *network support* (social integration). The perception of social support is considered to have positive effects on social relationships [155]. Moreover, it has been linked to positive outcomes in children's mental health and coping with traumatic events [149].

However, there are individual differences both in the perception and also on the employed types of support depending, for example, on people's gender and personality [61, 121].

Although empathy can be seen as an intrinsic characteristic of human behaviour, the prosocial behaviours triggered by empathic feelings can sometimes be influenced by external factors. Hoffman [70] argues that such limitations, also known as modulators, are due to two main reasons: *empathic over-arousal* and *relationship between the observer and the victim*. Empathic over-arousal occurs when the signs of distress are so intense that the observer's empathic distress is transformed into a feeling of personal distress. The relationship between the observer and the victim also strongly affects the type of prosocial action taken by the observer. For example, people tend to be more empathic towards friends and family members than towards strangers [92]. Other researchers suggest that empathic responses can be modulated by other factors, such as the observer's *personal characteristics* (past experience, mood, personality, gender, age...) or the *context of the situation* [42]. Likewise, in the OCC model [111], the intensity of the empathetic emotions "happy-for" and "sorry-for" are affected by four main variables, namely *desirability-for-self*, *desirability-for-other*, *deservingness* and *liking*.

In the literature, empathy has often been mixed with other similar concepts such as emotional contagion and sympathy. Preston and de Waal [116] distinguish these concepts by enumerating the main differences between them and the notion of empathy. For example, emotional contagion differs from empathy because there is no distinction between the self (observer) and the target of distress, and usually this phenomenon does not lead to prosocial actions and helping behaviour, but rather to personal distress in the observer. Regarding sympathy, the main difference is that sympathy typically focuses on feelings of "sorrow" for the victim, and such feelings are more focused in the situation itself and not properly in the victim. For this reason, the observer does not necessarily matches the other's emotional state. In other words, sympathy is a more cognitive process by which the observer does not really need to feel the emotion experienced by the target.

2.5 Concluding Remarks

Taking as an analogy Levinger's relationship stages [97] presented earlier in this chapter, our aim is to have the robot interacting with the user until the third stage of the relationship model (continuation). Even though dialogue is considered a very important factor for maintaining human relationships, non-verbal behaviours also play a major role. In our work, the main focus is on non-verbal behaviour, more specifically the role that empathy plays in the possible long-term relationship established between users and a social robot.

The studies on human-robot relationships illustrate how simple robotic or virtual agents, especially if they have human or animal-like characteristics either in terms of appearance or behaviour, can elicit feelings and social responses that people use when relating to other living beings. As Bickmore suggests, nothing seems to prevent computers from "eventually fulfilling the role of a relational partner" [19]. This poses several ethical questions regarding the authenticity and validity of such relationships, raising questions such as "what is the ethical status of building a device whose behaviour signals emotions that it does not actually feel?" [35] or "what kinds of relationships are appropriate to have with machines?" [146]. However, it does not necessarily mean that such relationships will be similar in every way to the ones humans establish with each other. Addressing the field of robotic pets, Friedman et al. [55] state that they challenge traditional boundaries between "who or what can possess feelings, establish an emotional connection or engage in companionship".

As one can see from the previous sections, every relationship is different and relies mainly on the type interactions of its constituents, so interactions with artificial companions may just increase the scope of the word *relationship*.

Chapter 3

Related Work

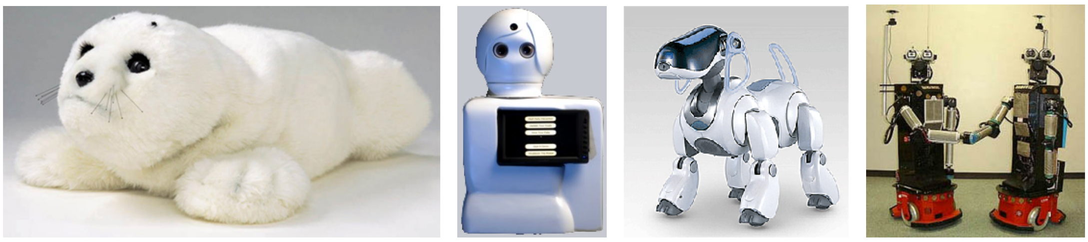
This chapter describes existing work relevant for this thesis, and is divided in two main sections. In the first section, we review existing long-term interaction studies, with special emphasis on studies using social robots. The second section is dedicated to empathy in virtual agents and social robots.

3.1 Long-term Interaction Studies

In this section, we present the state of the art on long-term interaction studies, organised by application domains. Although our focus is on long-term studies with social robots, the first work that will be presented includes a virtual agent, for being one of the first and most significant long-term studies carried out so far with an embodied character. We decided to organise the selected works by their application domains (health care and therapy, education, work, and home environments) because the robot's features and the study design (e.g., data collection methods) are more likely to be similar in the same domain, and therefore easier to compare.

3.1.1 Health Care and Therapy

One of the most promising application domains for long-term interaction between humans and embodied agents is health-care and therapy. As part of a system that tracks



(a) Paro (AIST, Japan)[150, 151, 152] (b) Autom[84] (c) AIBO[54] (d) Robovie [128, 79]

Figure 3.1: Robots used in the Health Care and Therapy related long-term studies (images used with permission of the authors).

users' healthy habits, Bickmore and Picard [19] developed Laura, a virtual conversational agent that employs many of the relationship maintenance strategies enumerated in section 2.1. Through more than 30 scripted dialogue lines, Laura uses relational behaviours such as small talk, empathy dialogues, meta-relational communication, humour and continuity behaviours (e.g. proper greetings and farewells), while politely asking users about their physical activities during the day. Several non-verbal behaviours were also included in the agent, for example hand gestures, head-nods and eyebrow raises. Laura was evaluated in a controlled experiment with approximately 100 users who were asked to interact daily with the exercise adoption system. After four weeks of interaction, the agent's relational behaviours increased the participant's perceptions of the quality of the working alliance (on measures such as liking, trust and respect), when comparing the results with an agent without relational behaviours. Besides, participants interacting with the relational agent expressed significantly higher desire to continue interacting with the system.

In this domain, there is also a great potential for social robots that interact with users for extended periods of time. *Socially assistive robotics*, as defined by Matarić and colleagues [102], are expected to "augment human care and existing robot-assisted hands-on therapy toward both improving recovery and health outcomes and making the therapeutic process more enjoyable". They argue that the physical embodiment of the robot, its personality and the ability to model some of the patient's motivational states

Table 3.1: Summary of the long-term studies in the Health Care and Therapy domains.

References	**Agent/ Robot**	**Capabilities**	**Exp. Design**	**Nr. Sessions**	**Main Results**
Bickmore & Picard (2005)	Laura	Small talk, empathy dialogues, meta-relational communication, humour and continuity behaviours	Subjects: 100 Measures: working alliance Methods: questionnaires, interviews	30 days	Quality of the working alliance increased; participants interacting with the relational agent expressed higher desire to continue interacting with the agent.
Wada & Shibata (2006, 2007)	Paro	Animal-like behaviour; responds to touch, sound and lights; limited-keyword recognition	Subjects: 12 Measures: degree of social interaction, stress levels Methods: video, interviews, urinary tests	30 (9 hours a day)	Increased social interaction between participants, stress levels reduced.
Heerink et al. (2008)	iCat	Providing information (e.g., weather forecast, TV programs) and telling jokes	Subjects: 30 Measures: enjoyment and intention of use Methods: questionnaires and interaction logs	1 to 9 (5 min. average)	Enjoyment and int. of use significantly affected by interaction times and length.
Kidd & Breazeal (2008)	Autom	Eye contact and small talk depending on time of day, state of the relationship with the user	Subjects: 45 (3 conditions; 17-72) Measures: weight loss, WAI, usage of the system Methods: questionnaire	50 (average)	Participants interacting with the robot reported their weight for more days and expressed more willing to continue interacting with the system.
Francois et al. (2009)	AIBO	Animal-like behaviour	Subjects: 6 (autistic children) Measures: children's progress during interaction Methods: video observation	10 (40 minutes each)	Children tended to express more interest towards the robot over time, with occasional displays of affect.
Sabelli et al. (2011)	Robovie	Remotely operated dialogues and child-like behaviours (e.g. "what is this?")	Subjects: 55 Measures: interaction patterns during interaction Methods: interviews, direct observations	15 to 35 (10 to 20 minutes each)	Robot was well accepted due to role as "child" and behaviours such as greetings and calling users by their names.

can have a positive impact in robots applied in this context. Additionally, the role of the robot (e.g., physical therapist or nurse's assistant) and the task that the robot is meant to achieve must be clear to the user.

Of particular importance is the work by Wada and Shibata [150, 151, 152] with the robot Paro (Fig. 3.1a), which can be considered one of the landmarks in the field of long-term human-robot interaction. Paro is a seal shaped robot specifically designed for therapeutic purposes. The robot's behaviour contains a reactive layer for responding to certain stimuli (e.g., touch, sounds and light) and a proactive layer triggered by the robot's internal needs. Paro is also able to recognise certain keywords that users may use more frequently around it (for example, when they give the robot a new name), and gradually adapt its behaviour according to the stimuli of the user. Two Paro robots were placed in common living rooms of a care house where elderly residents could interact with the robot over 9 hours a day. The interactions of the residents with PARO were video-recorded during this period. After one month, the results of 12 subjects indicated that PARO strengthened the social ties among the residents of the care house and that most residents established moderate or strong ties with the robot (e.g., greeting Paro when they pass by). Also, urine tests showed that after introducing PARO, the stress levels of the residents decreased. The long-term effects of Paro in nursing home residents were also investigated by other researchers with similar results (see, for example, the studies by Turkle et al. [147] and Giusti and Marti [57]).

Another example is the study by Heerink et al. [65], who investigated whether the perception of enjoyment and intention of use influenced the acceptance of a social robot by elderly users. They conducted a study in a nursing home, in which 30 participants interacted with an iCat robot and, after the first session, answered a questionnaire measuring enjoyment and intention of use. After this initial session, the robot was left in the nursing home for another 5 days and participants could freely interact with it at any time. By analysing the log files of the system, the authors found that 23 out of the 30 users interacted again with the robot at least once, up to a maximum of nine interactions (the average of subsequent interactions was 1,9 times of about 5 minutes each). Additionally, significant correlations were found between the enjoyment and intention of use scores reported by users in the first session and their actual usage in minutes and times of use at the end of the study.

More recently, Sabelli et al. [128] reported an ethnographic study in which a humanoid social robot (Robovie, displayed in Fig. 3.1d) interacted with 55 residents of an elderly care centre for three and a half months. The robot's behaviour was remotely operated to act as a conversational partner through basic dialogues that included greetings, questions about hobbies, travel experiences and other child-like questions such as "what is this?". The analysis to the interviews and direct observations of the interactions suggested that the robot was well accepted in the community. Behaviours such as greetings, calling participants by their names and the robot's role as a "child" were relevant for this result.

Social robots have also been employed successfully in autism related therapy [40, 106]. To investigate the impact of robotic companions in autistic children, François et al. [54] conducted a study in which six children played with an AIBO robot (Fig. 3.1c) for approximately 40 minutes once a week until a maximum of 10 sessions. AIBO was programmed to display several dog-like behaviours (e.g., wagging the tail, emitting "bark" sound or opening and closing the mouth) when users touched the robot's head, chin and back sensors. The experiment took place in a school setting and was inspired by non-directive play therapy, where the experimenter can participate in the trial but the child is the main leader in an unconstrained environment. The sessions were video-recorded and children's behaviour was later analysed according to three dimensions: play, reasoning and affect. Each child made progress in at least one of these three dimensions over the sessions. Children experienced progressively higher levels of play and developed more reasoning related to the robot (for example, by comparing AIBO to a real dog). Besides, they tended to express more interest towards the robot over the sessions, with occasional displays of affect.

In a different area, Kidd and Breazeal [84] studied the impact of social robots in terms of behaviour change while dieting. They developed a social robot, Autom (Fig.

3.1b), which is capable of establishing eye contact with the user and making small talk while helping individuals to keep track of their weight loss. The dialogue lines vary depending on the time of day, estimated state of the relationship with the user (initial, normal or repair), the time since the last interaction and the inputs of the user such as the number of calories. The study included 45 participants with ages between 17 and 72 years distributed through three different conditions: some participants interacted with the robot, others reported their weight loss in a computer and the third group used a traditional paper log. The main dependent variables were weight loss, usage of the system and the Working Alliance Inventory [72]. Although the weight loss results were not significantly different among the three groups, participants with the robot interacted significantly more days with the system (on average, 50.6 days against 36.2 and 26.6, respectively for participants who used the computer and participants using the log paper system) and expressed more willingness to maintain the interaction than participants in the two control conditions.

Discussion

All the presented studies (see Table 3.1 for a summary) found positive results regarding the long-term effects of social robots and virtual agents in therapeutic or health-related scenarios. However, the users who took part in these studies were very different (elderly, adults and autistic children), and so further research is needed to consolidate these results. Even when the target user group was the same, for instance in the studies with elderly users, the evaluation metrics were very different, from physiological measures to ethnographic observations. One key issue in the HRI community, which is the lack of common methodologies and metrics that allow the comparison of different robots and applications [138, 14], is reflected here. Despite these limitations, this area seems to benefit from the introduction of agents that can complement human activity and help users to achieve their goals, while receiving additional comfort or attention.

An important aspect in the studies presented here is that animal-inspired robots were used in two of them, namely in the work of Wada and Shibata [152] and François et al.

[54]. In fact, recently some researchers have been trying to transfer the positive effects of Animal Assisted Therapy (e.g., reduced loneliness and the development of attachment bonds) into HRI, by comparing the effects of social robots with those of real animals [8]. If such results are verified, robots can replace animals for example, in hospitals where for hygienic reasons animals are usually not allowed.

Table 3.2: Summary of the long-term studies in the domain of Education.

References	**Robot**	**Capabilities**	**Exp. Design**	**Nr. Sessions**	**Main Results**
Kanda et al. (2004)	Robovie	Identify users, recognising and speaking English	Subjects:228 Measures: length of interaction, English skills Methods: video observation, English tests	9 school days	Interaction after 1^{th} week declined; improvement of English skills in children who kept interacting with the robot.
Kanda et al. (2007)	Robovie	Identify users, pseudo-development mechanism, confiding personal information	Subjects: 37 (10-11 years) Measures: length of interaction Methods: questionnaire, video observation	32 school days	Children kept interacting with the robot after the 2^{nd} week.
Salter et al. (2004)	Wany	Obstacle avoidance, move in the environment	Subjects: 8 (5-8 years, male) autistic children Measures: activity around the robot Methods: video observation, analysis of interaction data	5	Children lost interest in the interaction from the third session.
Tanaka et al. (2007)	QRIO	Choreographed dance sequences and mimicking children's movements	Subjects: 11 (10-24 months) Measures: quality of interaction, haptic behaviour towards the robot Methods: video observation	15 (45-50 min. each)	Toddlers progressively started treating QRIO as a peer and exhibited several care-taking behaviours towards the robot.
Kozima et al. (2009)	Keepon	Display non-verbal behaviours (gaze, emotions, ...)	Subjects: 27 (3-4 years) Measures: children's responses Methods: video observation	20 (90 minutes each)	Robot played the role of social mediator; children maintained interest over the sessions.
Kozima et al. (2009)	Keepon	Display non-verbal behaviours (gaze, emotions, ...)	Subjects: 30 (2-4 years, autistic) Measures: children's responses towards the robot Methods: video observation	15	Although eye contact decreased, children gradually approached the robot more and established physical contact.
Hyun et al. (2010)	iRobiQ	move head and arms, navigate in the environment, express emotions	Subjects: 111 (5 years) Measures: children's perception of the robot Methods: interviews	10 (approx. 1 hour)	Robots are well accepted by children in educational settings.

3.1.2 Educational Settings

Another popular area already with a significant amount of long-term studies using social robots is education. Virtual pedagogical agents have been used for many years in this context (for a comprehensive survey, please consult [91]), and it is expected that social robots might have the same beneficial effects on students, especially due to their physical presence.

In 2004, Kanda et al. [78] performed a field trial evaluation for two weeks (9 school days) with elementary school Japanese students and two English-speaking interactive humanoid robots behaving as peer English tutors. The robots ("Robovie 1" and "Robovie 2") were capable of recognising children and calling them by their names using RFID tags, displaying several interactive behaviours such as greeting or hugging, as well as recognising 50 different English words and displaying some English utterances. The study revealed that the robots failed to keep most of the children's interest after the first week, mainly because the first impact created unreasonably high expectations in the children. However, children who kept interacting with the robots after the first week improved their English Skills. They also found that, very often, children interacted with the robots together with their group of friends. These results motivated a subsequent study where Robovie's capabilities were extended to better support long-term interaction with children [79]. The new capabilities include a pseudo-development mechanism (the more a child interacts with the robot, the more different behaviours the robot displays to that child), and self-disclosure behaviours (e.g., the robot may reveal its favourite baseball player). This newer version of Robovie interacted with children in Japanese in their classroom for 2 months (32 actual experimental days). In contrast to the results obtained in the previous experiment, Robovie was capable of engaging children after the second week (although with a slight decay), which the authors attribute to the new capabilities implemented in the robot. The children's motivations for interacting with the robot were also studied. Most children answered that their main motivation was to become friends with the robot. Regarding friendship estimation, the robot was also capable of estimating some of the friendship relations of children, by analysing the groups of children who interacted together with the robot.

Salter and colleagues [130] carried out a study to extract patterns of interaction between children and a commercial small robot equipped with infra-red sensors. Eight children with ages between 5 and 8 years old interacted with the robot five times each in individual sessions of approximately 5 minutes. In the first three sessions, children could freely play with the robot while it was performing obstacle avoidance, but they soon get bored with the interaction. Therefore, in the fourth session, researchers added some plastic stickers simulating eyes to the front of the robot with the purpose of regaining children's interest. However, this novelty factor quickly vanished. In the last session, the speed of the robot was increased, but again, this did not seem to increase children's engagement. The technical limitations of the robot also affected negatively the interaction, for example the robot getting "stuck" in the environment coincided with less activity of children.

With the goal of studying the dynamics of the interaction developed between children and social robots, Tanaka et al. [144] reported a longitudinal study where a QRIO robot (Fig. 3.2a) interacted with toddlers in a day care centre for 45 sessions of 45 to 60 minutes each. QRIO can display several behaviours including choreographed dance sequences and mimicking some of the children's movements. The study was divided in three different phases, in which the robot's behaviour varied slightly. Five independent coders annotated the video recordings of 15 sessions in terms of quality of the interaction. The quality of interaction increased during phase I, then decreased sharply on phase II and on phase III returned to the levels observed in phase I. The same pattern was verified for the amount of times children touched the robot. Moreover, when introducing two inanimate toys in the environment (a teddy bear and a toy very similar to QRIO), QRIO was still the most hugged by the children followed by the toy that looked like the robot. The results of this study suggest that toddlers progressively started treating the robot as a peer rather than as a toy, as they exhibited an extensive number of care-taking behaviours towards the robot.

Kozima et al. undertook a similar study to investigate the interaction between toddlers and Keepon (Fig. 3.2b), a small robot designed to interact with children through non-verbal behaviours such as eye contact, joint attention and emotions [90]. A group of 27 children interacted with Keepon in their class during 90 minutes for 20 sessions. As it was observed in Kanda et al.'s studies, Keepon often played the role of social mediator between children, who exhibited a wide range of spontaneous actions towards the robot. Children maintained their interest over the sessions, which was not observable in previous studies using the same robot with older subjects. The authors report that children's understanding of the robot changed over time, from a mere "moving thing" to a "social agent". Keepon was also introduced in the play room of a day-care centre for children with developmental disorders for more than 3 years. The results of the first 15 sessions (spanning through five months, approximately one hour per session) indicate that, even though eye contact between children and Keepon gradually decreased, children were able to spontaneously approach the robot and establish physical and social contact. As they gradually learn the meaning of the robot's actions and responses, dyadic, triadic and empathetic interactions start emerging.

Another research topic that has been receiving increasingly more attention is children's preconceptions and judgements of robots [15, 11]. With this goal, Hyun et al. [74] interviewed 111 children who had interacted with an iRobiQ robot (Fig. 3.2d) for around one hour over a two-week period. The interviews were around three main themes: robot's comparison to other media, a perception survey and appearance of robots. The results indicate that robots are well accepted in educational settings, and that mechanisms that promote social and emotional interactions between robots and children contribute to great extent to this acceptance.

Discussion

Table 3.2 summarises the studies within this domain. These studies show conflicting results that can be explained by the differences in the age groups of the subjects and by the variety in terms of complexity of behaviour of the robots employed in the experiments. Regarding age, younger children are more likely to engage with robots, possibly

(a) QRIO[144] ©Sony

(b) Keepon [90]

(c) iCat[95] ©Philips Research

(d) iRobiQ[74] ©iRobiQ Co.,Ltd.

Figure 3.2: Some of the robots used in the long-term studies in the Education domain (images used with permission of the authors).

because they do not have any preconceptions of what a robot is supposed to do, and they may see it simply as a very special/advanced kind of toy, as shown by the results of Tanaka et al. [144] and Kozima et al. [89]. In this last work, the authors even state that another experiment with the same robot but with older subjects was not so successful.

Another aspect that plays an important role for sustaining long-term engagement is the complexity of the robot's social behaviour, and the amount of diverse behaviours that the robot can display over the interaction. Salter et al. [130] found that a robot without any social capabilities was not capable of engaging autistic children after the third interaction. As studied by Kanda et al. [81], the incremental implementation of novel behaviours (e.g., self-disclosure) did play an important role for maintaining users engaged with Robovie. This can be explained by the habituation effect [73], a phenomenon characterised by children's tendency to, after a certain period of exposure, shift their preferences to novel (rather than familiar) stimulus.

In sum, even though social robots are well accepted by children in educational environments, they should be able to simulate complex and diverse social behaviours in order to engage children in the long-term. Considering the findings presented above, it seems to be the case that the older the children, the more complex, diverse and dynamic should the robot's behaviour be.

Table 3.3: Summary of the long-term studies in Work Environments and Public Spaces.

References	Robot	Capabilities	Exp. Design	Nr. Sessions	Main Results
Severinson-Eklundh et al. (2003)	Cero	Fetch-and-carry objects such as books or coffee cups	Subjects: 1 target user in a work group of 30 Measures: long-term effects of a service robot Methods: video and direct observation, system logs, pos-trial interviews	66	Social robots in public spaces should be able to interact with everyone, not just the main users.
Stubbs et al. (2004)	PER	Simulated scientific testing	Subjects: 11 Measures: people's cognitive model of the robot Methods: interviews	3 months	Regular interactions influence people's cognitive model of the robot.
Gockley et al. (2005)	Valerie	Reveal backstory, recognise people around the booth, limited natural language user interaction through text input	Subjects: 233 Measures: length of interactions Methods: analysis of interaction data	180	Many users kept interacting daily with the robot, but after a certain period only a few interacted for more than 30 seconds.
Kirby et al. (2007)	Valerie	Additional mood displays while telling stories	Subjects: 62 Measures: length of interactions Methods: analysis of interaction data, questionnaire	45 (8 hours a day)	Interaction patterns change according to the robot's mood and level of familiarity with the robot.
Kanda et al. (2010)	Robovie	Guiding, rapport building, identify repeated users, advertisement	Subjects: 162 Measures: intention of use, interest, perceived familiarity, intelligence and adequacy of route guidance Methods: questionnaire	2.1(average); from 2 to 18 sessions	Perception of the robot was positive; shopping suggestions of the robot were accepted by visitors.

3.1.3 Work Environments and Public Spaces

There are already some successful examples of commercial robots deployed in work environments and public spaces. Examples include Robotdalen's RobCab[1], a transportation

[1] http://www.robotdalen.se/en/Projects/RobCab---transportation-robot-for-hospitals/

(a) Cero [134] (b) Valerie [58, 86] (c) PER [140]

Figure 3.3: Robots used in the long-term studies in Work Environments and Public Spaces (images used with permission of the authors).

robot for hospitals, the Siga Robots[2] developed by YDreams to guide and interact with guests visiting the headquarters of Santander bank, and the robotic characters developed by Walt Disney Imagineering for the Disney parks[3]. As we will see from the long-term studies presented in this section, social robots deployed in these environments should be able to adapt to different and unexpected situations, as well as to different types of users.

Severinson-Eklundh et al. [134] reported probably one of the first long-term studies in a real-world setting involving a social robot. The goal of the study was to investigate social aspects of the interaction with a fetch-and-carry robot for motion impaired users in an office environment. The developed robot, Cero (Fig. 3.3a), was evaluated during 3 months in the workplace of a target user, a female academic with a walking disability. From the analysis of the videos recorded during the trial, the internal logs of the system and a post interview with the target user, the authors extracted patterns of how people interact and relate with robots in work environments. One interesting finding was that very often other people (e.g., office workers, cleaning staff, ...) wanted or needed to interact with the robot, but did not know how to do it. Thus, it is important that social robots immersed in public spaces can provide clear instructions on how to be operated

[2] `http://www.wired.co.uk/magazine/archive/2011/08/start/friendly-bank-bots`
[3] `http://disneyparks.disney.go.com/blog/tag/autonomatronics/`

and be “easy to use” by people who are unfamiliar with the robot. They also raised some issues for future research in long-term interaction such as the personality of the robot, the dialogue between users and the robot, and the relevance of group collaborations.

In a different context, Stubbs et al. [140] examined how people’s cognitive model of a robot changes over time. The target robot was PER (Fig. 3.3c), a robot designed to educate people about NASA’s Mars exploration robot, and the selected subjects were museum employees who interacted with PER on a daily basis. The study consisted on interviewing 11 museum employees at different stages of their relationship with the robot: the first interview was conducted before the PER exhibition was installed, followed by three other interviews, more precisely two weeks, one month and three and a half months after the installation. The results of the open-ended interviews indicate that regular interactions influence people’s cognitive model of the robot. Over time, references to anthropomorphisation increased significantly, together with discussions about the robot’s intelligence. Conversely, themes related to the technical capabilities of the robot became less frequent.

Gockley et al. [58] developed Valerie (Fig. 3.3b), a “roboceptionist” installed at the reception of one of the buildings at the CMU campus. The robot has a personality and a background story that is gradually disclosed to people through monologues. Students and university visitors interacted with the robot over a nine month period. The results indicated that while many users kept interacting daily with the robot, after a certain period only few of them interacted for more than 30 seconds. From the analysis of the interactions, the authors proposed some design recommendations so that Valerie (and possibly other robots) can be more engaging in the long term. Such recommendations include proper greeting and farewell behaviours, more interactive dialogue (rather than monologues, which did not attract visitors the way authors were expecting), a robust way of identifying repeated visitors and the ability to display emotions. This last design recommendation inspired another long-term study with the same robot [86] where Valerie also exhibited different moods (positive, negative or neutral) matching its life stories. The length and number of interactions were measured for a total of nine weeks, during which the robot operated eight hours per day, five days per week. Some users filled in a questionnaire that measured their subjective experience towards the robot. The analysis of the results was separated in two different groups, one for frequent users of the building (more familiarity with the robot) and another group for visitors (less familiarity). It was concluded that the robot's moods were easily recognised by all visitors, and that interactions were different depending on the level of familiarity: frequent users interacted more times when the robot was in a positive mood, but the amount of time they dedicated to the robot was higher when it was in the negative mood. The authors justify these results with the common ground theory (e.g., a smile can be understood as a positive signal that carries a certain amount of conversational content) and with the questionnaire answers, where participants found the positive mood robot less enigmatic. On the other hand, visitors spent less time interacting with the neutral mood robot, which may indicate that any form of affect display can be enough to sustain interactions. In short, the interaction patterns change according to the mood displayed by the robot, and how such patterns change depends on the person's level of familiarity with the robot. The authors use this argument to reinforce the idea that social robots need to properly identify users (to change their behaviour whether users are newcomers or repeated visitors), and that "a rich model of affect is necessary for forming long-term human-robot relationships".

More recently, Kanda and colleagues [80] also evaluated Robovie in a shopping mall. In this study, the robot was programmed with a different set of behaviours particularly relevant to a shopping mall environment: apart from building rapport with users by identifying them using RFID tags, employing self-disclosure mechanisms and adjusting the dialogues based on the previous dialogue history with each user, Robovie was also capable of offering directions and advertising specific shops and services of the mall. Although the long-term study considered 162 participants, 72 of them interacted with the robot no more than 2 times and only 23 participants interacted with the robot more than 5 times. The authors explain in the paper that this effect might have been caused by the continuous presence of many people (visitors of the shopping mall but not official participants of the study) around the robot. Due to the large cues, participants hesitated before deciding to interact with the robot. Questionnaires mailed to the study participants (even the ones who only interacted with the robot once) suggested that their perception of the interaction was positive, not only in terms of perceived familiarity, intelligence and interest towards the robot, but also regarding intention of use and adequacy of the route guidance behaviours. Moreover, repeated visitors provided significantly higher rankings in the questionnaire. In addition to these results, the study also concluded that people's shopping behaviour was influenced by the robot's suggestions.

Discussion

It is difficult to generalise and draw conclusions from the studies in this section because they were performed with very different robots, both in terms of embodiment and functionality (see Table 3.3 for an overview) and, apart from the studies with the Roboceptionist and Robovie, the sample size was small. Yet, these studies show that deploying social robots in public environments, where they can interact with almost every type of person, requires additional efforts in terms of usability and adaptation, so that they can better deal with the uncertainties of the environment. As the study by Severinson-Eklundh et al. [134] shows, people should easily learn how to interact with the robot and have access to its internal state, for example, to understand where the robot is going and eventually be able to help. To address this issue, Rosenthal and colleagues [127] are developing a robot that can ask a human for help to overcome some of its limitations (e.g., when it needs to pass through a door that is closed). On the other hand, robots should be able to adapt their behaviour to different types of users (for example, distinguishing between new and repeated visitors, as suggested in the "roboceptionist" study) so that interactions become more natural and intuitive.

Overall, the robots in these studies lack perceptual capabilities that would enable richer social interactions with users. While in the first two studies this was clear to users due to the robot's embodiment being extremely functional (and as such did not seem to affect the interaction), in the study by Gockley et al. [58] users were more disappointed by Valerie's monologues, as they expected more from its human-like embodiment. These results highlight the relevance of interactive experiences, where the user plays an important role influencing the robot's behaviour rather than being a mere spectator. Note that there are examples of rich social interactions with limited communication modalities (as seen, for instance, in interactions with Paro robot). However, if the robot's appearance suggests the existence of more human-like social communication capabilities (such as the embodiment of the Roboceptionist), then it should be able to interact with users using those modalities.

Nevertheless, robots in public spaces such as offices or shopping malls appear to be well accepted by users, and thus further long-term studies should be conducted within these application domains. These results are in line with a study performed by Takayama et al. [143] about occupations for which people consider that robots are qualified and desired. The results indicate that people envision robots performing jobs that require memorization, perceptual skills and service-orientation, in contrast with the notion that robots should only do dangerous, dirty or dull jobs.

3.1.4 At Home

Domestic environments are receiving increasingly more attention as an application for social robotics research. Even though robots at home were envisioned many years ago in science fiction, only recently technology has been robust enough to allow the execution of long-term evaluations in these settings. Long-term studies in this domain are even more challenging due to the privacy issues and consequent lack of control of users' activities.

In the work of Koay et al. [89], the habituation effects between users and a social robot were investigated during eight interaction sessions over a five week period. More precisely, their goal was to isolate certain aspects of participant's preferences that may be influenced by the habituation effect. The experiment was conducted at the Robot House, a naturalistic environment especially designed to study human-robot interactions outside the laboratory conditions. Participants expressed their preferences when a PeopleBot humanoid robot (Fig. 3.4a) approached them in several forms. The findings suggest that even though preferences did not change in the first two sessions, in the last sessions participants allowed the robot to come closer than in the previous interactions.

Most of the longitudinal studies in domestic settings employ commercial robots. For example, Sung et al. [141] empirically evaluated how people used and accepted Roomba vacuum cleaner robots (Fig. 3.4b) in their homes. Roombas were distributed through 30 households and one experimenter visited each household five times during a six-month period. The first visit happened a week before Roomba was introduced, the second visit when families unpacked the robot and used it for the first time, and the other three visits took place respectively, two weeks, two months and six months after Roomba was introduced. During the visits, several methods besides traditional interviews were used to better capture people's routines and acceptance of the robot, such as drawings, probing techniques and checklists of the activities they did with Roomba. Participants were also encouraged to report their experiences with the robot via e-mail. The authors argue that two months are enough for observing stable interactions between robots and households in a domestic environment. They also found that the combination of several data collection methods is extremely useful for capturing people's routines and interaction with the robot, especially in a domestic environment. Based on this study, the first steps for establishing a long-term framework were taken [142]. The framework includes four different temporal steps that contain key interaction patterns experienced while households were accepting the robot: *pre-adoption*, *adoption*, *adaptation* and *use and retention*.

Fernaeus et al. [51] reported a similar study with Pleo (Fig. 3.4c), a robotic toy dinosaur. Six families took a Pleo robot home for 2 to 10 months (each family decided for how long they wanted to keep the robot). They were also given a video camera to record moments of their interaction with the robot. Most families were interviewed at least two times after having Pleo in their homes. The study is focused on the discrepancy between previous expectations that participants had about Pleo, and how the robot met (or failed to meet) such expectations. Participants' initial expectations were really high due to the price, sophistication and the advertisements about Pleo. However, the robot's

(a) PeopleBot [89] (b) Roomba[141, 142] ©iRobot (c) Pleo [51] (d) Nabaztag [87]

Figure 3.4: Robots used in the long-term studies in home environments (images used with permission of the authors).

behaviour was not attractive enough to keep these expectations so high until the end of the study. After the initial novelty effect, participants did not interact with Pleo in a regular manner. After a while, Pleo was only switched on in special occasions, for example when friends were visiting the home. Even though at first Pleo was treated in a similar way to a real animal, with activities such as petting and choosing a name for the robot, over time, it failed to encourage regular interactions and started being treated as a regular toy. This study reinforces the findings of Jacobsson [75] which reveal that the majority of the blog and forum posts about Pleo only contain a few posts concerning the initial stage of interaction and after that people stop writing about the robot.

With participants from a different age group, Klamer and colleagues [87] conducted a preliminary study to understand how elderly people use social robots at home, and which factors are relevant for people to build a relationship with the robot. A Nabaztag robot (Fig. 3.4d) was programmed to talk about health related activities in a personalised way with the three participants who took part in the experiment. They could answer the robot using yes- and no-buttons installed near the robot. The study consisted in interviewing the participants after 10 days of interacting with Nabaztag at their homes. The interview questions were about the general use of the robot (usefulness, contrast with initial expectations, and so on), perceived enjoyment while interacting with the robot and other relational factors such as perceived trust, credibility and likeability. Even though the sample size was very limited, the study points out interesting utilitarian and social factors of robots that deserve further attention. For example, the participant who found the robot useful was also the one who named the robot differently and stated that she built a relationship with the robot, which suggests that the utilitarian aspect of the robot is a major determinant for people to establish a social relationship with it.

Discussion

A summary of the studies described above is presented in Table 3.4. As domestic robots are becoming a reality with the arrival of commercial products such as Roomba, significant research has been made towards studying people's acceptance of robots in home environments. In fact, the earliest study presented in this section [89] was not performed in users' own homes, but in a home-like environment. More recently, the existing long-term studies in domestic environments so far used commercial robots, as research robots are still not stable enough to run out of laboratory conditions for extended periods of time without supervision. However, commercial robots still have limited capabilities, causing a gap between people's initial expectations and what they really experience after the initial interactions with the robot. These were consistent findings in some of the studies presented here, for example in [141] and [51]. On the other hand, as suggested in the Robot House study [89], even when the novelty effect fades away people allow robots to come closer to them in later interactions. This result highlights that there is potential for social robots in home environments, as long as they are capable of engaging users over extended periods of time. To do so, the robot's functionality (in other words, how it can assist users in their home routines), appears to be a major determinant.

Evaluating users' subjective experiences with a social robot in their home requires more original data collection methods than in any other environment. Quantitative measures such as the number and duration of the interactions or video recordings are usually not suitable in this case (unless when controlled by the families as in the Pleo study). Therefore, researchers have to come up with novel ways to gather user information while keeping users' privacy. A good example on how to overcome this can be found in [141], where traditional interviews were complemented with other methods such as drawings, probing techniques and checklists of the user activities.

The framework proposed by Sung et al. [141] is particularly relevant not only to domestic settings, but also to the field of long-term human-robot interaction in general. The temporal steps identified in their study (*pre-adoption*, *adoption*, *adaptation* and *use and retention*) are similar to those reported in HCI long-term studies (for example, see [82]), which suggests that several methodologies and practises employed in HCI for measuring user experience over time can be applicable to HRI.

Table 3.4: Summary of the long-term studies in Home environments.

References	**Robot**	**Capabilities**	**Exp. Design**	**Nr. Sessions**	**Main Results**
Koay et al. (2003)	PeopleBot	Approach the user in several ways	Subjects: 12 (8 male and 4 female) Measures: proxemic preferences Methods: questionnaire, comfort level device	8 (aprox. 1 hour each)	People's preferences in terms of promixity change over time.
Sung et al. (2009, 2010)	Roomba	Vacuum cleaning, move around the house	Subjects: 48 (across 30 households) Measures: acceptance of robot Methods: observation, interviews, probing techniques, activity cards, small questionnaires	6 months	Two months is the time required for observing stable interactions between robots and households. Several techniques should be complemented to really capture people's routines at home.
Fernaeus et al. (2010)	Pleo	Animal-like behaviour	Subjects: 6 families Measures: exploratory study Methods: interviews, video recordings and pictures	2-10 months	Initial expectations about Pleo were not met. After the novelty effect, participants played with the robot only occasionally.
Klamer et al. (2011)	Nabaztag	Personalised health conversations; users interact using yes- and no-buttons	Subjects: 3 (50-65 years old, females) Measures: usage and acceptance of social robots Methods: interviews	10 days	Utilitarian and social factors seem important reasons for participants to accept social robots in domestic environments.

3.1.5 Summary of Results

The presented studies show that, in the last few years, significant research has been made towards studying how robots and virtual agents can be improved to engage users for extended periods of time. Nevertheless, this is a very recent area, which means that most of the presented studies were exploratory. The purpose of the majority of the experiments was to gain familiarity with the environment where the agents would be placed, and to better understand the nature of the situations that may arise when users interact with these agents for extended periods of time.

The preferred data collection methods reflect the exploratory nature of these studies: video observation and interviews were often preferred to other quantitative methods such as questionnaires. This also happened because in most studies the sample size is limited (which is understandable given the difficulties in recruiting participants for long-term studies) and it is more informative to analyse user's behaviour and open-ended answers. However, when video material was collected, a deep analysis of all the interactions is usually inexistent, because video annotation is very time consuming. For example, in the Keepon study, only the first 15 sessions (out of 30) were analysed, and the results of the Paro studies collected in the nursing homes have not been analysed or the results are not published yet.

Even though evidence suggests that people are willing to accept and interact with robots for extended periods of time, there is still a lot of work that needs to be done, for instance in terms of understanding which social capabilities should robots be endowed with to engage users in the long-term. We believe that, with this thesis, we are taking another small step contributing to the development of this field.

3.2 Empathy in Virtual Agents and Social Robots

This section is divided in two main parts. First, we survey existing computational models of empathy. Then, we provide a literature review on relevant studies about the effects of empathy both in virtual agents and social robots.

3.2.1 Computational Models of Empathy

McQuiggan and Lester [104] argue that there are two different approaches for developing computational models of empathy: analytical and empirical. The analytical approach comprises models inspired on the theoretical empathy research (e.g., psychology, neuroscience, etc..), while the empirical approach is based on the analysis of human-human empathic behaviours exhibited while doing a specific task.

As we will see from the works presented below, at the moment there is no preferred approach for building computational models of empathy. If, on one hand, analytical models are very difficult to specify due to the lack of a common definition of the internal processes of empathy in the literature, on the other hand, empirical approaches can be very domain dependent and thus difficult to generalise to other application scenarios.

Analytical Approaches

Rodrigues et al. [125] extended an affective agent architecture with an empathic model based on the theoretical work of Vignemont and Singer [42] and in the Perception-Action Model [116]. The model contains two main processes, *empathic appraisal* and *empathic response*. The *empathic appraisal* is activated when the agent observes emotional cues in another agent (note that they do not specify any particular cues, because different scenarios might benefit from different affect recognition techniques). The observed emotional cues, combined with a *self-projection appraisal* mechanism where the agent appraises the event that caused the emotional cues in the other agent assuming the other agent's perspective, will generate a potential empathic emotion. This potential empathic emotion will then be modulated by a set of factors, namely the degree of similarity between the two agents, their affective link, as well as the mood and personality of the agent. In the *empathic response* phase, the resulting empathic emotion obtained from the *empathic appraisal* will trigger domain-dependent empathic actions, which can be predefined in the agent's behaviour. A study where subjects observed two different videos of virtual characters interacting in a school scenario (one with the empathic model and another without) suggests that the empathic model had positive effects on people's perception of the videos. For example, most participants indicated that their favourite character was the one with strongest empathic responses, and the social relationships between characters (e.g., being a friend of) became easier to identify.

Boukricha and Wachsmuth [21] took a similar approach in EMMA, an "Empathic Multimodal Agent" designed to "feel" empathy for MAX, another virtual agent that acts as a tour guide in a museum. While MAX interacts with a human through pre-defined dialogue lines, EMMA is paying attention to MAX's emotional behaviour and its empathic processes are activated. EMMA's empathic model is composed of three main steps: *empathy mechanism*, *empathy modulation* and *expression of empathy*. First, the *empathy mechanism* is triggered when EMMA perceives that MAX is displaying an emotional facial expression. This is possible because both agents use the same repertoire of facial expressions to express their emotions. The *empathy modulation* happens in the second step, and it is influenced by factors such as EMMA's mood and relationship with MAX (e.g. familiarity, liking, etc.) and desirability-for-self. The modulation factors influence not only the intensity of the empathic emotion, but also the emotion type, which means that EMMA's empathic emotion might differ from the emotion perceived in MAX's facial expression in the previous step. In the third step, *expression of empathy*, the empathic emotion generated by EMMA is expressed through the agent's facial expressions, speech prosody and other behaviours such as eye-blinking and breathing. This model has some limitations for relying exclusively on facial expressions to infer the other agent's affective state. As acknowledged by [21], inferring emotions only from facial expressions can lead to misleading assumptions, and therefore they are also planning to include role-taking in their *empathy mechanism*.

Empirical Approaches

Concerning empirical models of empathy, McQuiggan and Lester [104] propose CARE, a data-driven framework for empathic assessment (when to be empathic) and empathic interpretation (how to be empathic). In this framework, training data was first collected by having users interacting with two trainers, a target and an empathizer, in a virtual environment where the CARE-models were then used. After this training phase, two empathy models were induced based on the interaction logs, one using a naïve Bayes classifier and another one with a decision tree classifier. The framework was tested by having subjects interacting with two virtual agents rather than the human trainers. The results suggested that participants considered the empathic behaviours inferred by CARE as appropriate and accurate as the behaviours of humans empathizing in similar situations.

The same research group used a similar method to investigate the effects of different types of affective feedback strategies in a learning virtual environment [124]. Two studies were carried out, the first one to evaluate users' responses to agents displaying parallel or reactive empathic feedback and, in the second study, the preferred strategies for particular situations and affective states of the user were employed by the virtual agents. The last study was focused on the quality of the agents' responses and on the effects of affective intervention. Even though the results were fairly similar in both studies, in the second study it was observed that when students were experiencing the affective state of confusion, if they received a helpful response they were more likely to remain in that state, whereas if receiving an inappropriate response they would probably change to a state of frustration. The studies also indicate that affective and motivational states such as flow, delight and boredom seem to be more vulnerable to the quality of feedback given than other emotions such as frustration.

Combining the two Approaches

There are also models that integrate the two approaches. For example, Ochs et al. [109] combined theoretical descriptions of emotions with the empirical analysis of real human-machine dialogues to extract causes and consequences of situations that have lead users to express emotions. The inferred model of empathic emotions was implemented in a virtual agent that users can interact with to obtain information about their emails. Users communicate with the agent using predefined sentences that the agent uses to infer the user's emotions and, if appropriate, generate empathic emotions in return. An evaluation was carried out in this scenario using three different versions of the agent: one version that did not display any emotions, another one displaying empathic emotions and finally a virtual agent displaying incongruent empathetic emotions. The results indicated that the empathic virtual agent was perceived as more expressive, jovial and cheerful than the agent that did not display any emotions. Moreover, the agent displaying incongruent emotions was perceived as less pleasant, more irritating, cold and stressful than the non-emotional version of the agent.

3.2.2 Studies with Empathic Agents

Virtual Agents

Several researchers have been addressing the effects of empathic agents in different application scenarios. One of the early studies was carried out by Klein and Picard [88], who developed an interactive text-based agent to relieve user frustration caused by an intentional faulty application. The agent used active listening, empathy and sympathy with the intention of helping to relieve the user's negative states. Even though the self-report data did not provide significant results, the behavioural analysis of users showed that they continued to interact with the faulty system significantly more after interacting with the affect-supportive agent, when compared to other two control conditions. These results suggest that affective-supportive agents can indeed undo some negative feelings of users. However, from this study it remains unknown which components of the agent's behaviour (e.g., active listening, empathy, etc.) contributed the most for this result. This hypothesis was further explored by Hone [71], but this time with an embodied virtual character. Through a series of three different experiments, not only it was found that subjects interacting with an empathic text-based agent experienced significant reductions in self-reported frustration levels, but also that a virtually embodied character is even more successful at achieving this purpose. The author attributes this effect to a "good match between the characteristics of the feedback strategy (human-human) and the characteristics of the entity delivering that feedback". The third study showed that there were no significant differences in the frustration levels of female and male participants while interacting with a female or a male embodied virtual agent.

Another example of an empathic agent capable of assisting users in stressful situations is the Empathic Companion [115], a virtual agent that aids users during the simulation of a job interview scenario. The application collects physiological signals from the user in real time and provides affective feedback using utterances such as "it seems you did not like this question so much". The results of an exploratory study indicate that users interacting with the empathic agent felt less stressed while the interviewer was asking questions, when compared to users that interacted with a "neutral" version of the agent.

The ability to express empathy and provide comfort to users is also very valuable in agents employed in health-related applications. With this in mind, Bickmore and Shulman [16] evaluated two different conditions of an embodied conversational agent: an empathic version, where the agent constantly asks the user how he/she is feeling and provides appropriate empathic feedback, and an expressive version which asks the same question to the user but, even though users can provide free speech input (in the empathic version they had to choose between a set of predefined answers), the agent always responds in the same way. The result of a trial with these two agents indicate that subjects preferred the empathic agent, even though this version restricted the way users could interact with the system. Note that in another virtual agent developed by the same author, Laura (see section 3.1), one of the relational strategies that the agent employed was empathic dialogue (e.g., if the users says that he/she is not feeling well, Laura might say "I am sorry to hear that"). More recently, other authors investigated the effects of empathic agents in comforting situations [108], concluding that empathic agents can improve user's attention and willingness to interact with a system. Moreover, human-like agents without empathic capabilities can lead to a negative user experience due to the expectations that users may create while interacting with such agents.

Many studies have also highlighted the benefits of employing empathic agents in game scenarios. For example, Brave et al. [22] showed that empathic agents in a blackjack game were perceived as more caring, likeable and trustworthy than agents without empathic capabilities, and that people feel more supported in the presence of such agents. Another study explored the impact of a virtual opponent by analysing users physiological responses while playing against the agent [114]. The results of this study indicate that, in a competitive game scenario, if the agent lacks emotional expression or displays only positive empathic behaviours (e.g. showing happiness when the user is winning against the agent), the user becomes more stressed, i.e., her skin conductance increases. An important conclusion from this study is the fact that how the agent behaves cannot be decoupled from the interaction scenario and the agent's role in that scenario.

Another area where very often virtual agents are endowed with the ability to recognise and respond to some of the users' affective states is Intelligent Learning Environments. These pedagogical agents [76] are designed to intervene usually when the student is experiencing a negative affective state, in the attempt to reverse this state and with the ultimate goal of improving the learning experience. For example, Pour et al. [2] investigated how positive, neutral and negative feedback responses of Autotutor, an Intelligent Tutoring System that is represented by a virtual agent, influenced the leaner's affective and physiological states. The results suggest that there is a correlation between the affective feedback provided by the agent and (1) the self-reported users' states and (2) the users' affective states obtained from physiological data, which indicates that pedagogical agents can influence the learner's state. In the same line of research, Chaffar and Frasson [30] tested the impact of different emotional coping strategies employed by a virtual tutor in a scenario where users faced two different situations. In the first situation, subjects experienced a misunderstanding in a topic that was being developed by the agent and, in the second situation, participants obtained marks in an evaluation test. They found that problem-focused strategies, such as providing a definition for the topic or giving an example, are most effective in situations where users did not understand the topic. However, these results were not observed when the agent employed emotion-focused strategies (e.g., encouraging, congratulating, etc.). This may have happened because EMG signals, the method used to assess participant's valence after the interventions, was not enough to capture the changes in the users' affective state.

Social Robots

In the field of social robotics, researchers have only more recently started to assess the effects of endowing robots with empathic behaviour. One possible reason for this is that it is relatively easier to infer the user's affective state in interaction scenarios with virtual agents than it is with human-robot scenarios. Usually, while interacting with virtual agents, the user is sitting in front of a computer screen and there is often the possibility of selecting predefined dialogues to inform the agent of the user's emotional state, as in [19]. The interaction with robots tends to be more open-ended and thus perceiving the user's affective state is more challenging. Nevertheless, along with the significant improvements in automatic affect recognition using different modalities such as vision, speech or physiological signals [156], the study of empathy in social robots has witnessed significant progress in the last few years.

The early studies in this field focused on mimicking the user's affective state, a particular aspect of empathy also designated as emotional contagion. For instance, Hegel and colleagues [67] present a study with an anthropomorphic robot that recognises the user's emotional state through speech intonation and then mirrors the inferred state using a corresponding facial expression. The results suggest that users who interacted with this version of the robot found the robot's responses adequate both in terms of appropriateness to the social situation and timing, than subjects who interacted with the robot without affective expressions. In another study [123], a robot with the form of a chimpanzee head mimics the user's mouth and head movements. When interacting with this robot, most subjects considered the interaction more satisfactory than participants who interacted with a version of the robot without mimicking capabilities.

More recently, Cramer et al. [36] studied how empathy affects people's attitudes towards robots. In a between-subjects design, two groups of participants saw a four-minute video with an actor playing a cooperative game with an iCat robot. The experimental manipulation consisted in causing the robot to express empathic behaviour towards the actor in an accurate or inaccurate manner (i.e., incongruent to the situation), depending on the control group. In this study, there was a significant negative effect on user's trust in the inaccurate empathic behaviour condition. Conversely, participants that observed the robot displaying accurate empathic behaviours perceived their relationship with the robot as closer.

A relevant finding in empathy research is that empathy is positively linked to prosocial behaviour [47, 69]. In another study using the iCat robot, Saerbeck and colleagues [129] investigated the effects of social supportive behaviours of the robot on student's learning performance and motivation. Among other supportive behaviours, the robot conveyed basic empathic responses through its facial expressions. The results indicate that simple manipulations in terms of the robot's supportiveness, while maintaining the same learning content, increased student's motivation and scores on a language test.

3.2.3 Discussion

In this section we reviewed existing computational models of empathy and studies on the effects of empathic agents (either virtual characters or social robots). Although there are some computational models of empathy, most of the works in this domain rely on *ad hoc* domain-specific behaviours, possibly because empathy is not fully explained in the areas of cognitive and social sciences. Also for this reason, the type of empathic behaviours and prosocial strategies used by empathic agents are very domain dependent, and results are difficult to generalise to other domains.

There are some limitations in the works presented in this section. First, most of the studies do not use reliable methods for assessing user's affect, which means that agents might be employing inappropriate empathic responses and some of the beneficial effects of empathy might be biased for this. Another issue is related to the difficulty of isolating which empathic behaviours (or other behaviours of the agent) influenced the user's perception of the agent. For example, in the study of Saerbeck et al. [129], several supportive behaviours were implemented in the robot, and it remains unknown which of these behaviours contributed the most for the end results. Furthermore, as described in the literature, empathic responses can go beyond simple emotional reactions, and very often they are realised through more prosocial actions or coping behaviour. However, little exploration has been made for example on the effects of prosocial actions that can be taken to reduce users' levels of distress. Finally, to our knowledge, in all the evaluations involving empathic agents or robots, users were exposed to such agents only for a short period of time or in a single interaction. Thus, long-term interaction studies are necessary to verify whether the benefits of empathic agents still apply after a certain period of time.

3.3 Concluding Remarks

In this chapter we reviewed existing related work on long-term interaction studies (with special emphasis on studies with social robots), computational models of empathy and studies on the effects of empathic virtual agents and social robots.

By analysing the presented long-term studies, we can conclude that most of the existing virtual agents or social robots create extremely high expectations in users, and their current capabilities are not enough to keep such users' expectations after repeated interactions. One of the features that is often missing is the capacity to understand, adapt and respond more appropriately to the user's affective and motivational states or, in other words, the ability to empathise with users. In human-human relationships, it is often argued that empathy facilitates the creation and development of social relationships [3], as discussed in the previous chapter. In addition, previous short-term studies using empathic agents and robots presented in this section yield promising results in endowing such characters with the ability to empathise. These results strengthen our hypothesis that empathy is crucial for social robots or agents to engage users for extended periods of time.

Chapter 4

Towards the Design of an Empathic Robot Companion

The goal of this thesis is to investigate the role of empathy in long-term interaction between children and social robots. The motivation for studying empathy comes not only from the extensive literature highlighting the role of empathy in social relations, but also from an exploratory long-term study that will be presented in this chapter. After presenting in more detail the scenario that will be used as a testbed for this thesis, we describe our first long-term exploratory study and another study to evaluate the influence of empathic behaviours on people's perception of a social robot.

4.1 Scenario

Our case study scenario consists of a Philip's iCat robot [148], an electronic chessboard from DGT Projects[1] and a computer where all the processing takes place (see Figure 4.1). The iCat acts as a game companion, playing chess with children that already have some basic chess knowledge. During the game, the robot can display several facial expressions such as happy, sad or surprised, by moving its eyebrows, eyelids, mouth, neck and body, and say utterances using a text-to-speech engine. Additionally, the robot's

[1] http://digitalgametechnology.com/

Figure 4.1: User playing with the iCat.

paws and ears contain lights and touch sensors.

4.1.1 Interaction Cycle

The interaction starts with the iCat waking up and inviting the user to play (e.g., by saying "let's play chess"). Users can play an entire game or a chess exercise, depending on the initial arrangement of the pieces on the chessboard.

The interaction follows the diagram presented in Figure 4.2. The user plays with the white pieces, and therefore is always the first one to start playing. After every valid move played by the user, the iCat provides feedback on the user's move by displaying an affective reaction. Then, the robot asks the user to play its move in chess coordinates. When the user plays the move requested by the robot, it replies with a confirmation signal (for example, a small utterance such as "ok, thank you" or a nodding animation). If the user does not play the right move, there is also a set of "disapproval" animations and utterances. The game continues until one of the opponents checkmates the other.

The architecture behind the robot is separated in three main modules: *game*, *emotion* and *animation*. In the remaining of this section, we provide more details on the *game*

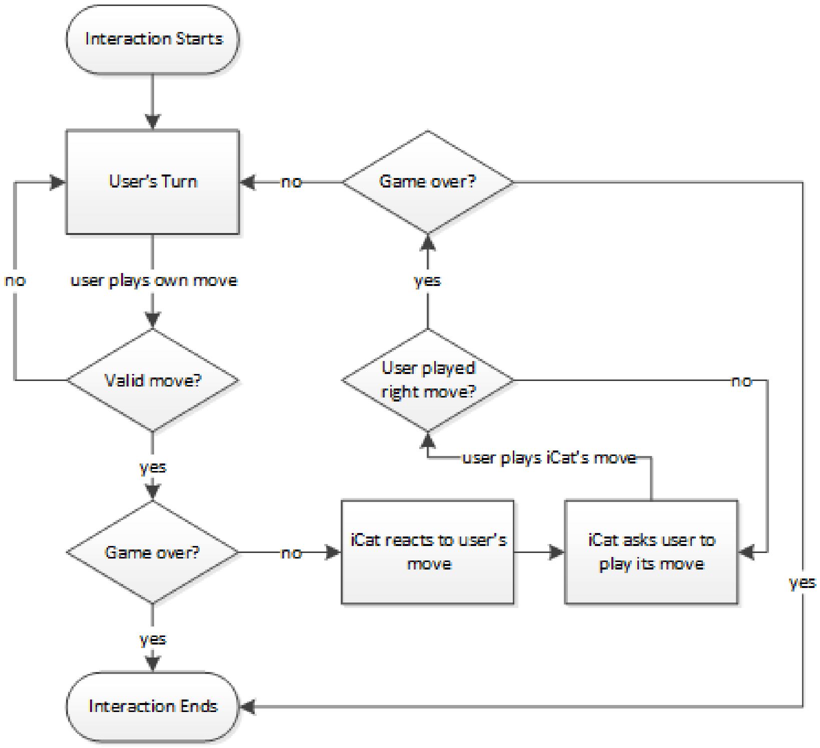


Figure 4.2: Overview of the interaction cycle.

and *animation* modules. The *emotion* module is the one responsible for generating the iCat's affective state and consequent expressive behaviour, influenced by the state of the game (retrieved from the *game* module). In the initial version of the scenario, evaluated in the preliminary long-term study of the following section, the iCat was competitive. This means that if the user was in advantage in the game, the robot would get sad, and if the user was losing the robot would become happier [96]. Since the generation of affect is one of the main focus of this work, more details on the *emotion* module of the robot will be provided in the following chapters.

4.1.2 Game Engine

The *game* module deals with the whole logic of the chess game. It contains an open source chess game engine[2] that is used not only to compute the moves of the iCat, but

[2] http://www.tckerrigan.com/Chess/TSCP

also to obtain a value that represents the evaluation of the game state after each user's move. The two most relevant methods of the Game Engine are the following:

- **play():** considering the state of the game, selects the best move for the robot to play in the next turn.

- **eval():** returns a value between -9999 and +9999 that estimates the value of iCat's position in the game. Positive values indicate that the robot is winning (the more positive, the better) and negative values that the robot is losing. The same value obtained from this function, but with an opposite signal, indicates the evaluation of the user's position in the game.

4.1.3 Animating the iCat

The main role of the *animation* module is to manipulate the iCat's body parts to convey the affective states and behaviours into the robot's embodiment. The degrees of freedom of the robot are depicted in Figure 4.3. The iCat's embodiment can be animated by two different modalities: predefined animations (scripts containing a temporal sequence of a set of body parts and correspondent values) or direct manipulation (i.e., by setting the value of single body parts in real time). These two modalities are used not only to express the robot's affective state, but also to convey other animations that increase the overall believability of the robot, such as blinking, looking at the user or looking at the chessboard when the user plays a move.

The iCat Research Platform includes the Open Platform for Personal Robotics (OPPR) software, which provides an animation engine that makes it possible to combine multiple robot animations at runtime execution. The animation engine supports the rendering of animations at 10 frames per second. Animations can include not only the movement the robot's mechanical parts, but also changes in lights, sound and speech, represented as *embodiment variables* in the OPPR system. This engine also allows direct manipulation, which are "dynamic" animations written in Lua Programming Language that directly change the values of the iCat's body parts. The iCat software comes with

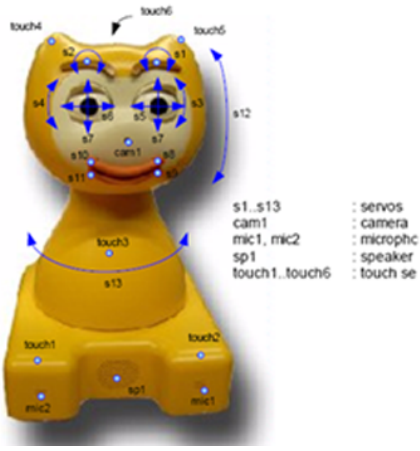


Figure 4.3: Degrees of freedom of the iCat robot.

a library of animations including some emotional animations (e.g. happy, sad and surprised) that have been previously submitted to tests ensuring that users perceive them correctly on the iCat's embodiment [10]. Additionally, the iCat software offers a lip synchronization mechanism that ensures the synchronization between the robot's mouth movements and the text-to-speech. The combination of multiple animations and body manipulations increases the believability of the character. This combination is possible notedue to the existence of ten animation channels with different priorities. Each channel controls the execution of a single robot animation.

When several channels are running animations at the same time and an embodiment variable is used concurrently, the chosen value will be the one set by the channel with highest priority. In our scenario, lip synchronization outranks the remaining animations in terms of priority in accessing variables (i.e., setting the value of a certain variable). After that, priority is given to the channel of the emotional expressions and finally to the "idle" animations such as looking at the chessboard.

4.2 *As Time Goes By*: Exploratory Long-term Study

Using the scenario described in the previous section, we conducted a first exploratory long-term study. The main goal was to analyse how children's behaviour towards the robot changed over time, in particular in terms of perceived Social Presence. Social Presence is defined by Biocca as "the degree to which a user feels access to the intelligence, intentions, and sensory impressions of another" [20]. This concept has been widely used to measure people's responses towards different technological artefacts, such as text-to-speech voices [93], virtual reality environments [66] and social robots [132]. In the following study, we evaluated whether children's social presence towards the robot changed over time and, if so, which dimensions of social presence were most affected.

4.2.1 Procedure

The experiment took place at a local chess club where every Saturday children between 5 and 15 years old take chess lessons from an instructor and play with each other. The class is composed of 7 children, even though we will only report the results of 5 of them (four males and one female) because the others missed more than one session. None of the participants had interacted with the iCat or with any social robot before.

At the chess club, the robot, the electronic chessboard and the computer were placed on a table. Participants were seated in front of both the iCat and the chessboard like in a regular chess game, as described in the previous section. A set of chess exercises was previously proposed by the chess instructor. He analyzed them and suggested modifications so that the difficulty of the exercises was adequate for each participant. In each session, subjects played a different chess exercise against the iCat. While the iCat was playing with one participant, the others were in the same room, and could be watching the game or playing against each other continuing their lessons. The idea was to integrate the robot in the group as one of their own. In this way, users were directly playing with the iCat and indirectly interacting with it during the remaining time. The experiment was performed over five consecutive weeks. All the sessions were video recorded for further analysis. At the end of the first and the last sessions, children were asked to fill a questionnaire that measures Social Presence.

4.2.2 Measures

We used a Social Presence Questionnaire based on the questionnaire developed by Harms and Biocca [62], which conceptualises social presence in six dimensions:

- **Co-presence:** the degree to which the observer believes she/he is not alone;
- **Attentional allocation:** the amount of attention the user allocates to and receives from an interactant;
- **Perceived message understanding:** the ability of the user to understand the message from the interactant;
- **Perceived affective understanding:** the user's ability to understand the interactant's emotional and attitudinal states;
- **Perceived affective interdependence:** the extent to which the user's emotional and attitudinal state affects and is affected by the interactant's emotional and attitudinal states; and
- **Perceived behavioural interdependence:** the extent to which the user's behaviour affects and is affected by the interactant's behaviour.

Two items of each dimension of the original Harms and Biocca's Social Presence Questionnaire were translated to Portuguese. The criteria for choosing the items was to select the possible sentences that children would find easier to understand. The final set of questions used in the experiment are described in Table 4.1 (translated back to English). Subjects were asked to express their agreement or disagreement regarding each item on a five-point Likert scale, in which zero meant "totally disagree" and five meant "totally agree".

In addition to the Social Presence Questionnaire, the videos from the first, second and fifth week of five users were analyzed using ANVIL video annotation tool [85]. We annotated the segments of video in which users were *looking at the iCat*, *looking sideways*, *talking to the iCat* and the *user's facial expressions*. We also distinguished the phase of the game in which users were looking at the iCat: *after the user's own move* – when the iCat performs an emotional reaction, *after playing the iCat's move* – when the user receives feedback from the robot, that confirms or disapproves his/her move, and while the user is *thinking* – when the iCat is performing idle behaviours such as blinking and looking sideways.

Table 4.1: Means of the Social Presence Questionnaire items in the first and fifth week, with the corresponding minimum and maximum values between brackets. The last three columns contain, for each question, the number of subjects who increased, maintained and decreased their ratings from the first to the fifth week.

	1st week	**5th week**	$\uparrow$	$\leftrightarrow$	$\downarrow$
Co-Presence					
Q1. I noticed iCat.	4 (4;4)	3,75 (3;5)	1	1	2
Q2. iCat noticed me.	3,75 (2;5)	3,75 (2;5)	0	4	0
Attentional Allocation					
Q3. I remained focused on iCat.	3,5 (3;5)	2,75 (1;4)	0	2	2
Q4. iCat remained focused on me.	3,75 (3;5)	3,25 (2;4)	1	2	1
Perceived Message Understanding					
Q5. iCat's thoughts were clear to me.	3 (2;4)	3,25 (2;5)	2	1	1
Q6. My thoughts were clear to iCat.	3,25 (2;5)	2,75 (2;3)	0	3	1
Perceived Affective Understanding					
Q7. I could tell how iCat felt.	3 (2;4)	3 (1;4)	2	1	1
Q8. iCat could tell how I felt.	2,25 (1;3)	2,5 (1;4)	2	0	2
Perceived Affective Interdependence					
Q9. I was influenced by iCat's mood.	3,75 (3;5)	3 (2;4)	0	1	3
Q10. iCat was influenced by my mood.	3,5 (2;5)	2,75 (2;3)	1	1	2
Perceived Behavioural Interdependence					
Q11. My behaviour was tied to iCat's.	3,5 (2;5)	2,25 (1;4)	0	2	2
Q12. iCat's behaviour was tied to mine.	3,5 (2;4)	2 (1;4)	1	0	3

4.2.3 Results

Social Presence Questionnaire

According to the results displayed in Table 4.1, in general, perceived social presence decreased after five weeks of interaction. Within the co-presence dimension, considering the means of Q1, there is slight evidence that users seem to notice the iCat less on the last week, which can be strengthened by the results of video observation. The amount of time that subjects looked at the iCat on the last interaction is lower than in the first ones. This may happen due to the novelty effect mentioned earlier, as none of the children had interacted with a social robot before. In spite of that, when asked if the iCat noticed them (Q2), all of them maintained their opinion. The turn-taking nature of the chess game may be the main cause for such result. Since the iCat reacts to children's moves and asks them to play its move, children might have interpreted those "reactive" behaviours as the robot noticing their presence.

After five weeks, the two items regarding attentional allocation, Q3 and Q4, did not increase. In fact, on average they decreased. From our observations at the chess club, when children are playing with each other, they comment their moves, refer to previous games, explain theories behind certain moves and sometimes even make fun of each other. The lack of some of these behaviours in the iCat could have been determinant for the lack of engagement of children after several interactions, especially the lack of "small-talk" related to the game.

In the perceived message understanding category, half of the users claimed to know better what the iCat was thinking after the five week period (Q5). Even so, when they were asked if their thoughts were clear to the iCat, most of them maintained their position. This also happens for the perceived affective understanding dimension (Q7 and Q8). For instance, on the last weeks of interaction, when the iCat reacted sadly to a good move from the user, some of them talked to the robot: "I know you don't like that". These results indicate that users felt that the iCat could not learn or adapt to their thoughts or affective states, and this affected the way they perceived the robot.

The last two dimensions (perceived affective interdependence and perceived behavioural interdependence, from Q9 to Q12) were the ones whose means decreased the most after repeated exposure. Over the weeks, the iCat seems to be perceived much more as an automaton, behaving independently of how users feel or act, only reacting to their moves: subjects were expecting the iCat to behave more like a companion than a mere chess interface.

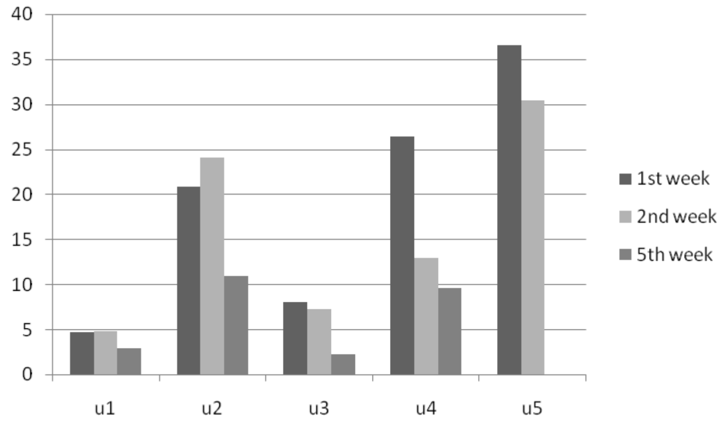


Figure 4.4: Total percentage of time that each user spent looking at the iCat in the first, second and fifth weeks of interaction (U5 did not attend in the fifth week).

Video Observation

From all the behaviours defined for annotation, looking at the iCat was the one with more annotations and also the one with more different results among different sessions. As such, it will be the main aspect of this discussion. As one can see in Figure 4.4, the novelty effect apparently does not fade away immediately. On most subjects there are no substantial differences between the first and second weeks. However, the total time that subjects spent looking at the iCat on the last session is, on average, half the time they spent on the first one. These results are aligned with the ones obtained in the Social Presence Questionnaire, especially with respect to the co-presence and the attentional allocation dimensions. A detailed analysis of the average time that users were looking at the iCat in each phase of the game is shown in Figure 4.5. Even though the average time decreases in all phases, *after playing the iCat's move* and *thinking* were the phases that decreased the most over the weeks. In *after user's own move* phase, this fall was not so pronounced (between the first and second weeks the values remained roughly the same). Furthermore, in the questionnaire, subjects maintained their opinion regarding the understanding of iCat's behaviour (Q5 to Q8). One possible explanation is that, over time, even though subjects keep looking at the affective feedback displayed by the iCat, they spend less and less time decoding its meaning.

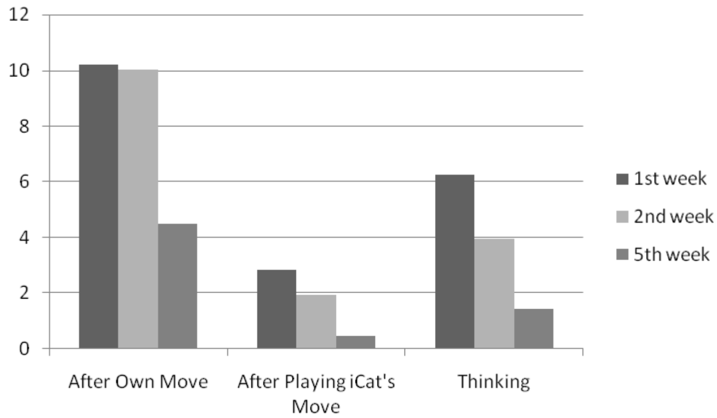


Figure 4.5: Average percentage of time that users looked at the iCat in the first, second and fifth weeks, broken down by the phases of the game. Since all the exercises had different durations, the results are presented as percentages of the total duration of the exercise.

The remaining annotations were considerably more sporadic than looking at the robot. For this reason, instead of presenting the quantitative results obtained from the annotation tool, we provide a brief qualitative discussion on each topic. In the annotations for users *looking sideways*, no significant differences were found between the first and the last weeks of interaction. In most cases, users looked away from the iCat or the chessboard due to some external event at the chess club (e.g. someone arriving at the club). Still, in the last weeks some participants expressed signals of boredom after playing their move, while waiting for iCat's affective reaction and consequent move. Some of them looked away, but only for short periods of time. Regarding user *talking to the iCat*, this annotation changed significantly among users and not so much over the weeks. Older subjects barely talked to the robot, but younger participants did. For example, when the iCat thanked users for playing its move, young subjects replied "you're welcome". Over the weeks, younger subjects felt less shy and started talking to the iCat even when it was not their turn to play. On the first week this did not happened, probably because they were not so comfortable in the presence of the robot and the experimenters. It remains to be validated if this behaviour would continue over subsequent interactions.

Concerning the user's facial expressions, we basically identified two different types: the ones users displayed when they did not understand an iCat's affective reaction and the ones performed in the end of the game. The expressions users displayed to show misunderstanding about the iCat's reactions decreased over the weeks, which again indicates that over time the perceived message and affective understanding dimensions of social presence tend to improve (or at least remain the same). There were no substantial variations on the user's expressions in the end game though. Usually, when winning the game, users showed happy faces and when loosing they made a sad expression or showed no expression at all.

4.2.4 Discussion

In this study, we investigated long-term human-robot interaction in terms of social presence. The results of the Social Presence Questionnaire were reinforced by the analysis of the video recordings of the interactions. The outcomes of the evaluation indicate that users' perception of social presence towards a social robot decreases after five weeks of interaction. We are aware that these results were obtained in a specific domain (a social robot in a chess game) and with a limited number of subjects, and therefore more experiments should be performed to see if the results can be generalised to other domains.

Overall, we concluded that the robot's current behaviour is not enough to create and maintain the perception of social presence after several interactions. Although it might appear believable and intelligent on the first impressions, *as time goes by*, users needed more.

One of the main contributions of this study that was valuable for this thesis upcoming work was the identification of the dimensions of social presence that decreased the most after five weeks of interaction. The questionnaire results indicate that attentional allocation, perceived affective and behavioural interdependence are the dimensions that decreased more over time. These results were reinforced by video observation. Furthermore, the video analysis suggested that co-presence decreases over time as well. The identified dimensions are mainly related to robot's believability and user's attention to the system. We observed that the attention that users dedicate to the robot decreased significantly over the weeks, which suggests that new mechanisms and behaviours must be developed in order to maintain the engagement. These results are also in line with the previous long-term studies presented in the related work chapter.

4.3 *iCat Watching Two Players*: Studying Empathic Behaviours

The outcomes of the previous study strengthened our hypothesis that the ability to understand and respond to the users' affective states is crucial for long-term interaction. Thus, we started exploring how specific empathic behaviours influenced people's perception of a social robot. For this study, a slight variation of the scenario described earlier was implemented, where the iCat watches and comments a chess match between two players. In this scenario, the iCat displays empathic behaviours towards one of the players, which will be referred to as "companion", and provides neutral comments towards the other player. Having the iCat in an outside position instead of directly as a player allowed us to evaluate empathic behaviours avoiding conflicts that may arise between expressing empathy and acting as an opponent (as observed in [114]). Also, with two players interacting at the same time we can simultaneously evaluate the two different conditions in the iCat's behaviour (empathic and neutral).

Even though there is not a common definition for empathy so far, most researchers (e.g., [47, 70]) agree that empathy can be achieved by two distinct phases: (1) inferring the other's affective state and (2) responding empathically to that state. Next we describe how the robot empathises with one of the players by inferring his/her affective state using perspective taking. After that, we present the study conducted to evaluate both quantitatively and qualitatively the influence of empathic behaviours on people's perception the robot.

4.3.1 Modelling Empathy: a First Approach

Figure 4.6 presents an overview of the first empathic model that we implemented in the iCat robot. When a new game event is received (i.e., when a valid move is played on the chessboard), the iCat infers the companion's affective state using perspective taking, that is, analysing the game considering the companion's point of view. The perspective-taking analysis influences the iCat's affective state, as the robot is intended to "feel" empathy for the companion player. The robot's empathic affective state then determines its empathic response, which is modulated by the relationship with the players. The responses are thus different whether the iCat is addressing its companion or the other player. In the following subsections, we describe in more detail how the perspective-taking mechanisms and empathic responses were implemented in this scenario.

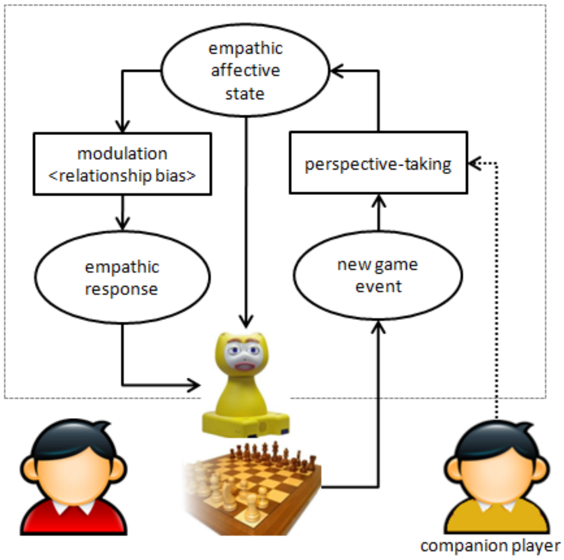


Figure 4.6: Overview of the empathic model in this scenario.

Inferring the User's Affective State: a Perspective Taking Approach

Following the literature on empathy research, the capacity of assuming the perspective of another person is a valid way of inferring that person's affective state. Our robot uses a perspective-taking approach to infer the possible affective states of its companion. The iCat may have a different perspective of the game than the companion, which may lead to wrong interpretations, but the same can also happen with humans. Previous work on automatic affect recognition has also shown that the contextual information of a chess game is enough to discriminate some of the user's affective states [28].

After every move played on the chessboard, the robot evaluates the new board position using a chess heuristic function (from the perspective of the companion player) and interprets it as if it were its own move. These values are the input of the robot's affective state, which is based on an anticipatory mechanism that generates an affective state based on the mismatch between an expectation of the next state of the game (*expected value*) and the real evaluation (*sensed value*) of the new board position [101, 94]. The expected value is computed based on the history of previous values obtained from the chess evaluation function, and the sensed value is the result of the current evaluation of the game. Using this mechanism, one of nine different affective states can be generated. For example, suppose that the companion is losing the game and plays a bad move; the iCat's consequent affective state is "expected punishment", meaning that the current state is bad, as the robot was expecting. Table 4.2 describes the list of possible empathic states and their meanings.

Empathic Responses

The affective state of the iCat influences the empathic responses the robot displays to the players. Empathic responses can range from simple facial expressions (often mimicking the other's facial expression) to more complex cognitive actions, including prosocial behaviours. However, the concrete prosocial behaviours associated with empathy are still not clearly defined. Eisenberg et al. [47] attribute this to the "failure of many investigators to differentiate among various modes of prosocial behaviour". Nevertheless, empirical studies that describe typical empathic and prosocial behaviours in particular contexts are starting to appear.

To re-create empathic responses in our robot, we took inspiration from one of such studies [34], where different characteristics of empathic teachers were identified and

Figure 4.7: Snapshots of the different empathic facial expressions of the iCat.

grouped by the following categories: body-language, positioning, teaching content, teaching method, voice, attitudes, facial characteristics and verbal responses. Given the limitations of our scenario in terms of the robot's embodiment and interaction context, we only considered differences in facial characteristics and verbal responses.

Facial Expressions

Empathic teachers constantly reflect the student's emotions in their facial expressions, while non-empathic teachers are often not in tune with what they are saying. Likewise, our robot expresses its empathic state using a facial expression that tries to reflect the companion's affective state by assessing his/her situation in the game. For instance, if the companion is constantly playing bad moves and plays another bad move, the iCat's affective state would return the empathic state of "expected punishment", which means that the change in the current state of the game is bad for the companion player, and the robot was already expecting that. In this case, the iCat would express a low intensity (because it was already expected) sad expression as a result of its empathy towards the disadvantaged companion. However, as the robot analyses the game and reacts according to the companion player's perspective, if the other player makes a mistake in his/her next move, causing the companion to recover some disadvantage, the iCat's affective state would change to "weaker punishment". In this situation, the iCat would display a subtle positive facial expression congruent with the possible affective state of the companion. Figure 4.7 presents an overview of the robot's empathic facial expressions.

Empathic teachers also tend to use more eye contact. This characteristic was also modelled in our robot: while players are pondering their next moves, the iCat displays idle behaviours, such as looking at the chessboard or the players. These behaviours were weighted such that the robot looks at its companion twice as often as the other player.

Verbal Comments

As mentioned above, empathy is positively correlated with prosocial and helping behaviours. We modelled prosocial behaviours in our robot by providing verbal feedback to the players. However, as our aim is to vary the empathic responses based on the relationship of the iCat with different players, the verbal comments depend not only on the iCat's empathic state, but also on the player who made that move. In other words, the verbal empathic responses of the robot are modulated by the relationship between the robot and players.

After reacting empathically to an event on the chessboard, the robot verbally comments on that move. As the comments depend on the iCat's empathic state and on the player who made that move, two sets of utterances were defined for each affective state: *empathic* utterances, to be used when the robot addresses its companion, and *neutral* utterances, to be used when the robot comments on the moves of the other players. While neutral utterances merely indicate the quality of the move in a very straightforward way (e.g., "bad move", "you played well this time"), empathic utterances often contain references to the companion's possible feelings in the attempt to encourage and motivate the companion (e.g., "you're doing great, carry on!"). Table 4.2 contains examples of possible utterances of the robot in both conditions. The utterances are sometimes combined with the name of the players either at the beginning or end of the sentence.

As with eye contact, when speaking to its companion, the iCat uses the companion's

Table 4.2: Description of the possible affective sensations experienced by the iCat, with examples of the verbal comments of the robot for both conditions. In the neutral utterances, the comments only vary depending on "reward" and "punishment" sensations.

Affective State	**Meaning**	**Examples of Verbal Utterances**	
		Empathic	**Neutral**
Stronger Reward	Better than expected	"Great move! Even better than I was expecting!"	"Good move." "You played well this time."
Expected Reward	As good as expected	"Nice move, you played what I would have played!"	
Weaker Reward	Not as good as expected	"I believe you could have played better."	
Unexpected Reward	Good, unexpected	"You played very well."	
Unexpected Punishment	Bad, unexpected	"Oh... I wasn't expecting that move..."	"This move wasn't that good." "Bad move."
Weaker Punishment	Not as bad as expected	"Well done, you are recovering your disadvantage!"	
Expected Punishment	Bad, as expected	"Don't worry, you didn't have better options."	
Stronger Punishment	Worse than expected	"You're making me sad... you could have played better.."	

name twice as often as when speaking to the other player. In addition to feedback after the moves, the iCat also congratulates the companion when she/he captures an opponent's piece and encourages her/him in the critical moments of the game – for example, when the chances of winning (or losing) become evident. At the end of the game, the robot's behaviour also differs depending on who won the game. If the companion wins, the robot says "Congratulations! You are a very good chess player!" (while displaying a happy facial expression); if the other player wins it simply says "Checkmate, see you on the next match" (displaying a sad facial expression).

4.3.2 Procedure

A between-subject experiment where participants played a chess game against each other with the iCat commenting on their games was performed. Forty subjects, 36 male and 4 female, with ages ranging from 18 to 28 years, took part in the experiment. Since this was an exploratory study, and was planned for a time when children would be on holiday, we decided to use young adults. To recruit participants for the study, an email asking for volunteers who knew how to play chess was sent to the student mailing list of

Figure 4.8: Users interacting with the iCat.

the Computer Science Department at a Portuguese University. The first 40 participants who replied to this email and met the necessary requirements for the study (i.e., knew how to play chess and never interacted with the iCat before) were selected.

The selected participants were assigned to a time slot (2 participants in each slot) depending on their availability. At the scheduled time, they were asked to sit in front of each other, having the iCat on their side commenting on the game (in Portuguese), and played an entire chess game (see Figure 4.8). Interactions lasted an average of one hour. At the end of the game, participants were guided to different rooms where they filled out a questionnaire and answered some open-ended questions. The initial part of the questionnaire contained some demographic questions, including participant's age and chess expertise (beginner, intermediate or advanced). 21 participants considered themselves as beginners, 17 as intermediate and 2 considered themselves to have advanced chess knowledge.

4.3.3 Conditions

The behaviour of the iCat varied whether the robot was addressing the player controlling the white or the black pieces. Participants playing with the black pieces belonged to the *empathic* condition, which means that the iCat behaved towards them according to the empathic behaviours described in the previous section. Participants playing with the white pieces were part of the control condition (*neutral*), and the robot addressed them in a neutral way. The assignment of subjects to the different conditions was done randomly.

Table 4.3: Friendship Questionnaire items used in this study (translated to English) with the corresponding means and standard deviations for each item.

Function	**Assertions**	**Empathic**		**Neutral**	
		M	***SD***	***M***	***SD***
Stimulating Companionship	I had fun playing with iCat by my side.	4.4	0.5	4.0	1.2
	iCat behaved as my companion during the game.	3.6	0.7	2.9	0.8
	iCat made me laugh.	4.8	0.4	4.8	0.9
	I enjoyed talking to iCat.	3.4	1.2	2.6	1.2
	It was nice being with iCat.	4.0	0.7	3.8	0.7
Help	iCat helped me during the game.	3.0	1.0	3.1	0.8
	iCat's comments were useful to me.	3.8	0.8	3.8	0.7
	iCat showed me how to do things better.	2.7	1.1	2.7	1.4
	iCat's comments were helpful to me.	3.0	1.2	2.9	1.1
Intimacy	iCat knew when something was bothering me.	2.9	1.2	2.1	1.0
	iCat knew when I was upset.	2.9	1.2	2.1	0.8
Reliable Alliance	iCat was loyal to me.	4.1	0.7	3.1	1.2
	iCat would still be my friend if we did not see each other for several months.	3.6	1.0	2.7	1.2
	iCat would still be my friend even when we did not agree on something.	4.0	0.8	3.3	1.1
Self-Validation	iCat encouraged me to play better during the game.	4.0	0.7	2.9	1.3
	iCat praised me when I played well.	4.8	0.4	4.3	0.7
	iCat made me feel intelligent.	3.4	1.0	2.8	0.9
	I felt that I could play better in the presence of iCat.	3.9	0.9	3.8	0.9
	iCat enhanced the aspects that I am good at.	3.4	1.1	3.1	0.9
	iCat made me feel special.	2.6	0.9	2.7	1.0
Emotional Security	If I were worried, iCat would make me feel better.	3.0	1.0	2.7	1.2
	If I were nervous, iCat would make me feel more calm.	3.4	1.1	3.0	1.4
	If I were upset, iCat would make me feel better.	3.3	1.0	3.0	1.3

4.3.4 Measures

To quantitatively measure the influence of empathic behaviours on the relationship between participants and the iCat robot, we took inspiration from the McGill Friendship Questionnaire [105]. This questionnaire quantifies the degree to which a friend fulfils the functions existing in most friendship definitions:

- **Stimulating Companionship:** doing enjoyable or exciting things together;
- **Help:** providing guidance and other forms of aid;
- **Intimacy:** being sensitive to the other's needs and states and being open to honest expressions of thoughts, feelings and personal information;
- **Reliable alliance:** remaining available and loyal;
- **Self-validation:** reassuring, encouraging, and otherwise helping the other maintain a positive self-image;
- **Emotional Security:** providing comfort and confidence in novel or threatening situations.

A friendship relation, having similar patterns as human-human friendship, cannot be established between a human and robot, especially in an interaction lasting approximately one hour. However, as empathy is a characteristic that people value in friendships [1], we believe that by analysing the friendship functions individually, some indicators on improving long-term interaction between humans and robots can be retrieved. Thus, we hypothesise that subjects towards whom the robot displays empathic behaviour will rate some of these friendship functions more positively.

As the participants' mother language was Portuguese, we used a validated Portuguese version of the McGill Friendship Questionnaire [136]. The questionnaire contains a set of five-point Likert scale items for each friendship function, ranging from 1 (totally disagree) to 5 (completely agree). We removed some assertions that were not applicable in this experimental set-up and slightly adapted some existing assertions to become more in tune with the type of interaction that users had with the iCat. Table 4.3 contains the English translation of the assertions used in the questionnaire and the corresponding average rankings that users provided for each item separated by the control conditions of the study.

In addition to the perceived friendship measure, we also investigated the overall goals and expectations of users after interacting with the iCat. To do so, we analysed participants' responses to the following open-ended questions: "I liked that iCat...", "When I played bad, iCat...", "When I played well, iCat...", "When I was feeling insecure about the game, iCat...", "What would make me interact again with iCat would be...".

4.3.5 Quantitative Results

To test our hypothesis, we first performed a Cronbach's alpha test to evaluate the internal consistency of the adapted assertions of the friendship functions. The index was reliable for stimulating companionship (α=.79), help (α=.86), reliable alliance (α=.69), self-validation (α=.8) and emotional security (α=.71), but not for intimacy (α=.56).

We noticed that a small group of participants took much less time to fill in the questionnaire than the majority of participants. These subjects also provided some inadequate responses to the open-ended questions and, in some cases, did not answer some questions, which suggests that they did not devote enough time to provide reliable answers. Therefore, we decided to remove these participants from the analysis. 4 participants were excluded from the *empathic* condition and 5 from the *neutral* condition; our final sample was 16 in the *empathic* group and 15 in the *neutral* group. None of the 4 female participants were removed.

After this step, we ran a Mann-Whitney test for each friendship function[3]. Each function was calculated as the average of all questionnaire items associated with that friendship function.

[3] Non-parametric tests were used because the data did not follow a normal distribution.

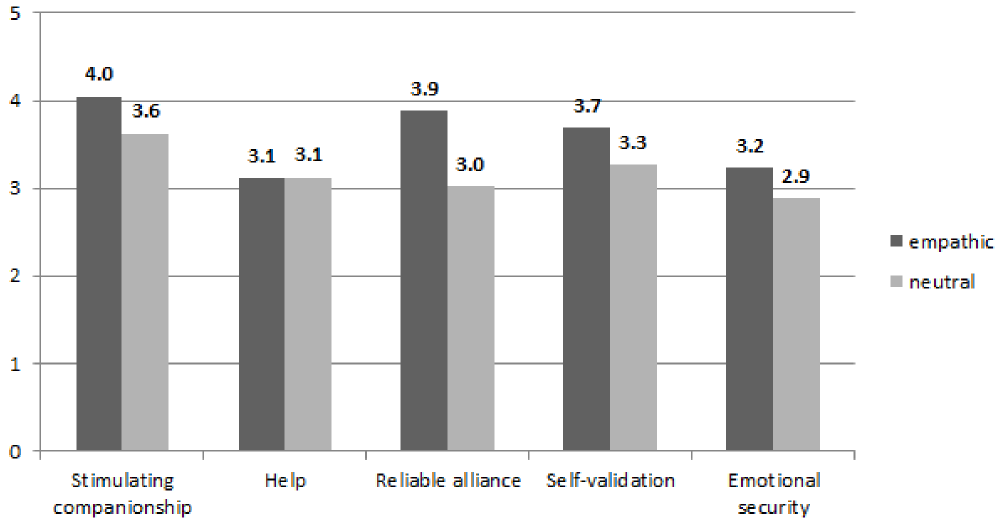


Figure 4.9: Mean values of the friendship functions (except intimacy) in the *empathic* and *neutral* conditions.

Stimulating Companionship

Subjects in the *empathic* condition rated this function significantly higher than subjects in the *neutral* condition (U = 72.5, p <0.05, z = -1.893). The chart in Figure 4.9 shows the stimulating companionship function means in each condition.

Help

Figure 4.9 illustrates that *empathic* and *neutral* conditions are not significantly different with regard to the help function (U = 118, p = 0.47, z = -0.79), suggesting that adding empathic behaviours does not seem to affect the feedback provided by the robot. It is also important to note that the ratings for the questionnaire items related to help were rather low in both conditions, as depicted in Table 4.3.

Intimacy

Since we could not guarantee the reliability of this function, we present the results of the proposed assertions for Intimacy separately (Figure 4.10). The ratings of the two statements, "iCat knew when something was bothering me" and "iCat knew when I was

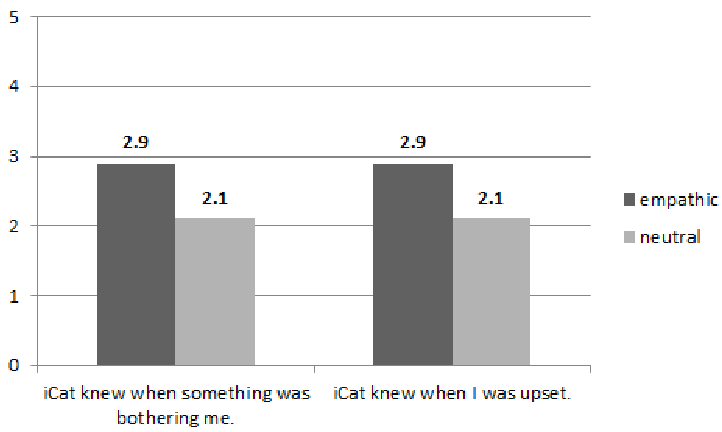


Figure 4.10: Mean values of the intimacy assertions in the *empathic* and *neutral* conditions.

upset", differed significantly with better ratings for the *empathic* condition (U = 69.5, p <0.05, z = -2.058 and U = 71.5, p <0.05, z = -2.033, respectively).

Reliable Alliance

The ratings of the reliable alliance function were significantly higher in the *empathic* condition than the *neutral* condition (U = 46, p <0.01, z = -2.954). This can also be verified in the chart displayed in Figure 4.9.

Self-Validation

In this function, there was also a significant difference in the direction of the *empathic* condition (U = 66.5, p <0.05, z = -2.131), as illustrated in Figure 4.9. The questionnaire item that contributed the most for this difference was "iCat encouraged me to play better during the game" (see Table 4.3).

Emotional Security

There was no significant difference between the two groups in the emotional security function (U = 95, p = 0.159, z = -0.998), despite slightly lower ratings by participants in the *neutral* condition in all questionnaire items of this function (see Figure 4.9).

4.3.6 Qualitative Results

In this subsection, the most relevant findings collected throughout the content analysis of the open-ended questions are presented. For each condition, participants' answers were grouped into several categories and the frequencies of these categories were analysed. As we ended up with an unbalanced number of subjects, the results are also presented as percentages.

"I liked that iCat..."

In both groups, participants highlighted the fact that the iCat provided feedback on their moves: 50% in the *empathic* condition and 67% in the *neutral* condition. Some participants also enjoyed that the robot used their names when speaking – 3 in the *empathic* condition (19%) and 2 in the *neutral* condition (13%). In the *empathic* group, 8 subjects (50%) mentioned that they liked the iCat because it encouraged them in the difficult moments of the game: *"iCat knew exactly the best moves I should play, and even when the game was almost lost, it kept giving me hope to continue (...)"*. One participant in this condition even mentioned that the iCat elicited an empathic feeling on him: *"I liked that the iCat used my name and commented on my moves. Its facial expressions and movement made me feel empathy."*

"When I played bad, iCat..."

Most users in both conditions recognised that the robot warned them about their bad moves – 10 (63%) in the *empathic* condition and 13 (87%) in the *neutral* group. One participant in the *empathic* condition even answered that the iCat got sad when he played poorly. The opposite answer was given by 2 participants (13%) in the *neutral* condition (they noticed that the iCat got happy when they played bad). 4 participants (25%) in the *empathic* condition also mentioned that when they played bad moves, the robot encouraged them to play better.

"When I played well, iCat..."

A substantial number of subjects answered that the iCat congratulated them when they played good moves (44% from the *empathic* condition and 67% from the control condition). 3 participants (20%) from the *empathic* condition stated that the robot got happy when they played good moves, while the same number of subjects in the control condition said that the robot got sad. As in the previous open-ended question, 8 subjects (50%) from the *empathic* condition also added that the robot gave them positive feedback: *"When I played a good move, iCat demonstrated his support."*

"When I was feeling insecure about the game, iCat..."

Nearly one-third of the participants indicated that they did not feel insecure in any part of the game and therefore did not answer this question. For those who answered, the opinions differed among conditions. While nearly half of the participants in the *neutral* group (47%) did not notice any differences in the iCat's behaviour, 6 subjects (38%) in the *empathic* group stated that the robot encouraged them when they felt insecure during the game: *"When I felt insecure during the game, the iCat tried to make me calm, so I could better play my next moves."* In fact, participants from the *neutral* condition recognised that the iCat supported their opponent more: *"It didn't help much... I got the feeling that iCat was supporting my opponent the whole time and didn't care about me."*

"What would make me interact with iCat again would be..."

Participants from both conditions proposed several motives for interacting again with the iCat. While in the *neutral* condition 7 subjects (47%) would like to interact again with the robot just for fun, 6 participants (38%) in the *empathic* group would like to play another game in this same setting to improve their chess skills. 7 participants stated that they would like to play against the iCat, instead of having the robot commenting on their game - 3 participants (19%) from the *empathic* group and 4 (27%) from the *neutral* condition. Additionally, 5 participants from the *empathic* condition (31%) would like to play another match in this same setting, in contrast to only 2 (13%) from the control version.

4.3.7 Discussion

This study addressed the role of empathic behaviours on users' interactions with a social robot. Empathic behaviours reported in the literature were modelled in a social robot capable of inferring some of the user's affective states, reacting emotionally to such states and commenting on a chess game.

Several friendship functions were measured using a questionnaire. The results suggest that our initial hypothesis is valid for most measured friendship functions. Participants who interacted with the empathic version of the robot provided significantly higher ratings of companionship, reliable alliance and self-validation. Considering that companionship is about spending time and doing things together, these results suggest that empathic robots can provide more enjoyable interactions. As such, users may eventually spend more time interacting with the robot, which is important if we aim to build companions capable of engaging users in long-term interactions. However, these assumptions must be verified by a long-term study. The results regarding reliable alliance can be explained by the iCat being empathic to the user in the empathic condition, remaining loyal to that player whether she/he is winning or losing the game. For self-validation, the encouraging behaviours that the iCat displays towards participants in the empathic version seem to have a positive effect on users from this group. Nevertheless, it is important to stress that, in both conditions, the iCat says whether the player did a good or bad move, which is also a form of reassurance.

Our initial hypothesis was not validated only in help and emotional security. However, given the experimental scenario design and the interaction length, such results were not totally unexpected. The lower and non-significant different scores on help are understandable, because even though the iCat gives feedback on both players' moves (in a more empathic manner for subjects in the empathic version), it does not provide any concrete suggestions on how to play the game better. For emotional security, this may be a function that requires long-term interaction before users consider that the iCat or another entity can comfort them. Therefore, a long-term study is needed to clarify this result.

Although the results for intimacy cannot be analysed within the scope of the friendship questionnaire, the empathic group provided significantly higher ratings in the two assertions individually. As we are inferring the user's affective state considering his/her situation in the game, these results support our previous work [28] in which we found that task-related features are important for discriminating among user's affective states.

The overall impression and understanding of the interaction was investigated by a content analysis of participants' answers to open-ended questions. This analysis revealed that subjects towards whom the iCat behaved empathically considered the robot more encouraging and sensible to their feelings. These results reinforce the findings obtained in the questionnaire, particularly that the empathic behaviours implemented in the robot were well understood and accepted by users. Some subjects in the neutral condition recognised that the iCat's behaviour was more empathic towards their opponent, but they still valued the robot's feedback on their moves. Players from the empathic group not only valued the robot's feedback, but they considered the robot's messages encouraging. More subjects from the empathic condition would like to interact again with the robot in this setting, and they would do so to improve their chess skills. These results indicate that empathy might be a relevant characteristic in social robots for educational purposes (e.g., robots who act as tutors).

4.4 Concluding Remarks

In this chapter, we introduced our application scenario and presented the two foundation studies that contributed to the solution that will be proposed in the following chapter. The first study investigated the changes in children's perception of a social robot after several interactions. We analysed the same group of children playing an entire chess game with the iCat over five sessions (once a week). The results suggest that social presence decreased over time, especially in terms of perceived affective and behavioural interdependence (the extent to which users believe that the behaviour and affective state of the robot is influenced by their own behaviour and affective state). The outcomes of this experiment, in particular the questionnaire results, highlighted our hypothesis that the ability to understand and respond to user's affect is crucial for long-term interaction, and motivated the second study that evaluated the influence of empathic behaviours on user's perception of a social robot. For this study, a slight variation of our application scenario was implemented, in which the iCat observes and comments a chess match between two players. The robot exhibits empathic behaviours towards one of the players and neutral comments to the other player, through facial expressions and verbal comments. The results of this study suggest that players towards whom the iCat displayed empathic behaviour perceived the robot as friendlier.

These two studies inspired our proposed solution in distinct ways. While the long-term study was important to acknowledge that perceptual and empathic capabilities are important for a social robot that aims to engage users for extended periods of time, the empathy study showed that, through the manipulation of simple empathic behaviours in the robot, people perceive the robot differently.

Chapter 5

An Empathic Model for Long-term Interaction

This chapter describes an empathic model for supporting long-term human-robot interaction. The model was inspired not only by theoretical work on empathy research, but also by the user studies presented in the previous chapter. In the following sections, we start by delineating the scope and inputs of the model, and then describe each one of the components individually. Finally, we present an implementation of the empathic model in the iCat scenario.

The generic empathic framework in which the model was established is depicted in the diagram of Figure 5.1. The diagram follows a traditional Perception and Action loop. The perceptions of the agent are the user's affective state and the information about the task the agent and the user are performing together. Using this data, the agent evaluates the context of the task and generates an empathic affective state. Additionally, considering the predicted affective state of the user, empathic actions are triggered by the agent.

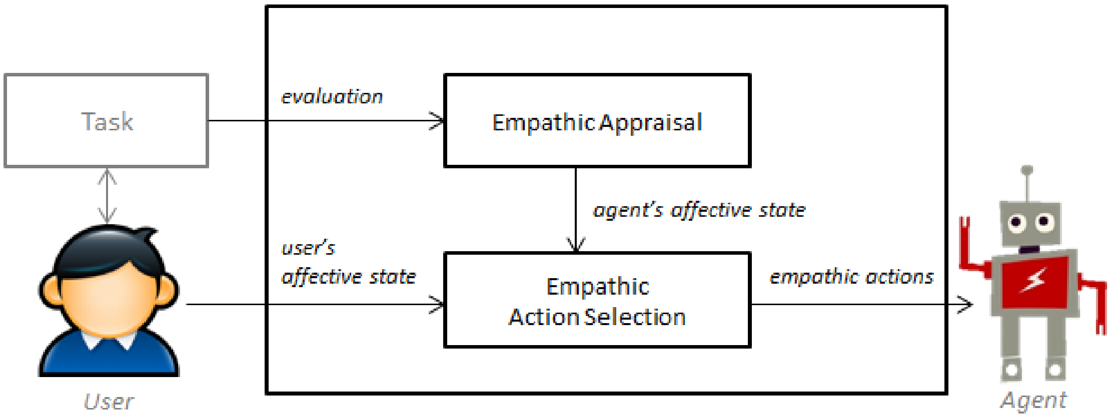


Figure 5.1: Generic empathic framework for long-term interaction.

5.1 Scope, Assumptions and Inputs

Before describing in more detail the main components of the proposed empathic model, we first need to determine its scope and describe the main assumptions and inputs. The empathic model is tailored for social robots that play games with users in real-world settings for extended periods of time. Although the model was developed specifically for Human-Robot Interaction, it also supports the interaction between users and virtual characters, since it is mainly a cognitive model and the behavioural layers, i.e., the way specific actions are conveyed by the agent, are outside of the scope of the model. As such, in the remaining of this section, we will use the term *agent* to describe the autonomous entity (robot or virtual character) who is interacting with the user.

To perceive information about the user and the context of the task, the model has two main inputs: a Game Engine and an Affect Recognition System. The Game Engine computes the agent's next moves and informs the Empathic Appraisal component about the state of the game (in the perspective of the user) through the variable $gamestate \in \mathbb{R}$. Although our focus in on games, the Game Engine component could be replaced by other tasks or perfect information turn-based games, as long as it was possible to translate the progress of such task in a value that could inform the Empathic Appraisal component about the user's progress. Being a turn-based interaction brings some advantages. For example, the agent has its own turn in the game, and therefore it knows *when* to inter-

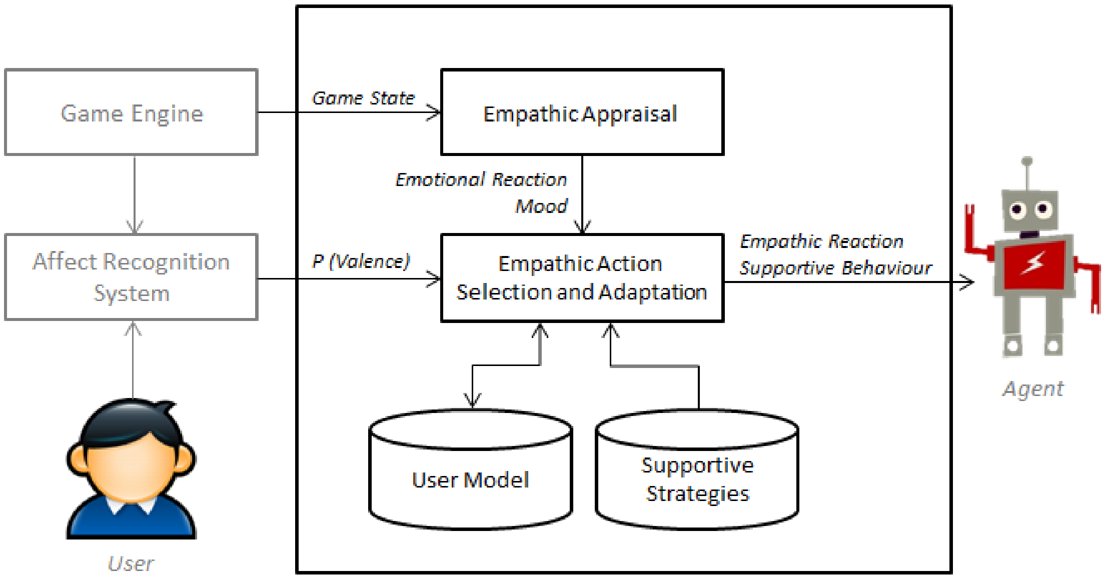


Figure 5.2: Overview of the proposed Empathic Model.

vene. The proposed model allows the agent to decide *how* to intervene, that is, which actions should the agent take to empathize with the user in an appropriate manner.

The Affect Recognition System is the other main input of our model. Since context can play a very important role in the prediction of the user's affective state [28], this component is connected to the Game Engine. We assume that this component infers the affective state of the user in real-time using several modalities and returns the estimate of the user's valence of feeling $P(valence) \in D = \{positive, negative, neutral\}$, where $\sum_{x \in D} P(valence = x) = 1$.

When a new game event is received and interpreted by the Game Engine, the Empathic Appraisal component selects an empathic affective state based on the user's performance in the game. After that, the Action Selection component determines, considering the user's affective state and on the previous history of interactions with that user, if additional Supportive Strategies are needed (besides the display of an empathic affective reaction). Finally, the User Model component is updated. In the following sections, the main components of the model are described in more detail.

5.2 Empathic Appraisal

Most of the definitions of empathy agree at least in one aspect: empathy is an affective response in tune with the other's affective state. As such, the Empathic Appraisal component calculates the agent's affective state using "perspective-taking", that is, appraising the situation that the user is experiencing from his/her own point of view, generating an empathic affective state. By taking inspiration from Scherer's work [131], which divides the affective states in five different categories (emotion, mood, interpersonal stances, attitudes, and personality traits), the empathic affective state of the agent incorporates the first two, emotion and mood. We decided not to include the remaining three categories because we do not have enough information from the outside world of the agent to allow a proper representation of those states. In this section, we describe how emotions (which we refer to as Emotional Reactions) and Mood are generated in our model.

Emotional Reactions

Emotional reactions are the immediate emotions experienced after a relevant event in the game (e.g., when the user plays a new move). They are relatively brief episodes in response to the evaluation of an external or internal event of major significance [131]. Although being of short duration, they are quite explicit.

These reactions are computed based on the *emotivector* model [101], an anticipatory mechanism coupled with a sensor that: (1) uses the history of the sensor to anticipate the next sensor state (*expected value*); (2) interprets the mismatch between the predicted *expected value* and the *sensed value*, by computing its attention grabbing potential and associating a basic qualitative sensation with the signal; (3) sends its interpretation along with the signal. At time $t-1$, the emotivector value is $x_{t-1} \in [0,1]$, Using its history at time $t-1$, the emotivector estimates a value for next time $t(\hat{x}_t)$, and predicts that its value will change $\Delta\hat{x}_t = \hat{x}_t - x_{t-1}$. At time t, a new value is sensed (x_t) and a variation $\Delta x_t = x_t - x_{t-1}$ is actually verified.

In our approach, the *expected value* is computed based on the history of previous *game state* values. As such, it can be interpreted as a prediction of how good or bad the game will be for the user after his/her upcoming move. Calculating the *expected value* requires an estimator for the prediction of the next state. Diverse estimators can be used, for instance Kalman filter or Moving Averages algorithm. For simpler approaches, one could even assume that the value for the next state is equal to the previous sensed value, but this would lead to larger mismatches between the estimated and the sensed value. By confronting the predicted *expected value* with actual *sensed value* (i.e., current *game state*), and using a model inspired in the psychology of attention, the *emotivector* returns one out of nine different sensations.

Table 5.1: Sensations elicited by the *emotivector* model [101].

Expected	**Sensed**		
	More Punishing	**As Expected**	**More Rewarding**
Reward (R)	Weaker R	Expected R	Stronger R
Negligible	Unexpected P	Neutral	Unexpected R
Punishment (P)	Stronger P	Expected P	Weaker P

Table 5.1 depicts the nine different sensations that can be generated by the *emotivector* model. If the current expected value is higher than the previous one, the generated affective signal will belong to the first row of sensations the table (more Reward); if there are no significant changes, a signal from the second row is selected, and if the value is lower, it belongs to the third row (more Punishment). Depending if the sensed value is higher, within a threshold or lower than the expected value, three different sensations can be elicited: stronger, expected and weaker. The threshold value is computed based on the history of mismatches between the *expected* and the *sensed* values. Computing the threshold value has the advantage of not requiring fine-tuning, which is fundamental to our model, since the signals picked up by the *emotivector* can vary depending on the game that is being used. Thus, the only adjustment that needs to be done when implementing this model is to define the negligible variation for the nine sensation model.

The sensations elicited by the *emotivector* model will be sent to the Action Selection component, where they will be mapped to an affective expression that reflects a certain emotion (depending on the available Supportive Strategies).

Mood

Mood is a relatively lasting affective state. It is less specific, often less intense and thus less likely to be triggered by a particular stimulus or event. Moods generally have either a positive or negative valence effect and are longer lasting [145], working like a background affective state when no other emotions are occurring. Based on this definition, we represent mood as a valence variable M that ranges between [-1;+1]. The magnitude of M represents the intensity of the mood. Positive values are associated to positive affective states (e.g., happiness), whereas negative values are related to negative affective states (e.g., unhappiness).

The *game state* in the perspective of the user is also the main stimulus for the mood variable. However, some pre-processing is required. First, because the limits of the *game state* variable may not lie between [-1;+1]. Second, because even if the previous condition is true, this mapping would only lead to linear correspondences, which can be undesirable in some cases. For example, consider that the values sent by the Game Engine correspond to an evaluation function that ranges between -1 to +1, where -1 means that the user is going to lose the game in the next turn and +1 means that the user is certainly going to win. In this type of functions, the boundary values only come up in the endgame. Thus, most part of the signals would be around [-0.3;+0.3], especially in the beginning of the game. Directly mapping these values would result in low intensity moods during the most part of the game. To overcome these situations, mood can also be represented as a function $M(gamestate) \in [-1;+1]$, which filters the values received directly from the Game Engine.

When a new *game state* is received, $M(gamestate)$ is computed and the mood variable gradually increases until it reaches the desired value. However, in the absence of new stimuli, mood tends to come back to zero. A *decay rate*, which determines how fast the valence decays over time, as well as the number of seconds before the decay should start (since the last event), must be defined and tuned.

5.3 Empathic Action Selection and Adaptation

As explained in chapter 2, empathy not only encompasses affective responses (e.g., facial expressions), but also prosocial actions to reduce the other's distress, especially when the person we are empathizing with is experiencing negative emotions. As such, the Action Selection component is responsible for realizing the affective signals and moods sent by the Empathic Appraisal component (i.e., translating them into concrete behaviours for the robot to display), but also for selecting supportive behaviours when the user's affective state is negative.

The diagram of Figure 5.2 shows that the Action Selection is influenced by the results of the Empathic Appraisal, the user's affective state obtained from the Affect Recognition System, the available Supportive Strategies and the User Model. Therefore, before going into more details on how the Action Selection works, we need to describe the two components that were not introduced yet: the Supportive Strategies and the User Model.

5.3.1 Supportive Strategies

This component contains all the available Supportive Strategies for the agent to display empathy towards the user. The supportive strategies are divided in two main groups: emotional expressions and supportive behaviours.

Emotional Expressions

The emotional expressions are basically the mapping between the empathic affective state computed in the Empathic Appraisal component (Emotional Reactions and Mood) into concrete expressions in the agent's behaviour. The database of emotional expressions depends on the expressivity of the agent's embodiment. Since our goal is to propose a generic model independent of the agent's embodiment, we do not propose concrete expressions at this level, but only general guidelines for future application of the model.

For the Emotional Reactions, this component contains at least one emotional expression for each one of the nine sensations resultant from the *emotivector* model. For example, the Unexpected Reward sensation can be mapped into one (or several) animation(s) reflecting surprise.

Since Mood is represented as a variable between [-1;+1], this module contains at least two different body parametrisations for the agent's expression, one positive and one negative, each one corresponding to a limit of this variable. The agent's mood is calculated as an interpolation from these two parametrisations, depending if the value is positive or negative.

Supportive Behaviours

We define supportive behaviours as the set of actions $S_1, S_2, ..., S_n$ that the agent can display (one at a time) with the goal of changing the user's affective state from negative to positive, that is, maximizing $P(PositiveFeeling)$ and consequently minimizing $P(NegativeFeeling)$[1]. These behaviours include not only verbal utterances (e.g., comments), but also concrete actions (e.g., game-related actions). Like the emotional expressions, prosocial actions depend on the agent's capabilities. Moreover, they depend heavily on the interaction scenario.

5.3.2 User Model

The User Model stores all the relevant user information to support long-term interaction. The following data is updated for each user at the end of every interaction:

- User's name
- Date of last interaction
- Total number of interactions

[1]Note that $P(PositiveFeeling) + P(NegativeFeeling) + P(NeutralFeeling) = 1$, which means that by maximising $P(PositiveFeeling)$, $P(NegativeFeeling) + P(NeutralFeeling)$ will decrease.

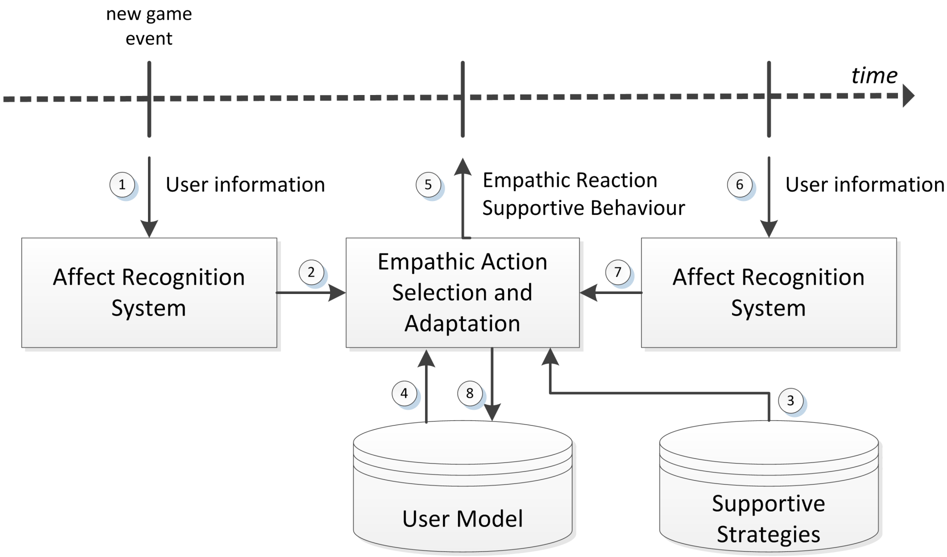


Figure 5.3: Partial view of the model with a timeline that exemplifies how the reward values are updated during the interaction.

- Previous game results (number of victories, defeats and draws)

Additionally, the User Model also stores specific information about the Supportive Strategies. In particular, the list of all the Supportive Behaviours employed by the agent during the game are saved, as well as the user preferences regarding the different support behaviour categories. Using this information, the agent can adapt to the user by selecting the most effective strategies for that user. The mechanisms that allow long-term adaptation to the user are described in the following subsection.

5.3.3 Adaptive Decision-Making

After a new game event, the Affect Recognition System informs the agent about the new affective state of the user. Consequently, the Empathic Appraisal updates the affective state of the agent and the Action Selection maps this affective state into concrete animations in the agent's embodiment. Additionally, when the user is experiencing a negative affective state with a probability above a certain threshold, one of the available supportive behaviours is also displayed by the agent. However, all the available supportive behaviours are appropriate for the agent to display when the user is experiencing a negative affective state. Therefore, a relevant question arises: which behaviours are more effective for a particular user? As discussed earlier in chapter 2, different behaviours may be more effective for different users depending, for example, on each user's personality, gender, and so on.

Taking this into account, we propose an approach based on Reinforcement Learning (RL), so that the agent can learn by trial and error which are the best strategies for a particular user, and adapt its empathic behaviour accordingly over time. Similar RL approaches were applied successfully, for example, to induce pedagogical strategies in an Intelligent Tutoring System without requiring a large training corpus [32]. But how can we apply RL techniques into our model? In other words, how to measure the success of each supportive behaviour? And, more importantly, how to select the most effective empathic strategy at a given moment?

One key issue in personalizing user experiences is deciding between trying strategies that have yet to be used with a particular user (exploration) and using the seemingly best strategy up to that moment (exploitation). To address this issue, we adopted a well-known *no-regret learning* algorithm that was designed to optimally balance exploration and exploitation. Algorithmically, the advantage of no-regret learning algorithms over other learning approaches is the focus of the former in maximizing performance *during* learning. In terms of our application scenario, this can be translated in a more effective adaptation as perceived by the user. This is a classical problem in Machine Learning research commonly designated as Multiarmed Bandit Problem [6]. In this problem, a gambler chooses a slot machine to play (a "one-armed bandit"), which yield rewards $X_{i,n}$ for $1 \leq i \leq K$, where i is the index of a machine. Consecutive plays of machine machine i return rewards $X_{i,1}, X_{i,2}, ...$ The goal of the gambler is to maximise the sum of rewards he receives from a sequence of pulls.

Formalising this problem into our model, the agent has a set of supportive behaviours $1 \leq i \leq S$, where each i is the index of a Supportive Behaviour and S is the number of available behaviours. When selecting the behaviour S_i at the time step t, this action yields a reward $x_i(t)$. Since the goal of the agent is to increase the user's estimation of being in a positive affective state, the reward function should reflect how the user's affective state changed after the agent has employed the supportive behaviour:

$$x_i(t) = P(PositiveFeeling)_t - P(PositiveFeeling)_{t-1} \tag{5.1}$$

$P(PositiveFeeling)_t$ is the new valence estimate and $P(PositiveFeeling)_{t-1}$ represents the estimate of the user's positive feeling before the strategy i is displayed to the user. Figure 5.3 contains a temporal diagram that shows how the reward values are updated. When a new game event is received, the Affect Recognition System infers the user's affective state (step 1) and informs the Action selection of $P(PositiveFeeling)_{t-1}$ (step 2). Considering the history of previous interactions with that user stored in the User Model (step 3), the agent selects a supportive behaviour (step 4) and displays it to the user, along with the facial expression resultant from the Empathic Appraisal (step 5). A few moments after, and to capture the user's response to that supportive behaviour, the Affect Recognition system infers the user's affective state again (step 6) and informs the Action Selection about $P(PositiveFeeling)_t$ (step 7). The Action Selection then makes the necessary updates in the User Model (step 8).

In order to support adaptation to the user, the Action Selection component needs to integrate a *policy*, that is, an algorithm that allows the agent to select the next supportive behaviour considering the previous history of employed behaviours and consequent rewards. In the literature, there are several optimal policies that solve the Multiarmed Bandit problem (e.g., [6, 7]). The solutions vary mostly on the statistical assumptions that can be made about the application domain. In the following chapters, we show how two different policies for selecting the strategies were implemented in our scenario.

Since the obtained rewards are being constantly updated in the User Model, only in the first interaction the agent starts “from scratch”. In the subsequent games, the agent has already some knowledge about the user's preferences on the different supportive behaviours. This knowledge is refined and updated over time.

5.3.4 Recalling Past Interactions

As discussed in the Related Work chapter, the ability to memorise and recall past events is an important feature for social robots who interact with users for extended periods of time. Besides giving the user the impression of behavioural coherence, we believe that this capacity will contribute to the establishment of rapport between the user and the agent. Consequently, this may cause in the user empathic feelings towards the agent. As such, this component is part of our empathic model for long-term interaction.

All the information in the User Model is saved persistently for each user over the interactions. In the beginning of the game (except in the first one), the User Model is loaded and the agent uses that information to summarise the previous interactions with that user. For example, the agent can refer to the number of days since the last interaction and to the result of their previous game:

> "Hi <*user*>, it has been <*N*> days since we played together. In our last game I won. Let's see if I can keep up this time!"

Similarly, the information in the User Model can also be used at the end game to summarise the user's evolution in the game:

> "<*user*>, congratulations, it was a good game! You're doing really well. In the <*total number of interactions*> games we played together, you beat me <*number of victories*> times. Bye <*user*>, see you in the next game!"

5.4 Applying the Empathic Model to the iCat

In this section, we explain how the model proposed in the previous sections was implemented in the scenario introduced in section 4.1. We start by presenting the overall architecture of the scenario. After that, some implementation details on the iCat's Empathic Appraisal and Action Selection components are given.

An overview of the architecture of the scenario is displayed in Figure 5.4. As explained in the model, to convey empathy the agent, in this case the iCat robot, needs

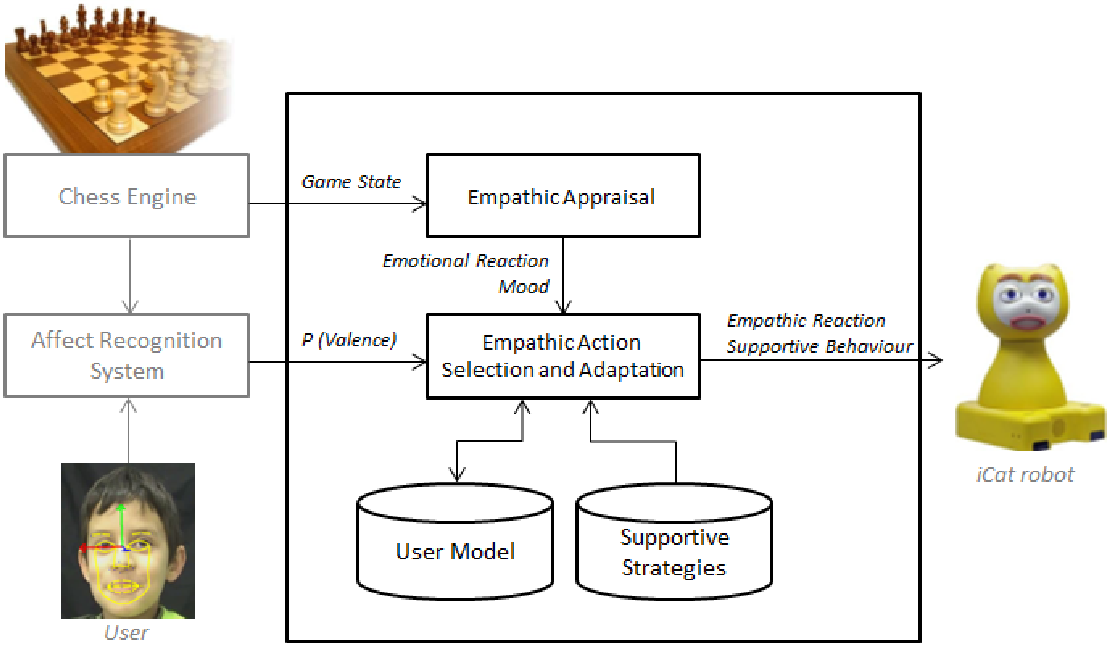


Figure 5.4: Instantiation of the Empathic Model in the iCat scenario.

to interpret the game environment and infer the affective state of the user.

In this scenario, the interpretation of the game is done by a Chess Engine previously described in section 4.1. The other input is the Affect Recognition system [27], which was developed in the context of a research project[2]. It consists of a valence detector based on Support Vector Machines (SVM) that combines a set of visual and contextual features and returns the estimated values for valence (i.e., whether the user is more likely to be experiencing a positive, neutral or negative feeling). The visual features are obtained by a face tracking software[3], using a standard Logitech webcam placed in front of the users (near the iCat)[4]. The contextual features are retrieved from the Chess Engine. The system was trained with data from Inter-ACT [27], a corpus collected during previous studies of children playing with the iCat robot. The corpus contains videos that capture children's interaction with the iCat from different perspectives, and includes

[2]http://lirec.eu/
[3]http://www.seeingmachines.com/product/faceapi/
[4]The iCat also has a built-in webcam, but since the robot is always moving due to the idle behaviours, it is impossible to perform face tracking with that camera.

synchronised contextual information about the game and the behaviour displayed by the robot. The final set of features used by the SVM are described in Table 5.2.

When the user plays a new move, the Chess Engine interprets that move and sends a value that represents the state of the game to the Empathic Appraisal component. This triggers the whole empathic process, which will be described below.

Table 5.2: Visual and contextual features for the discrimination of user's valence.

	Visual features	**Contextual features**
Positive feeling	Looking at the iCat, smiling	User winning, user improving in the game, user capturing a piece
Negative feeling	Looking at the chessboard, looking elsewhere	User losing, user getting worse in the game, iCat capturing a piece

5.4.1 Empathic Appraisal Implementation

Using the information received from the Chess Engine, the iCat interprets the new state of the game in the perspective of the user and updates its affective state. This empathic affective state consists of two main parts: emotional reactions and mood. In this section we describe how the affective state is calculated in this scenario.

Emotional Reactions. The emotional reactions of the robot are computed using the *emotivector* system, which generates affective sensations based on the mismatch between a sensed and an expected value. In this case, the sensed value is the new game state value received from the game engine. To calculate the expected value, a predictor based on the history of the previous game state values is required. We have used the moving averages algorithm [37]. Moving averages' underlying purpose is to smooth a data series and make it easier to spot trends. The two most popular types of moving averages are the simple moving average (SMA) and the exponential moving average (EMA). We have applied the exponential moving averages, also called weighted moving averages, which reduces the lag by applying more weight to recent values relative to older values. Using EMA, the expected value $\hat{s}(n)$ can be calculated as follows:

$$\hat{s}(n) = (1 - \alpha).\hat{s}(n-1) + \alpha.s(n) \tag{5.2}$$

$$\alpha = \frac{2}{1+N} \tag{5.3}$$

where $\hat{s}(n-1)$ is the expected value computed in the previous move and $s(n)$ is the current sensed value. α is a smoothing factor and N is usually the specified number of periods. We used $N = 2$ and thus our smoothing factor was $\alpha = 0.67$.

When using this equation to estimate the first expected value, there is no available value to use as $\hat{s}(n-1)$. Therefore, the first expected value is calculated using a simple moving average. A simple moving average is an arithmetic mean over a specified number of periods.

After calculating the expected value employing equation 5.2, and using the game state as the sensed value, the emotivector system computes the proper empathic sensation of the iCat. The threshold value that decides whether the sensation is expected or unexpected is dynamic. It is estimated using an EMA based on the history of mismatches between expected and sensed values. This means that, the larger the mismatch, the larger the threshold value tends to be in the future. On the other hand, if the mismatch between the expected and sensed values are low, the threshold tends to decrease over time.

Mood. The game state value also serves as the input for the robot's mood. After being normalised to the [-1;1] interval, Mood is determined by the following filter function:

$$M(x) = \begin{cases} min(1, \log(2x+1)), & \text{if } x \geq 0 \\ max(-1, -\log(-2x+1)), & \text{otherwise} \end{cases} \tag{5.4}$$

A graphical representation of this function is shown in Figure 5.6. When the output of the function is larger than 1 or smaller than -1, mood takes the value of 1 or -1, respectively. This function was selected due to the particularity of the chess heuristic function of the Chess Engine, which only returns values closer to -1 and 1 when the

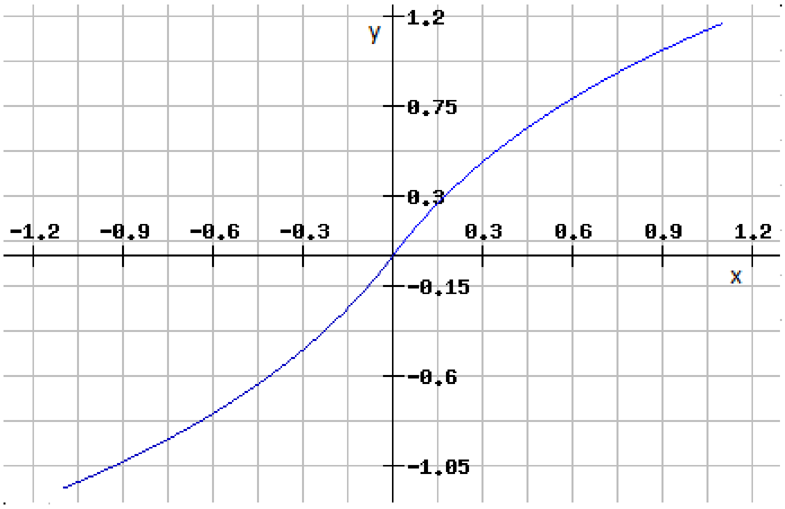


Figure 5.5: Mood filter function.

game is about to end. This way, we can have higher intensity moods not only in the end game, but during most of the game.

Implemented Supportive Strategies

The database of supportive strategies of the iCat include facial expressions, verbal utterances and game-related actions, divided in two main groups: emotional expressions and supportive behaviours.

Emotional Expressions. The emotional expressions display the iCat's emotions and mood resulting from the Empathic Appraisal component, and reflect the state of the game in the perspective of the user.

The emotional reactions generated by the emotivector system were mapped into animations existing in the library of animations that comes with the iCat software, since these were previously submitted to validation user studies (see section 4.1.3). We made a correspondence between the animations and each one of the nine emotivector sensations, as displayed in Table 5.3. The choices for the sensation-animation mapping were based on the meaning of the sensations. For example, "stronger reward" means that the user experienced a reward much better than she/he was expecting, and therefore the correspondent emotion is "excitement". However, the existing emotional animations

Table 5.3: Mapping between the *emotivector* sensations and animations in the iCat's library. The animations marked with * were slightly modified to recreate the desired state.

Sensation	**Meaning**	**Animation**
stronger Reward	Better than expected	excited
expected Reward	As good as expected	confirm*
weaker Reward	Not as good as expected	happy
unexpected Reward	Good, unexpected	arrogant
negligible	Not worth considering	think*
unexpected Punishment	Bad, unexpected	shocked
weaker Punishment	Not as bad as expected	apologize*
expected Punishment	Bad, as expected	angry
stronger Punishment	Worse than expected	scared

were not enough to map all the sensations. Therefore, we adapted other animations to some sensations. For instance, "expected reward" means that the user experienced a sensation as good as she/she was expecting. Thus, the emotional reaction corresponds to a gentle smile with a nod (adapted from the existing "confirm" animation).

Unlike the emotional reactions, that only occur after the user plays a move, mood is implemented as a behaviour that is constantly running. Two sets of parametrizations of the iCat's embodiment variables, one for the positive and another one for the negative mood, were defined. These parametrizations influence the iCat's eyebrows, eyelids and lips (see Figure 5.6). The values for the parametrizations were obtained from the happy and sad iCat animations. After the Empathic Appraisal process, the robot's body parts are changed as a linear interpolation between the values in one of these two parametrizations (depending if the mood value is positive or negative) and a "neutral" face. Using this method, we can have different intensity moods. For example, if the mood variable is 1, the resultant facial expression is the one displayed in Figure 5.6. However, if the value is 0.5, the facial expression of the iCat is still positive, but not as intense as the one in Figure 5.6.

Supportive Behaviours. Taking inspiration on the literature of empathy and prosocial behaviour (e.g., the work of Cooper et al. [34]), on the previous user studies, especially the one reported in section 4.3, and considering the particular context of the

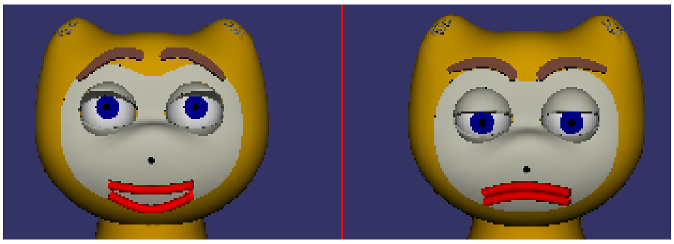


Figure 5.6: Two different body parametrizations for the iCat's mood (positive mood on the left and negative on the right).

chess game, the initial set of supportive behaviours implemented in the robot was the following:

- **Encouraging:** includes encouraging utterances, for example, "come on, I believe you can still recover the disadvantage!". The selected utterances vary depending on the current affective state state of the robot computed in the Empathic Appraisal module.

- **Scaffolding:** comments the user's move, while providing a brief explanation. For example, if the move the user just played was not very good, the iCat says "that was not a very good move, because now I can play my queen to the a5 square". If the user plays a good move, one possible comment would be "well done, you played what I would have played".

- **Suggest Move:** the iCat suggests a good move for the user to play, but only if the user accepts that suggestion. In the user's turn, the iCat offers help by saying something such as "Need help? Touch one of my paws so I can suggest you a move". If the user touches one of the iCat's paws, the iCat tells the user the best move for the user's according to the Chess Engine. This is the only behaviour that is not displayed right after the user's move, but rather in the user's turn, a few seconds after the user plays the iCat's move.

- **Play Bad Move:** the robot plays one of the worst moves of the list of valid moves, rather than the optimal one (as usual). Often this move allows the user to capture one of the iCat's pieces, and therefore the user gets some advantage in the game.

5.4.2 Adaptive Action Selection Implementation

After the Empathic Appraisal process, the Affect Recognition system informs the iCat about the new affective state of the user. If the user is experiencing a negative feeling with a value higher than 0.5 (this value was selected after some fine-tuning and testing with a small group of users), one of the supportive behaviours described above is selected and the robot displays it to the user (after showing the empathic emotional expression). In this section, we present the policy for selecting the most appropriate supportive behaviour type for a particular user.

Consider a vector $sb[n]$ that contains the $n = 1, ..., K$ possible supportive behaviour types of the robot (in this particular case, $K = 4$). The policy for selecting the most suitable behaviour S_i is computed according to Algorithm 5.1.

Algorithm 5.1 UCB1 Algorithm [6] for selecting the supportive behaviours.

```
if NOT all behaviours selected once then
   for i = 1 → K do
      S_i = sb[i]
   end for
else
   loop
      Select strategy S_i that maximises: x̄_i + sqrt(2 ln n / n_i)
      where:
      x̄_i → average reward obtained from strategy i
      n_i → number of times strategy i was selected so far
      n → overall number of strategies selected so far
   end loop
end if
```

UCB1 is a deterministic policy proposed by Auer and colleagues [6] which assumes that the rewards obtained from selecting the behaviours are within the [0;1] interval. Since our rewards can vary between -1 and +1 because they are calculated as the difference between two probabilities, a normalisation was applied to the reward values.

When the robot employs one of the supportive behaviours, the average reward $\overline{x_i}$ for that behaviour, stored in the User Model, is updated six seconds after it is displayed to the user. This is done after six seconds because it was the duration of the video segments used to train the smile features in the valence detector. This way, the robot can adapt its empathic behaviour by trial and error considering the previous reactions of a particular user to an empathic strategy.

5.5 Concluding Remarks

In this chapter, we presented an empathic model tailored for agents (robots or virtual characters) that interact with users during extended periods of time. This model, inspired by literature on social sciences and empirical studies, relies on the assumption that empathy can go beyond mere emotional responses to the affective state of another. Instead, we also consider that, when the user's affective state is negative, the agent should select a supportive action in the attempt to reverse that state. Since our goal is to apply the model to long-term interactions, the agent should understand which strategies have a higher impact on changing the user's affective state and adapt its behaviour accordingly. Additionally, because of the long-term goal, the agent should also recall previous events in order to establish a closer relationship with the user.

The model has two main inputs: a Game Engine and an Affect Recognition System. When a new game event is received, the Game Engine interprets the new state of the game, while the Affect Recognition System infers the user's affective state. After that, based on the state of the game in the perspective of the user, the agent generates an empathic affective state (Empathic Appraisal component). If the user's affective state is negative with high probability, in addition to expressing the empathic affective state, the Action Selection mechanism of the agent selects a Supportive Strategy based on the User Model, which contains information about the user's preferences. To build rapport with the user, at the beginning and at the end of each interaction, the agent recalls previous game events stored in the User Model.

The proposed model intends to be generic and applicable to a diversity of turn-based scenarios, with special focus on turn-based games. As such, several design decisions were left open, so that they could better suit the needs of the target application and the agent's embodiment. For example, we do not specify how the agent should express its affective state (e.g., facial expressions, verbal utterances), as this depends heavily on the agent's embodiment. For being scenario-dependent, the supportive strategies can also vary and thus are only regarded as a set of generic strategies that can change the user's affect from negative to positive.

After presenting the model, we described how it was implemented in the iCat scenario. In particular, we described how the affective sensations are predicted and generated, and how the robot's mood is calculated. The robot's affective state is displayed to the user through emotional expressions, while some facial features of the robot are constantly being updated because of the mood. Additionally, if the user's affective state is negative with a high probability (in this case, above 0.5), the robot also exhibits a supportive behaviour (encouraging, scaffolding, suggest move and play bad move). When a supportive behaviour is displayed, a few seconds after the iCat evaluates the effect that this behaviour had on the user's affective state. The difference between the previous and current affective states of the user work as a reward function for the policy that selects the most suitable supportive behaviour for a particular user. This scenario was evaluated in a preliminary short-term study reported in the following chapter. After the analysis of the results of this study, some refinements were made in terms of implementation, before conducting the final long-term study described in chapter 7.

Chapter 6

Short-Term Evaluation

To evaluate the solution presented in section 5.4, which describes an initial version of the proposed architecture in the iCat robot, we performed an exploratory study. This study had two main objectives. First, to evaluate how users react to the different empathic behaviours displayed by the robot, in a scenario where the robot takes an active role as a player (in contrast to having a passive role as in the study reported in section 4.3). The second main goal of this study was to analyse the viability and the impact of the adaptive supportive behaviours in real time, while collecting data that could be used to further improve the model.

6.1 Method

In this section, we start by describing the three different study conditions and the procedure. Finally, we present the measures employed in this study.

6.1.1 Conditions

We evaluated three different versions of the robot's behaviour:

- **Neutral**: the iCat's affective behaviour is competitive, i.e., it reflects the state of the game in its own point of view rather than from the user's perspective. The robot also provides feedback on the user's moves, but as in the *neutral* version of the study presented in section 4.3 (e.g., "good move", "you didn't play so well this time"). In this condition, the iCat does not recognise the user's affective state and consequently does not exhibit any empathic behaviour.

- **Empathic**: in order to show empathy for its opponent, the iCat's affective behaviour takes into account the state of the game in the perspective of the user. Besides displaying empathic emotions, when the robot detects that the user's estimation of negative feeling is below a certain threshold, it employs a supportive strategy. In this case, the robot selects one out of the possible set of strategies *randomly*. In a similar way as the study presented in section 4.3, the iCat employs the user's name two times more than it does in the *neutral* condition.

- **Adaptive Empathic**: this version is similar to the empathic version, except in the way of selecting the supportive behaviours, which are selected according to the learning algorithm described in section 5.4.2.

6.1.2 Procedure

This scenario was evaluated with children from an elementary school where students have two hours per week of chess lessons. A total of 84 subjects, with ages between 8 and 10 years old, participated in the study. From these, 40 children (19 male and 21 female) played an individual chess exercise with the iCat, and 44 (22 male and 21 female) interacted with the robot in groups of 4 or 5 students.

In the **individual interactions** (see Figure 6.1a), participants were alone in a room with the experimenter and played a chess exercise with the iCat for approximately 10 minutes[1]. After that time, depending on the state of the game, the iCat either gave up (if it was in disadvantage) or proposed a draw (if the child was loosing or if none of the players had advantage), arguing that it had to play with another user. Participants were

[1]The exercise was suggested by the chess instructor, so that the difficulty was reasonably appropriate for all the students.

then oriented to another part of the room where they filled in a questionnaire. Finally, they were guided to another room and were interviewed by another experimenter.

Participants who played individually with the iCat were randomised over the three study conditions. Since we followed a between groups design, our final sample consisted of 14 children in the *adaptive_empathic condition*, 13 in the *empathic* and 13 in the *neutral* condition. As the number of subjects in each condition is not equal, when appropriate, we present the results using percentages rather than frequencies.

The **group interactions** differed from individual interactions in several aspects. First, instead of playing alone, participants collaboratively played the chess exercise against the iCat, with each child playing a few moves (see Figure 6.1b). These interactions took approximately 15 minutes. In this case, only the *empathic* version of the robot was used. Users need to be sat down, face-to-face with the robot, so that the visual features of the user's expression are properly tracked. Since in group interactions participants were standing up in front of the robot, in a way that all of them could see the iCat's behaviour, the Affect Detection system only took into account the contextual features of the game. In this situation, the reward values of the *adaptive empathic* are usually zero, considering that in 6 seconds the state of the game is unlikely to change. For this reason, we opted for having only the *empathic* condition in the group interaction and use these results to capture children's reactions to the several empathic strategies displayed by the robot. Since the questionnaire was tailored for individual interactions, after playing with the robot, these participants were only interviewed (also in groups).

Despite the limitations in terms of affect recognition, the group interactions were extremely useful for improving the iCat's behaviour. By having other colleagues around, children become much more extrovert and often express their opinions aloud, discussing them with the other students. The post observation of the group interactions allowed us especially to refine some concrete behaviours of the iCat. For example, some of the empathic utterances were not perceived as so by the children, or they were too complex for children of that age to understand. We noticed this, for instance, when children asked their colleagues what the robot tried to say.

(a) Individual interactions

(b) Group Interactions

Figure 6.1: Children playing with the iCat during the study.

6.1.3 Measures

Through the questionnaire, we measured user's **engagement with the robot**, **perceived affective and behavioural interdependence**, **help** and **self-validation**. For each measure, users expressed their opinion on a set of assertions using a 5-point Likert scale, where 1 meant "totally disagree" and 5 meant "totally agree" (see Table 6.1).

Engagement is a metric that has been extensively used both in human-robot and human-agent interaction and has been defined from several perspectives. For example, Sidner et al. [135] defined engagement as "the process by which two (or more) participants establish, maintain and end their perceived connection". More recently, Bickmore and colleagues [17] proposed a definition of long-term engagement: "the degree of involvement a user chooses to have with a system over time". Even though the later definition is more appropriate for the long-term goals of this work, in this study, users only interacted with the robot for a short period of time, and thus we will adopt the first definition, which has more to do with short-term engagement. The questionnaire items regarding engagement in our study (see Table 6.1) are based on the questions used by Sidner et al. [135] to evaluate users' responses towards a robot capable of using several social capabilities to attract the attention of users.

Perceived affective and behavioural interdependence are dimensions of the social presence questionnaire previously used in the long-term study described in section

Table 6.1: Questionnaire items used in this study for each measure, translated to English.

Measure	Questionnaire Items
Perceived Affective and Behavioural Interdependence	Q1. I knew how iCat was feeling. Q2. iCat knew how I was feeling. Q3. iCat's mood influenced my own mood. Q4. I influenced iCat's mood. Q5. iCat's behaviour influenced my own behaviour. Q6. I influenced iCat's behaviour.
Engagement	Q7. iCat made me participate more in the game. Q8. It was fun playing with iCat. Q9. Playing with the iCat caused me real feelings and emotions. Q10. I lost track of time while playing with iCat.
Help	Q11. iCat helped me during the game. Q12. iCat's comments were useful to me. Q13. iCat's comments were helpful when I needed them.
Self-Validation	Q14. I felt that I could play better in the presence of iCat. Q15. iCat praised me when I played well.

4.2. In fact, these two dimensions were part of the ones who decreased the most over the interactions. With the integration of the empathic behaviours in the robot, different results were expected this time.

We also decided to measure two of the friendship dimensions that were used in the study of section 4.3, **help** and **self-validation**. We decided not to evaluate all the friendship functions since we wanted to keep the overall number of questions in the questionnaire reasonable to 8-10 years old children. For this reason, we selected only help and self-validation, because we believe that these dimensions may be affected by the empathic strategies employed by the robot in this scenario.

To further analyse how users perceived the interaction with the robot we also performed **open-ended interviews**. The main objectives of the interviews were to capture how children perceived the affective behaviour of the iCat (e.g., check if they mention the empathic capabilities of the robot) and to find possible ways of improving the scenario. The interviews were semi-structured, containing initial yes-or-no questions followed by open-ended questions that allowed children to justify and elaborate their answers. The structure of the interview is depicted in Table 6.2. Each interview lasted on average 10 minutes. The interviews were audio recorded and later on transcribed. The interview transcriptions were used as primary sources of data for content analysis.

Table 6.2: Structure of the interview.

Question	**Response Type**
1. Do you believe that iCat recognised how you felt during the game?	yes-no
1.1. Why do you think so?	open
2. Did you experience any particular emotions during the game?	yes-no
2.1. Which emotions?	open
2.2. Do you think the iCat felt similar emotions?	yes-no
2.3. Why?	open
3. Do you think the chess game is more fun because of the iCat (than playing with a computer)?	yes-no
3.1. Why?	open
4. What are the main differences of playing against the iCat compared to playing against your colleagues?	open
5. How could we further improve the iCat?	open
6. Identify advantages and disadvantages of playing with the iCat.	open

6.2 Results

6.2.1 Questionnaire

While answering the questionnaire, most participants asked the experimenters what was the meaning of the word “influence”, used in questions 3, 4, 5 and 6. All these questions relate to the perceived affective and behavioural interdependence dimensions of social presence. Thus, we decided to exclude these measures from the quantitative analysis and study them only taking into account the open-ended interviews. For this reason, in this section we only present the questionnaire results for engagement, help and self-validation. For each measure, we considered the average ratings of all the questionnaire items associated to it. We verified that the distribution of the data was not normal by an initial run of Kolmogorov–Smirnov test, so non-parametric tests were applied.

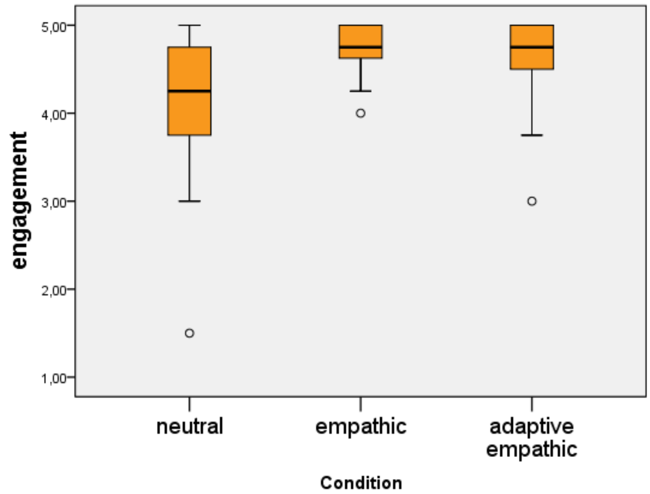
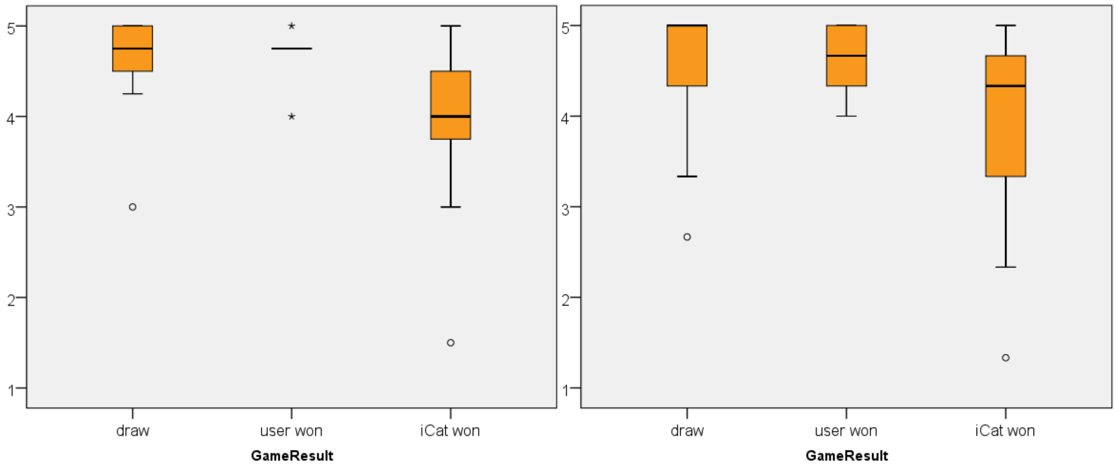


Figure 6.2: Boxplot of the Engagement ratings in the three conditions.

Engagement

Even though the boxplot chart of Figure 6.2 indicates that participants in the *empathic* and *adaptive empathic* conditions found the interaction more engaging, these results were not significantly verified by a Kruskal-Wallis test with the three groups. Thus, to follow up this finding, two pairs of Mann–Whitney tests were ran, respectively *neutral* vs. *empathic* and *neutral* vs. *adaptive empathic*. The results of these two tests indicate that participants in the *empathic* group found the interaction significantly more engaging than participants in the *neutral* group (U=46, r=-0.4, p<0.05), and participants in the *adaptive empathic* condition also provided significantly higher ratings in terms of engagement than the control group (U=62, r=-0.28, p<0.05).

We also tested whether the end-game result had any effects on children's engagement. Interestingly, considering the subjects from all conditions (N=40), the end game result affected how engaging participants found the robot ($H(2)$=8.17, p<0.05), with participants who lost the game (N=14) rating the interaction as less engaging than participants who won (N=6) or ended the game in a drawing position (N=20). These results are illustrated in Figure 6.3a.



(a) Engagement (b) Help

Figure 6.3: Boxplots of the game result.

Help

The Kruskal-Wallis test indicates that empathy affected how helpful participants found the robot ($H(2)$=10.53, p<0.05), as shown in Figure 6.4. We also run Mann–Whitney tests to compare each group against the control condition. The results were again significant, showing that participants in the *empathic* (U=33, r=-0.53, p<0.05) and *adaptive empathic* (U=34, r=-0.56, p<0.05) conditions found the robot more helpful than participants the *neutral* group. These results differ from the ones presented in the study of section 4.3 regarding help, where this dimension was perceived equally by users from the neutral and empathic conditions. It is possible that this difference might have been caused by the prosocial actions implemented in the robot. While in the previous study the iCat acted as a mere spectator commenting the quality of the players' moves, in this case the robot pro-actively assisted the user, for example by suggesting a good move or by playing a deliberately a bad move.

As expected, the end game result also affected this variable. As it happened with engagement, participants who lost the game against the iCat also found the robot less helpful ($H(2)$=8.65, p<0.05). These results can be strengthened by Figure 6.3b.

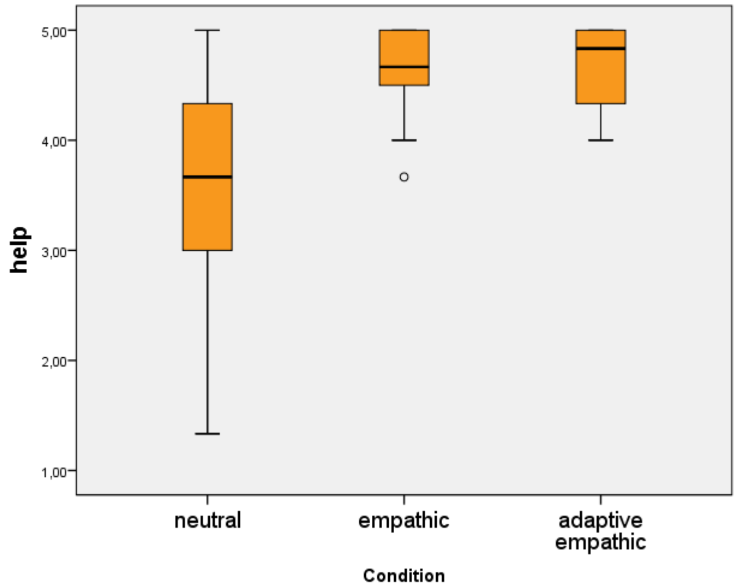


Figure 6.4: Boxplot of the Help ratings in the three conditions.

Self-Validation

Similarly to the engagement results, the Kruskal-Wallis test was inconclusive for self-validation, but Mann–Whitney tests between each empathic condition and the control group yielded significant results. Participants in the *empathic* group gave significantly higher ratings than participants in the *neutral* group (U=45, r=-0.37, p<0.05), and the same happened with the *adaptive empathic condition* (U=60.5, r=-0.3, p<0.05). The boxplot chart of Figure 6.5 also illustrates these findings, which are also in line with the previous results on self-validation (see section 4.3).

The end game result did not have any significant effect on this dimension. As opposed to what happened with engagement and help, participants found that the robot encouraged them during the game, despite the final result. Since the empathic behaviours of the robot are more often employed when the user is loosing the game (because it is one of the indicators of the user's affective state being negative), this suggests that the iCat's empathic behaviour had a positive effect on users.

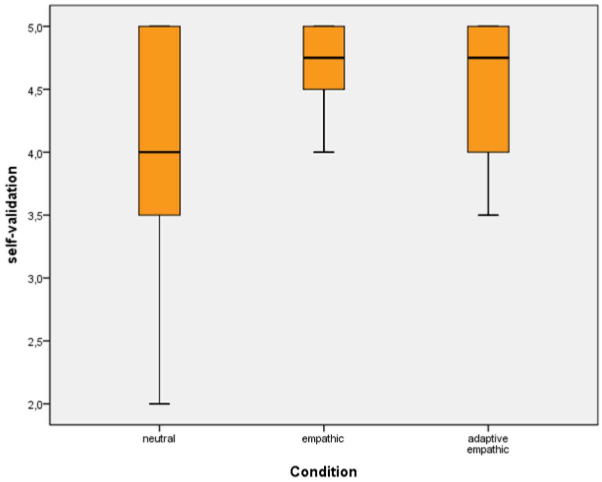


Figure 6.5: Boxplot of the Self-Validation ratings in the three conditions.

6.2.2 Open-Ended Interviews

The interview transcriptions were coded according to the methodology proposed by Chi [33], using an iterative coding process. An initial coding scheme was obtained while reading and highlighting the main concepts in the text. Subsequent iterations allowed us to refine the initial coding scheme. After that, related codes were grouped into themes and categories. To ensure reliability during the coding process, two complete coding iterations were performed with an interval of one month by the same coder. Differences between these two coding steps were addressed by asking another coder to rate that segment. Finally, using the developed coding scheme, a comparative analysis between the three different control conditions was performed to understand how the robot's empathic behaviour affected children's perception of the interaction.

The results presented in this section are divided by the main themes in the open-ended questions of the interview (see Table 6.2). In the first two themes (questions 1 and 2), our goal was to understand whether children perceived the empathic behaviours expressed by the robot and how such behaviours had an impact on children's own emotions. After that, we looked for the main differences between playing with the iCat and playing with a computer or a colleague (questions 3 and 4). Finally, children were asked about the advantages, disadvantages and ways of improving the robot, with the aim of gaining a better understanding on how, in the view of children, would be the ideal interaction with a social robot (questions 5 and 6).

Perceived Empathy

In the first question of the interview, our goal was to investigate whether children believed that the robot "perceived" how they were feeling during the game, and see if they refer to the implemented empathic behaviours to justify their answer. In other words, we wanted to see if they considered that the robot was being empathic towards them. As depicted in Figure 6.6, when asked whether they believed that the iCat knew how they were feeling, almost all the children from the *adaptive empathic* condition and nearly 77% in the *empathic* answered "yes", while in the *neutral* condition only about 50% of the children provided an affirmative answer. The iCat's comments and helpful behaviour seemed to play an important role for children who answered "yes", especially in the two *empathic* conditions, as these two participants stated:

> P18: *"I think so, because when I played bad moves the iCat warned me about that. He said what I was thinking."*
>
> P21: *"Yes, because for example when I played a bad move, he [the iCat] told me that I should do better [...] and when I didn't know what to play sometimes he told me that he could help"*

While in the *empathic* conditions most participants used the robot's behaviour and facial expressions to justify their affirmative answers, most of the participants from the *neutral* condition considered that the robot knew how they were feeling due to the cameras and because the robot was placed in front of them. For example:

> P34: *"Yes, because he saw my face and managed to memorise it"*

This effect might have happened due to the initial calibration phase of the face tracking software, in which children were asked to move their heads to different sides while looking at the robot. This procedure was done in all the conditions, even when the affect detection system was not being used, i.e., for children in the *neutral* condition.

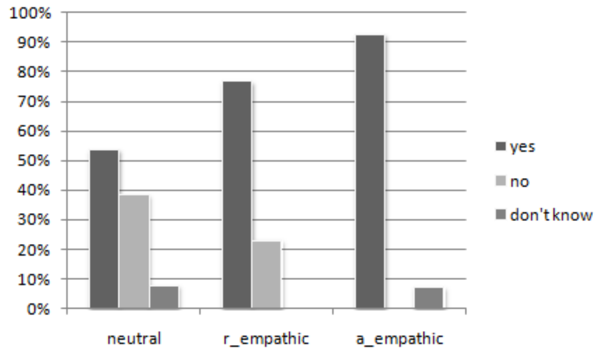


Figure 6.6: Children's answers to the question "Do you believe that the iCat recognised how you felt during the game?" in the three groups of users.

As for the participants who answered "no", most of them were not able to justify their answer. In the *neutral* group, some participants argued that the iCat was too focused on its own game to have time to understand their emotions. Being a machine or a robot was another frequent reason given by participants who answered negatively:

> P25: *"No, because it's a robot and I don't think he feels the same things that people feel..."*

These results suggest that the implemented empathic behaviours influenced how children perceived the robot, as they frequently mentioned these behaviours to justify their affirmative answer. On the other hand, children who answered "no" often referred to the mechanical features of the robot. This leads us to believe that without the empathic behaviours, children perceive the robot more as a machine, and are less willing to suspend their disbelief during the interaction.

Experienced Emotions

The second group of questions was about children's experienced emotions while playing the game, and whether they believed that the iCat felt similar emotions. As in the previous group, we were looking for signs of perceived empathy: if children considered that the robot felt similar emotions, this would be an indicator of empathy. Most of the children in the three conditions answered affirmatively to the question "did you experience

Table 6.3: Frequency of the reported affective and motivational states in the three conditions.

Category	**neutral**	**empathic**	**a_empathic**
happiness	3	1	7
nervous/anxious	1	3	0
enjoyed playing/fun	6	3	2
can't explain	1	2	2

any particular emotions during the game?" (respectively, 77% in the *neutral* condition, 85% in the *empathic* and 79% in the *adaptive empathic* condition). When asked about particular emotions, we obtained a variety of affective and other motivational states, which in the final coding phase were grouped into the categories presented in Table 6.3. Children who interacted with the *adaptive empathic* version of the robot reported more frequently "happiness" as one of the experienced emotions:

> P33: *"I was happy! I have waited so long to play with the iCat but it was worth it."*

Other participants reported that they felt anxious or nervous during the game. From the observations of the interaction, we noticed that the experimental apparatus (children being alone in the room with the experimenters, the presence of the cameras, etc.) may also have influenced children to feel this way. For example, one participant from the *empathic* condition described his affective state as follows:

> P38: *"I felt weird... I was playing with someone I've never played before [...] I was a bit nervous..."*

In addition to the reported affective states, a significant part of the children also described other positive states and reactions when asked about emotions, which can be explained by the fact that children's understanding of emotions is still under development at this age. As argued by Denham and colleagues [43], some children are more skilled than others in experiencing and using emotional vocabulary due to several factors such as gender, parent's emotional displays of affection and children's own past experience. Therefore, they used terms such as "enjoyed playing" or "I felt that I was having

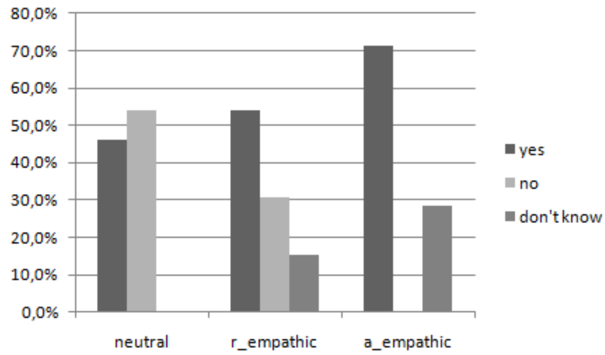


Figure 6.7: Children's answers to the question "Do you believe that the iCat felt emotions similar to the ones you felt?" in the three groups of users.

fun" to describe their emotional experience. Others even highlighted the fact that they were playing with a new robotic product, which is still not available to everyone. When asked about the emotions experienced during the game, one participant from the *neutral* condition answered:

> P34: *"I felt a big honour for playing with a robot that is not for sale yet!"*

As illustrated in Figure 6.7, more participants from the two *empathic* conditions considered that the iCat felt similar emotions to the ones they have reported, whereas in the *neutral* group this effect was not so substantial. Moreover, the affirmative answers in the *neutral* condition seem to be biased by an effect that we were unable to control: as children were very curious and expectant about the presence of the robot, the ones who played first spoke about their experience to the others. We noticed this because some of the children, at the end of the game, asked us questions such as "one of my colleagues told me that the iCat suggested him a good move... why didn't he do that with me?". Also, some children who played with the *neutral* version of the robot reported behaviours that do not happen in that condition (for example, the robot offering help). One of the participants even said that directly:

> P22: *"Yes, because I was told that he can do that."*

In sum, nearly all the children experienced emotions while playing chess with the iCat. However, the type of reported emotions changed slightly among conditions: while happiness appears in almost all the descriptions of children in the *adaptive empathic* condition, children in the *neutral* and *empathic* conditions reported less positive emotions. An interesting finding was that negative affective states (e.g., anxious, nervous, etc.) were mentioned more often in the *empathic* condition, which may indicate that the introduction of empathic behaviours should be done carefully under the risk of confusing the users.

Comparing Robots with Computers

To gain a better understanding of how children perceive social robots and how the presence of the robot might influence their opinion of the chess game, we asked them what did they consider more fun, playing against the iCat or playing against a computer. All the participants except one (P18) agreed that playing the game against the iCat was more fun. The only participant who did not answered this way gave the following answer:

> P18: *"For me it's the same, playing chess is always fun. But I had a lot of fun playing against the iCat".*

As for the reasons why playing with the iCat was more fun than playing with a computer, there were some differences in the answers provided by participants from the different conditions. In both *empathic* conditions, children often referred to the helping behaviours of the robot as the main difference between playing with the robot when compared to playing with the computer. Also, some participants mentioned that playing with the iCat is more similar to playing with another person:

> P10: *"because the iCat is like a person but it's a robot."*
>
> P16: *"because he [the iCat] can say if I played well or bad, what he thinks about our moves... and also because we can talk to him."*
>
> P24: *"[the iCat] has emotions, and computers don't. He speaks and actually is a good friend and player."*

There were some aspects that were evenly used by participants in the three conditions. For example, the fact that they never played with a robot before, they need to play the iCat's moves while the computer plays automatically, the robot speaks, and also the robot's emotional and facial expressions.

Comparing Robots with Humans

After asking about the differences between robots and computers, we asked children what were the main differences between playing with the iCat or playing with their colleagues. There were differences between the answers provided by subjects from the *adaptive empathic* condition and the two other conditions. Most children from the *adaptive empathic* condition indicated as the main difference the fact that the iCat provides help and comments, or simply that playing with the robot it more fun than playing with their colleagues (for example, because the iCat allows them to take back moves). Conversely, many participants from the *neutral* and *empathic* conditions mentioned embodiment aspects of the robot as the main difference, and compared the robot's skills to the skills of a computer:

> P16: *"the iCat doesn't have arms and we need to play his moves, and I need to say yes when he proposes a draw. I need to play really well to beat him."*
>
> P38: *"the main difference is that [the iCat] is a robot, and my colleagues are regular people like me. Also we don't think the same way [...] the iCat thinks the way computers think."*

The expressivity and emotional behaviour of the robot was also frequently mentioned as one of the main differences, especially by participants from both *empathic* conditions:

> P26: *"[my colleagues] are not so fun, they don't smile when they are winning and don't get sad when they are loosing... I've never seen their feelings, but at least their faces... they never get very sad."*
>
> P17: *"[my colleagues] don't get so excited, they don't show so many emotions... and the iCat gives us advice, and people don't."*

Some children also highlighted the fact that the iCat is more focused in the game than their colleagues but, at the same time, does not try to rush the game:

P33: *"my colleagues often say 'come on, play!' while I'm still thinking... and the iCat doesn't, he waits for my move, lets me think..."*

P35: *"the iCat is very calm, he is always there focused in the game. [my colleagues] are always looking at something else and I am waiting, looking at the [chess] clock."*

Advantages, Disadvantages and Suggestions

The last part of the interview contained questions in which children were encouraged to provide suggestions on how we could improve the robot, and indicate advantages and disadvantages of the interaction. The answers to the question "how can we further improve the iCat?" were coded into 4 different groups: suggestions regarding embodiment, behaviour and chess skills of the robot, and another category for the ones who considered that there is no need for improvement or were unable to name any suggestion. The category regarding embodiment included suggestions such as "the iCat should make less noise", "the robot should speak more slowly" or references to the addition of arms and other body parts, so that it could play its own moves. These suggestions regarding embodiment were more frequent in participants from the *neutral* condition. On the other hand, children from both *empathic* conditions provided more suggestions related to the chess skills of the robot. For example, some participants stated that to improve the iCat it should play chess more often, others that it should play an entire chess game instead of playing a chess exercise. As for the suggestions regarding the behaviour of the robot, while participants from the *neutral* condition proposed that the robot should be more helpful (or provide help when asked), children from the *empathic* conditions suggested improvements related to the expressivity and affective behaviour of the robot, for example:

P34: *"he is joyful, and we get happy... it's like playing with real cats... it's very funny because he smiles and gets very emotional about our moves".*

P33: *"you could include a system for the iCat to say what he really feels about the moves I've done. I wouldn't be upset if sometimes the iCat say what he really feels..."*

This last statement suggests that, even though the empathic behaviours of the robot improve the interaction in may ways, some users interpret these behaviours as less natural when compared to the behaviour of their chess mates. This result reinforces the idea that the robot should adapt its behaviour to the preferences of the child.

In terms of the advantages of playing with the robot, we did not found substantial differences in the advantages enumerated by children among the three conditions of the study, except for the fact that some children from the *empathic* conditions referred to the helping behaviour of the robot as one of the advantages. The most referred advantage indicated by the participants was that the iCat is a novel experience, as none of them had played with a robot before:

> P16: *"he is more funny because it's a robot, and we don't see that everyday."*
> P19: *"it is a novel experience, the iCat is a robot and I've never played with a robot before."*

Children also highlighted that they learn more when playing with the iCat due to the robot's comments during the game (when compared to playing with their colleagues or playing chess in the computer). Also, they considered that playing with the robot is more fun, which can be explained by the novelty effect of interacting with a social robot.

The majority of the participants said that there were no disadvantages or were unable to list any disadvantage. Only 6 children indicated disadvantages, 3 from the *neutral* condition, 1 from the *empathic* and 2 from the *adaptive empathic*. Some children considered that playing the robot's moves is a disadvantage. However, others referred that as a positive aspect in earlier phases of the interview. Also, some children said that the robot speaks too fast, for example:

> P39: *"the disadvantage is that he is a robot, and does not have arms or legs. Legs aren't a problem, but the arms... sometimes I didn't understand what he said, it was a bit difficult to understand."*

We believe that the novelty associated with the interaction with a social robot contributed to the low number of disadvantages indicated by children. Also, we noticed that most children were not very comfortable in talking about disadvantages to the interviewer, which they considered one of the "builders" of the robot.

6.3 Discussion

In this chapter, we presented a first evaluation of an autonomous social robot capable of performing an "affective loop", that is, recognising some of the user's affective states in real time and displaying appropriate empathic behaviour. The study was conducted in a school setting, more precisely in the room where children usually have their chess lessons. The results of the questionnaire suggest that children perceived the robot in both empathic versions as more engaging, helpful and also provided higher ratings in terms of self-validation. The end game result affected participant's engagement with and how helpful they perceived the robot, but no effect effect was found on the self-validation dimension. However, a more controlled study, using subjects with very similar chess levels, is needed to further investigate the end game effects.

The findings obtained in the open-ended interviews reinforce the idea that empathy is an important mechanism so that social robots can interact naturally with people. Overall, children enjoyed the interaction and often mentioned the empathic behaviours of the robot as positive aspects. These results are in line with previous empathy studies with adults, in the sense that empathy facilitates the interaction and affects positively the perception of the robot. However, if not chosen appropriately, the robot's empathic behaviours might have an opposite effect: while most of the children interacting with the *adaptive empathic* version of the robot reported happiness as the most prominent emotion during the game, children who interacted with the *empathic* version of the robot reported more negative emotions such as "anxious" or "nervous". These findings suggest that the selection of the empathic behaviours should be done carefully, under the risk of having the opposite effect. A similar result was discovered by Cramer et al. [36] (see section 3.2.2), where the presence of inaccurate empathic behaviours had a negative effect on adults' trust towards the robot. Also, some children felt that the robot was too nice for them, saying that they would not mind if the robot showed its "real feelings". These observations are important for deciding the "degree of empathy" in a social robot in two ways. First, the empathic behaviours (and the number of times they are employed) should be suitable to the application scenario where the robot is immersed. Second, as this might be a decision that depends a lot on the particular preferences of users, the robot should be able to learn the best strategies for keeping that user in a positive affective state.

The amount of speech and comments of the robot seem to be appropriate for this type of scenario. The analysis of the group interactions contributed significantly for understanding this, and also to refine some particular behaviours of the robot. None of the children seemed to notice that the robot does not have speech recognition capabilities, nor did they expressed any desire to interact with a more talkative robot. These results suggest that the perceptions of the game and the affective states of children are enough for generating believable verbal behaviour in social robots that act as game companions for children.

Despite the positive results in this study, the differences between the two empathic conditions, *empathic* and *adaptive empathic*, were still unclear. First, we believe that a long-term study is needed to properly evaluate the adaptive empathic model. During the whole interaction, each child played between 10 to 20 moves, which was also the number of times that the robot assessed the children's affective state. The amount of times that a supportive behaviour is employed is only a small part of this, as the perceived user's affective state needs to be negative with high probability. For this reason, some users might not have had enough time to realise that adaptation was taking place or, also very likely, the inherent noise of the reward function did not allowed for the robot to really capture the users' preferences in such a short number of interactions.

Another possible reason concerns with the method for calculating the reward function. The Affect Detection system takes into account contextual information of the game and visual information such as the probability of user being smiling and user's head direction, and all the features have roughly the same weight. However, in the six-seconds interval between the time that the user's affective state is first measured and the time that the reward function is calculated, the game state is unlikely to change (i.e., players usually take more than six seconds to play another move to change the state of the board). This may lead to situations where even if the supportive behaviour had a positive effect on users, the reward function does not totally captures this outcome. For example, if the user has a great disadvantage in the game, but a supportive behaviour employed by the iCat caused a strong positive reaction (e.g. made the user smile), this value is always attenuated by the contextual information of the game. Thus, to calculate the rewards, the affect recognition system should only take into account users' non-verbal behaviours and discard the context of the game.

Finally, the list of supportive behaviours available for the robot to display to the user was not very extensive. Since users only interacted with the robot once, this did not seem to affect much children's perception of the robot. However, in long-term interactions, we believe that the available supportive behaviours need to be more diversified and better structured.

Although more empirical work is necessary to validate the findings presented here, this study is an important contribution for the iterative process of designing an empathic robotic companion for children. Most of the discussed issues will be addressed in the long-term study presented in the next chapter.

Chapter 7

Long-term User Evaluation

This chapter describes the long-term experiment conducted to evaluate the final version of the proposed empathic model. Before presenting the methodology and results of the final study, we describe the main changes in the scenario as a result of the preliminary short-term study reported in the previous chapter.

7.1 Scenario Refinements

Considering the results of the short-term study from the previous chapter, some refinements were made to the iCat's behaviour. The two main refinements were related to the structure of the supportive behaviours and the adaptive algorithm for selecting the supportive behaviours of the robot, but some other minor improvements were also made. To address the suggestion given by one of the children that the robot should speak slower, we reduced the velocity of the text-to-speech, ensuring that all the robot's utterances could be understood by the children. Additionally, the memory of past interactions stored in the User Model component was used by the robot, since in this study the same users interacted with the iCat over repeated interactions.

7.1.1 Supportive Behaviours

One of the conclusions of the previous study was related to the need of structuring the supportive behaviours of the robot. Some of the previous support behaviour types were too similar which, among other reasons, contributed to the unclear tendencies regarding the preferences on the support behaviour types. For example, suggesting a good move to the user or the scaffolding behaviour were perceived as almost the same by the children.

Taking this into account, we decided to redesign the supportive behaviours based on the framework defined by Cutrona et al. [38], which separates social support in different categories: *information support* (advice or guidance), *tangible assistance* (concrete assistance, for example by providing goods or services), *esteem support* (reinforcing the other's sense of competence), *emotional support* (expressions of caring or attachment) and *network support* (social integration). This framework not only give us suggestions on how to generate different types of supportive behaviours, but also (and more importantly) gives us structure, since similar supportive behaviours are grouped in the same category. For example, there is a clear separation between task-oriented (information support and tangible assistance) and relationship-oriented behaviours (emotional, esteem and network support). Cutrona and colleagues even argue that there are individual differences, such as personality or gender, in the preferences of the support behaviour categories. This suggests that it may be easier (and more appropriate) for the robot to learn which support behaviour categories users prefer the most, rather than adapting to specific supportive behaviours.

Table 7.1 provides examples on how the different support behaviour categories were implemented in the iCat. With the exception of network support (for not being applicable to this particular scenario), supportive behaviours from all the other remaining categories were adapted to the robot's behaviour.

Social Support Category	**Supportive Behaviours**	**Examples of implementation in the iCat**
Information Support	Suggestion/advice	"Need help? Touch my paw so I can suggest you a move."
	Teaching	"That wasn't your best move, because now I can capture your Queen."
Tangible Assistance	Direct Task	(Play a bad move)
	Tension Reduction	"Do you want to start the exercise all over again?" "I always say, lucky in love, unlucky in chess."
Esteem Support	Compliment	"That was professionally done!"
	Validation	"Well done, you played what I would have played!"
	Relief of Blame	"Don't worry, you didn't have better options."
	Reassurance	"Something's not quite right here, but it will get better for sure."
Emotional Support	Relationship	"I really enjoy playing with you!"
	Understanding	"I understand how you're feeling, I've been through similar situations."
	Encouragement	"Come on, I still believe in you!"

Table 7.1: Examples of supportive behaviours implemented in the iCat based on the framework of Cutrona et al. [38].

7.1.2 Empathic Action Selection and Adaptation

The second major improvement was related to the RL policy used to learn and adapt the supportive behaviours to particular users. With the previous policy [6], described in section 5.4.2, the iCat always selected the behaviour with highest average rewards (with a small decay factor). Although this policy was a simple solution for addressing adaptation, it had some limitations which would become even more visible in long-term interactions.

One of the main limitations of the previous policy was that, if one of the behaviours received a high reward in the beginning of the interaction, that behaviour would be selected many times before the robot could select another type. This lead to situations in which the behaviour of the robot was too repetitive. Even if the selected behaviour type was the optimal one, this kind of repetitiveness is not desirable when users interact with the robot for extended periods of time. Moreover, this policy assumes that the user's behaviour is predictable. In other words, it assumes that the user's response to a particular strategy follows a well-behaved stochastic process. Users responses are likely to change over time, and their behaviour is difficult to represent in a probabilistic manner.

To address these limitations, we adopted a more sophisticated approach, Exp3 algorithm [7], that makes no statistical assumptions about the generation of rewards. Exp3 does not assume that the user's behaviour is predictable in the sense that his/her response to a particular strategy will follow a predictable distribution. In this policy, action selection is probabilistic, which means that even though the robot can learn the most effective support behaviour category for a particular user, behaviours from the other categories will also be employed, resulting in less repetitive behaviours.

Algorithm 7.1 describes the new policy for selecting the most appropriate supportive behaviours for a particular user. The intuition behind this algorithm is simple: the robot will select more often support behaviour categories with higher average rewards, since they have higher probability values in $p(t)$. However, the other behaviours also have chances of being selected, because even though they had lower rewards in the past, the user may find them more suitable in the future. For example, consider the case where the child already played one game with the iCat and is playing the second game. In the previous game, the child won the game because the exercise was very easy, but in this second game she/he is experiencing some difficulties in the game. Even though the child did not need the tangible assistance behaviours in the previous game, in this game she/he may find them useful because the game is more difficult. This new algorithm is more flexible in addressing this type of situations. Also, the repetitiveness of behaviours is less visible, which is also important for long-term interaction.

Algorithm 7.1 Exp3 [7], the algorithm used for selecting the adaptive supportive behaviours in the final study.

Parameters: $\gamma \in [0, 1]$

Initialisation:
for $i = 1 \rightarrow K$ **do**
  $w_i(1) = 1$
end for

loop
  for $i = 1 \rightarrow K$ **do**
    Set $p_i(t) = (1 - \gamma).\frac{w_i(t)}{\sum_{j=1}^{K} w_j(t)} + \frac{\gamma}{K}$
  end for
  Draw i_t randomly according to the probabilities $p_1(t), ..., p_k(t)$.
  Receive reward $x_{i_t}(t) \in [0, 1]$
  for $j = 1 \rightarrow K$ **do**
    $\hat{x}_j(t) = \begin{cases} x_j(t)/p_j(t) & \text{if } j = i_t \\ 0 & \text{otherwise} \end{cases}$
    $w_j(t+1) = w_j(t)exp(\gamma\hat{x}_j(t)/K)$
  end for
end loop
where:
$K \rightarrow$ number of possible support behaviour types
$x_{i_t}(t) \rightarrow$ obtained reward if behaviour i is chosen at time step t
$\hat{x}_j(t) \rightarrow$ estimated reward at time step t
$p_i(t) \rightarrow$ probability of selecting behaviour i at time step t

Note that the method for calculating the reward values is still the same (difference between new and previous affective state of the user). However, and also to address one of the issues discussed in the previous study, in such cases the Affect Recognition system only takes into account visual features, leaving out all the contextual features of the game. Thus, the contextual features are only used to assess the user's affective state after the user's move, that is, to decide whether the robot should employ or not a supportive behaviour.

It is also relevant to stress that this algorithm still needs to chose between four different types of actions or support behaviour categories, making $K = 4$ in this case too. Even so, while in the first study these actions were concrete behaviours (encouraging, scaffolding, suggest move and play bad move), in this version they are support behaviour categories that contain a higher variety of behaviours with a similar purpose, as described in the previous section. When the robot selects one of the behaviour categories using this algorithm, one of the behaviours of that category is selected depending on the context of the game. For example, if the esteem support category is selected, the robot only congratulates the user for capturing a piece if the user actually captured a piece in the previous move. When several behaviours from the selected category are eligible to display to the user, one of them is selected randomly.

We could have gone even further and also adapt to the particular behaviours within the same support behaviour category. However, since these behaviours are very similar, we believe that not only too many interactions would be necessary for the algorithm to converge, but also that no significant improvements in the interaction would occur.

7.2 Method

After implementing the refinements proposed in the previous section, a final long-term study was conducted with this scenario. In this study, children interacted with the iCat robot in their school environment within a five week period. The main hypothesis of this study was:

> *The proposed empathic model will have a positive impact on children's perception of the robot during long-term interaction.*

Several measures and data collection methods were used to investigate the effects of the developed empathic mechanisms in children's perceptions of the robot over time. We start by describing the methodology employed in this study, and then present and discuss the obtained results.

7.2.1 Participants

All participants in this study belonged to a Portuguese elementary school in which children have chess lessons as part of their extra-curricular activities. The school was the same of the short-term study reported in the previous chapter, but since different classes were selected, none of the children had interacted with the iCat before. A total of 16 participants from the 3rd grade were selected: 9 girls and 7 boys. Their ages varied between 8 and 9 years old (M=8.5) and their chess level was roughly the same, as all of them had chess lessons at least since the 1st grade. The study took place in the elementary school after the official school hours (from 4 p.m. to 6 p.m.), with the 3rd grade children who stayed in the school during that period doing their homework and other activities supervised by a teacher.

7.2.2 Procedure

The study was carried out over five consecutive weeks. Each child played a total of five chess exercises with the iCat – one exercise per week. The exercises consisted in playing from a predefined chess position until the end of the game (i.e., either the child or the iCat checkmates the other) and were suggested by the school's chess instructor so that the difficulty was appropriate to the chess level of the children. If after approximately 20 minutes none of the players has checkmated the other, the iCat either proposes a draw to the user (if it is in advantage) or gives up (if it is in disadvantage). The difficulty of the exercises varied over the sessions: in the first, third and fifth weeks the exercises were easier (i.e., the child started with advantage), whereas in the second and fourth weeks the exercises were more challenging to the child, since the iCat started with some advantage.

The procedure was similar every week: at the scheduled time, the child was guided to a room where she was alone with two experimenters and was asked to play a chess exercise with the iCat. In some rare exceptions due to public holidays or incompatibilities in children's schedule, they played two exercises in the same week, but always with at least two days of interval. Each game lasted, on average, 20 minutes, ranging between approximately 10 to 25 minutes. After playing with the robot, in the first and last weeks of interaction children filled in a questionnaire and were interviewed in a different room by another experimenter. After the fourth interaction session, participants filled in a Personality Questionnaire. All the interaction sessions were video recorded. Additionally, for each child, interaction logs were automatically saved in every interaction. The logs contain not only information about the game (e.g. all the moves played by both the child and the iCat, captured pieces, the game results, etc.), but also information related to the affective states of the children and the empathic behaviours employed by the robot in the different moments of the game.

7.2.3 Measures

Most of the measures have already been used in the preliminary user studies reported in chapter 4, to allow consistency and comparison with previous results. In this section, we briefly describe the main measures of the study. A summary of the measures and their corresponding data collection methods is presented in Table 7.2.

Social Presence

Social Presence measures "the degree to which a user feels access to the intelligence, intentions, and sensory impressions of another" [20]. This measure was also used in our first long-term study reported in section 4.2, using the same questionnaire items as the ones employed in this study. In that earlier study, some of the social presence dimensions decreased more over time, namely *perceived affective and behavioural interdependence* (the extent to which the user's emotions and behaviours affect and are affected by the robot's emotions and behaviours) and *attentional allocation* (the amount of attention the user allocates to and receives from the robot). In addition to the questionnaire, we also complemented the evaluation of attention towards the robot through video observation of all the interactions, by analysing the amount of time that users spent looking at the robot.

Engagement

Engagement is an important metric for long-term Human-Robot Interaction. If users are engaged, it is more likely that they will keep interacting with the robot for longer periods of time. We measured Engagement through questionnaires and video observation, by analysing the amount of time that children spent looking at the robot over the sessions. The questionnaire items for Engagement are based on the questions used by Sidner et al. [135] to evaluate users' responses towards a robot capable of using several social capabilities to attract the attention of users. The same questionnaire items were used in the short-term study reported in chapter 6.

Help and Self-Validation

Help and Self-Validation are dimensions of the Friendship Questionnaire employed in the study where the iCat observes and comments the chess match between two human players (see section 4.3) and also in the short-term study of chapter 6. Help and Self-Validation were measured through questionnaire items. With these two measures, we intend to evaluate how helpful children perceived the robot, and to what extend they consider the iCat as encouraging and able to help children to maintain a positive image of themselves. These measures are somewhat related to the perception of social support that will be described next.

Social Support

Perceived social support can be defined as “the belief that, if the need arose, at least one person in the individual’s circle would be available to serve one or more specific functions” [38]. This can only be measured after repeated interactions. As stated by Cutrona and colleagues [38], “it is necessary for the person to have experienced a number of interactions with the individual that communicates support”. As such, we measured perceived support only in the final questionnaire (5^{th} interaction session). The questionnaire items that measured perceived support were adapted from the Social Support Questionnaire for Children (SSQC) [60], a self-report measure designed to evaluate children’s social support via five different scales: parents, relatives, non-relative adults, siblings, and peers. In this case, we adapted the Peer scale by translating the items from English to Portuguese and changing “a peer” to “iCat”.

Adequacy of Adaptive Behaviour

One of the goals of this long-term study was to understand the impact that the different support behaviour categories – information support, tangible assistance, esteem and emotional support – had on children, and whether the robot's capacity to adapt its behaviour to the preferences of the child positively affected the interaction. As such, in the final interview, we gave participants four different cards containing a picture of the iCat and some speech bubbles containing sentences that the iCat says when it is employing behaviours from one of the behaviour categories (see Figure 7.1). Participants were asked to order the four cards from the one they liked the most to the one they preferred the least. From the interaction logs, we could obtain the final weights attributed to the different support behaviour categories for each particular child. By analysing the similarities between these two rankings (the one provided by the children during the interviews and the one obtained by ordering the categories from the one with highest weight to the one with lowest weight), we were able to evaluate the effects of adaptation during the interaction.

Figure 7.1: Cards containing different supportive behaviours displayed by the iCat.

Personality

We used the picture-based Personality Questionnaire for Children (PQC) developed by Maćkiewicz and Cieciuch [100] to measure participant's Big Five personality dimensions: *extroversion*, *neuroticism*, *openness to experience*, *consciousness* and *agreeableness* [103]. The questionnaire contains a set of behaviour-oriented items with two pictures containing two opposite behaviours (high and low intensity) for a particular personality trait. We used the version for younger children (tailored for elementary school children), where children rate these image-based items using a 3-point scale.

The validated English version of the PQC was translated to Portuguese. Then, a native English speaking person fluent in Portuguese translated the Portuguese version into English again. The two English versions (original and translated) were then compared and some minor adjustments were made. The final Portuguese version of the PQC is presented in appendix B.

Table 7.2: Summary of the methods and measures of the study.

Measures	**Questionnaires**			**Video Observation**	**Interviews**	**Interaction Logs**
	1^{st} week	**4^{st} week**	**5^{th} week**			
Social Presence	✓		✓	✓	✓	
Engagement	✓		✓	✓	✓	
Help and Self-Validation	✓		✓			
Social Support			✓		✓	
Personality		✓				
Adaptive Behaviour					✓	✓
Perceived Role Over Time					✓	

7.3 Results

This study had some limitations due to its long-term nature, and several design decisions had to be taken due to practical limitations both in terms of time and allocated resources. First, there was only one study condition. We tried to overcome this limitation by approximating as much as possible the experimental design to the one in the

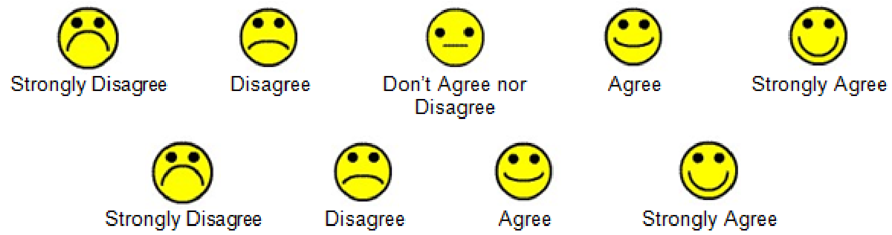


Figure 7.2: 5-point and 4-point Likert scales with *Smileyometer* [118] used in the questionnaire (translated to English).

first exploratory long-term study, to allow the comparability of some results. The second limitation was the limited sample size. Due to this limitation, sometimes we could not obtain statistically significant results. However, we tried to counter balance this by deeper qualitative analysis of the data (e.g., video observation and interviews).

As outlined above, questionnaires, video observation, in-depth interviews and interaction logs were combined to better understand the impact that the developed empathic mechanisms had on long-term interaction between children and a social robot. Given that several qualitative methods were used, not only results directly related to the main measures presented in the previous section are reported, but also other findings that emerged from the analysis of the collected data.

In this section, we present and discuss the results of this study. We start by reporting the results of the questionnaire data, followed by the video analysis, the interviews and the interaction logs. Finally, we present a general discussion of the obtained results.

7.3.1 Questionnaires

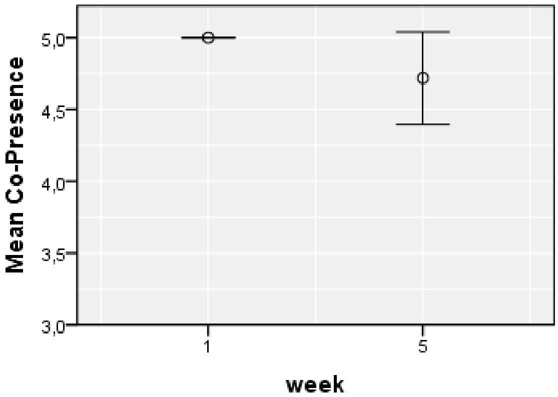
To measure Social Presence, Engagement, Help and Self-Validation, the questionnaire contained a set of assertions that children had to rate using a 5-point Likert scale. The Social Support items had a 4-point Likert scale. The Likert-scales were anchored by "Strongly disagree" and "Strongly agree" and by a *Smileyometer* [118], to help children interpret the meaning of the scale (see Figure 7.2). The final questionnaire used in this study is presented in appendix A.

We analysed the questionnaire data by comparing the results from the 1st and 5th weeks of interaction. The only exception was in the Social Support measure, because these items were present only in the questionnaire of the 5th interaction session. Our hypothesis is that participants' perception of Social Presence, Engagement, Help and Self-Validation would remain constant after five weeks of interaction. Since the goal is to show that the questionnaire results of the first session are indistinguishable from the results of the last session, traditional statistical tests of significance are not appropriate in this case [153]. As such, we performed equivalence tests by comparing the means and confidence intervals between the two groups (1st and 5th weeks).

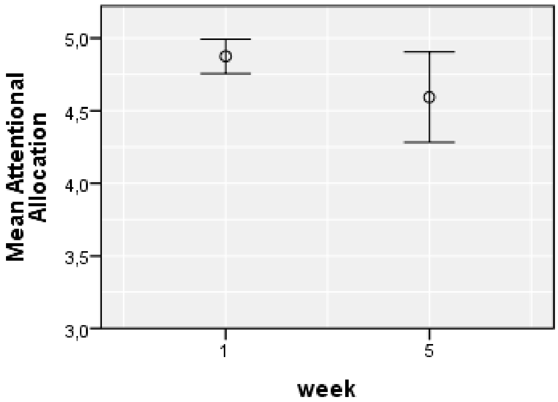
In addition to the measures that will be reported below, the final item of the questionnaire contained an assertion about whether users wanted to interact again with the iCat ("I would like to play with the iCat..."). Children had to mark one of the following options: (a) every day, (b) once a week, (c) once a month or (d) never again. In the first week, all the 16 kids except one selected option (a), and the only exception chose option (b). In the final questionnaire (5th week), all the children answered that they would like to play again with the iCat every day.

Social Presence

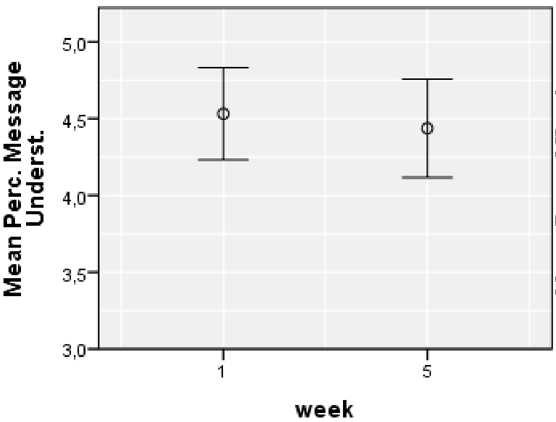
The means and confidence intervals of the Social Presence dimensions between the 1st and 5th week of interaction are displayed in Figure 7.3. In contrast with the results obtained in the first long-term study, the ratings remained roughly the same between the 1st and 5th interaction sessions, even in the dimensions that decreased over time in the earlier study – Attentional Allocation, Perceived Affective Interdependence and Perceived Behavioural Interdependence. In fact, for Perceived Behavioural Interdependence, the ratings even increased from the first to the last week (see Figure 7.3f). Given that the confidence intervals overlap in most of the cases, there is strong evidence that



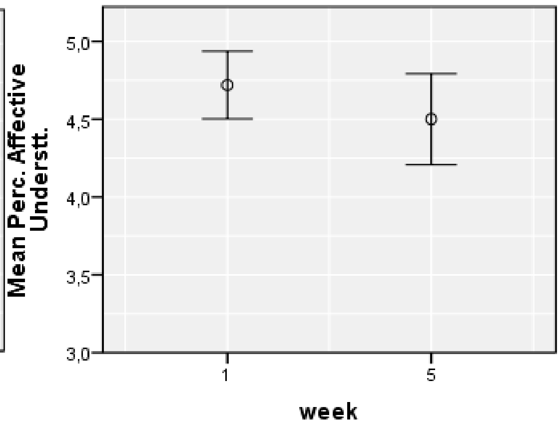
(a) Co-Presence



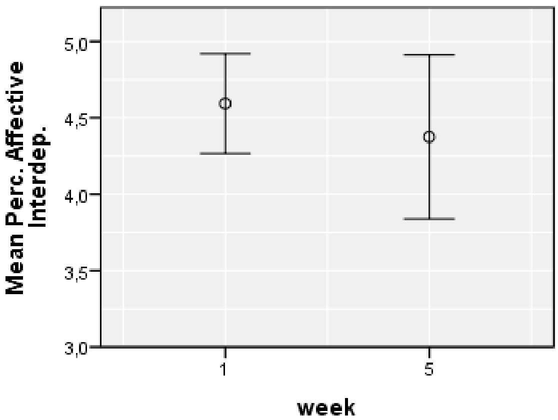
(b) Attentional Allocation



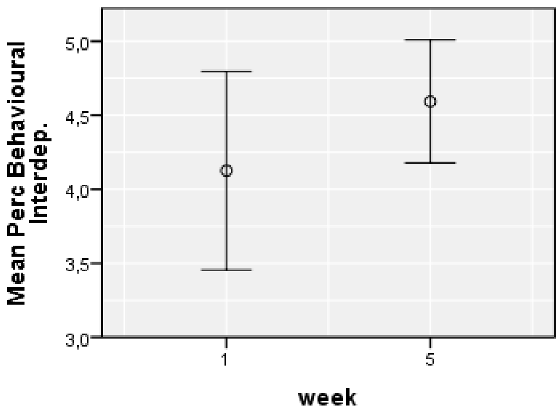
(c) Perc. Message Understanding



(d) Perc. Affective Understanding



(e) Perc. Affective Interdependence



(f) Perc. Behavioural Interdependence

Figure 7.3: Means and 95% confidence intervals for the ratings of the Social Presence dimensions in the 1st and 5th weeks of interaction.

in the two conditions (first and last week), children provided equivalent answers in terms of Social Presence.

Engagement

Similar results were obtained for Engagement, as we can see from Figure 7.4. After five weeks, children's ratings of Engagement were very similar to the ones given in the first week. To better interpret these results, we complemented these results with the analysis of the video recordings of the sessions (see section 7.3.2).

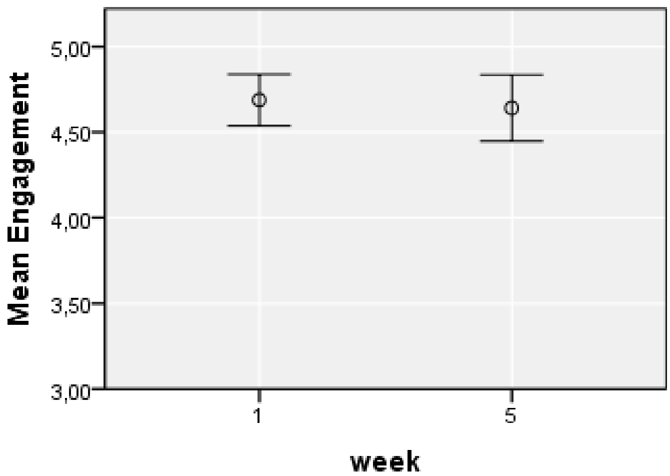
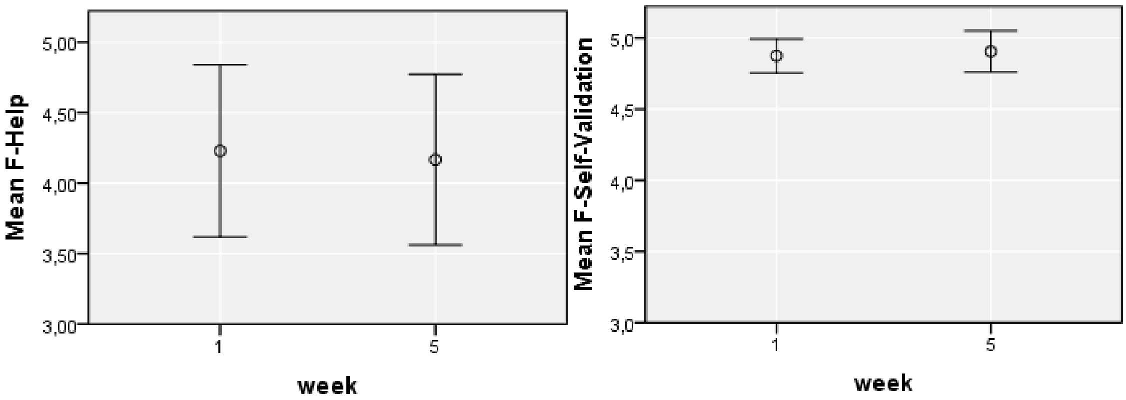


Figure 7.4: Means and 95% confidence intervals of the Engagement ratings in the 1st and 5th weeks of interaction.

Help and Self-Validation

The results for Help and Self-Validation followed the same trend as the ones for Social Presence and Engagement, as illustrated in Figure 7.5. In particular, the ratings for Self-Validation were considerably high in both sessions for all users, and the confidence intervals are very small (see Figure 7.5b).



(a) Help

(b) Self-Validation

Figure 7.5: Means and 95% confidence intervals of the Help and Self-Validation ratings in the 1st and 5th weeks of interaction.

Perceived Support

As explained earlier, the Perceived Social Support items were only included in the final questionnaire. Since it was the first time that we used this measure, we first ran Cronbach's alpha test to examine the internal consistency of our adapted version of the SSQC. The results revealed an acceptable consistency ($\alpha = 0.52$), although the original Peer scale from the SSQC had a higher reliability ($\alpha = 0.91$).

Table 7.3 contains, for each questionnaire item, the Means and Standard Deviations obtained in our study and the values obtained by Gordon [60] with a sample of 416 American children during the final stage of the questionnaire validation. In this latter sample, the mean age of children was 13 years old and, for this particular set of questions (Peer scale), children were asked to answer thinking on "anyone around your age who you associate with such as a friend, classmate, or teammate", whereas in our case they were asked to answer in relation to the iCat. Despite these differences, the positive results obtained in our study (all of them above the baseline mean values) suggest that, in this particular setting, the robot was perceived as supportive in a similar extent to what children in general consider being supported by their peers.

	Obtained Results		Baseline	
Questionnaire Items	**Means**	**St.Dev.**	**Means**	**St.Dev.**
iCat comforts me when I am upset.	2.88	0.48	2.25	0.82
iCat cares about me.	2.75	0.75	2.27	0.85
iCat gives me good advice.	2.44	0.86	2.15	0.83
iCat accepts me for who I am.	3.0	0.0	2.47	0.73
iCat supports my decisions.	2.94	0.24	2.27	0.76
I can count on iCat.	3.0	0.0	2.51	0.69
iCat encourages me.	2.81	0.39	2.35	0.8
iCat understands me.	2.63	0.78	2.46	0.74
iCat praises me when I've done something well.	2.94	0.24	2.21	0.89

Table 7.3: Means and Standard Deviations of the Social Support questionnaire items obtained in our study (2^{nd} column) and the baseline values (3^{rd} column) obtained by Gordon [60]. The scale ranges from 0 (Strongly Disagree) to 3 (Strongly Agree).

Personality

For each child, we averaged the scores of the items associated to each personality trait: Extroversion, Neuroticism, Openness to Experience, Conscientiousness and Agreeableness. The results of the personality questionnaire are displayed in Table 7.4. There were 6 children whose main personality trait was Agreeableness, 4 with Openness to Experience and 3 who were mainly Extroverts. Besides, there were some children with similar scores in two or even three personality traits. For example, 3 participants had the highest scores both in Conscientiousness and Agreeableness.

7.3.2 Video Analysis

The videos from the five interaction sessions of the 16 participants were analysed using ANVIL annotation tool [85]. As in the exploratory long-term study, we coded the

Participant ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Extroversion	2,5	2,7	2,2	2,7	2,8	2,7	**2,8**	2,5	2,7	**2,5**	2,8	2,2	2,5	2,3	**2,7**	1,8
Neuroticism	1,5	2,0	1,8	1,3	1,0	1,3	1,7	1,0	2,0	2,0	1,0	1,5	1,3	1,8	1,0	1,3
Openness to Exp.	2,8	2,5	**3,0**	2,7	2,2	**3,0**	2,7	**3,0**	2,5	2,3	2,3	2,0	2,2	**2,5**	1,8	**2,7**
Conscientiousness	2,4	2,8	**3,0**	2,6	2,8	2,8	2,6	2,8	**3,0**	2,4	**3,0**	**2,6**	2,4	2,0	1,8	1,4
Agreeableness	**3,0**	**3,0**	**3,0**	**3,0**	**3,0**	**3,0**	2,4	2,6	**3,0**	2,4	**3,0**	**2,6**	**2,8**	2,2	2,4	2,2

Table 7.4: Results of the personality questionnaire. The bold values indicate the most salient personality trait(s) of each child.

moments of the interaction in which users were *looking at the iCat*, *looking sideways*, *talking to the iCat* and also *children's affective expressions*. We also distinguished the phases of the game in which users were looking at the iCat: *after the user's own move*, when the iCat provides feedback to the user and employs empathic behaviours; *after playing the iCat's move*, when the user receives feedback from the robot after playing its move, and *during the game*, which includes all the other moments, for example, when the user is thinking about his/her move, when the interaction starts, ends, and so on. The following results are presented as percentages because the lengths of the interactions varied depending on the exercise and on the child.

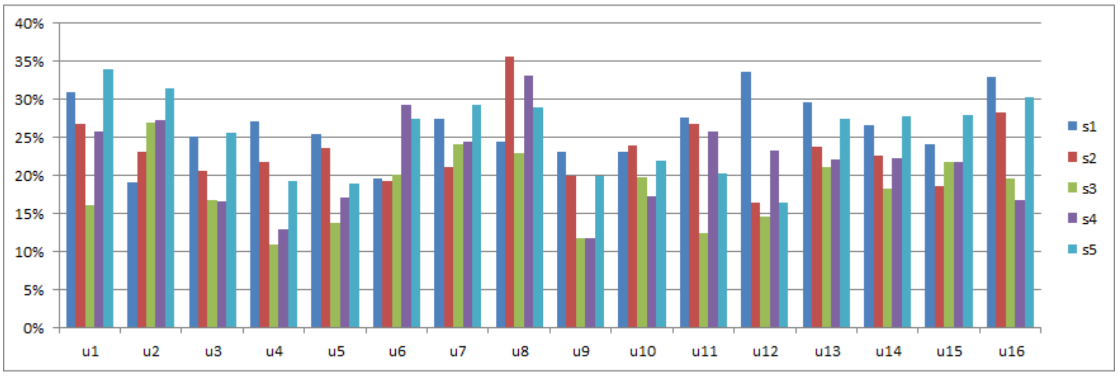


Figure 7.6: Percentage of time that each child (u1,...,u16) spent *looking at the iCat* in every interaction with the robot (s1,...,s5).

Figure 7.6 shows the percentage of time that each child spent looking at the iCat during the five interaction sessions. These values are summarised in Figure 7.8, which shows the average values for all users in each interaction session. Figure 7.7 displays

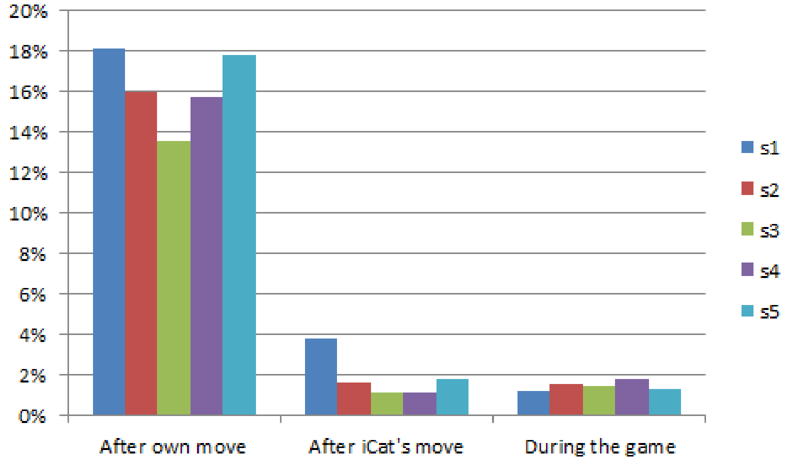


Figure 7.7: Average percentage of time that users spent looking at the iCat after playing their moves, after playing the iCat's moves and during the remaining parts of the game.

the same data but broken down by the phase of the game. In contrast with the results obtained in the first exploratory study, the values do not decrease in a linear way over time. In fact, there is almost no difference between the results from the 1st and 5th sessions, but in the middle weeks the average percentage of time that children looked at the robot decreased.

These results can be explained by the difficulty of the exercises in these sessions. As mentioned earlier, the difficulty of the exercises varied among sessions, resulting in different game results. As Table 7.5 suggests, the exercises from sessions 1 and 5 were easier (almost all children won), the exercises from sessions 2 and 4 were more difficult (more than half the participants lost the game) and the exercise from session 3 was the most challenging, since nearly half of the participants won the game and there was also a significant amount of draws. The chart from Figure 7.7 shows that the average percentage time that users spent looking at the iCat, especially after their own moves, is similar for exercises of the same difficulty. Therefore, it seems that the attention that children pay to the robot is related to the difficulty of the exercise: the more disputed the exercise was, the more children look at the chessboard rather than to the robot; and when they have greater advantage or disadvantage, they tend to be more engaged with the robot and look at it more often.

The results for the remaining annotations were considerably less frequent, as it happened in the initial long-term study. The *looking sideways* dimension only occurred when something disrupted the normal course of the interaction (e.g., someone entered the room by mistake) or when children looked at the experimenter if they did not understand something the iCat had said or done. This latter behaviour was more frequent in the first interactions. Apart from these situations, children were either looking at the robot or looking at the chessboard during the whole interaction.

The annotations regarding *talking to iCat* varied more between users rather than over time. Six out of the 16 children who took part in the study talked to the robot from the first until the last week of interaction, whereas the others usually did not (more details on the interview section). The most frequent words that children said to the robot were "thank you", usually after some praising comment of the iCat, and "check", a word that chess players say out loud when they play a move that threatens the other player's king.

The moments of the game where children exhibited more intense *emotional expressions* were at the beginning and at the end of the interactions. In the first minute of the each interaction, nearly all children displayed positive affective states (e.g., enthusiasm or joy), especially when the iCat mentioned their previous games together. The affective states displayed at the end of the interaction depended on the game result. Usually, children exhibited happiness when they won or draw the game, and showed embarrassment when they lost the game. During the remaining moments of the interaction, the most salient affective expression was surprise, usually when the iCat said something that children were not expecting or when they were in check (iCat threatening their king).

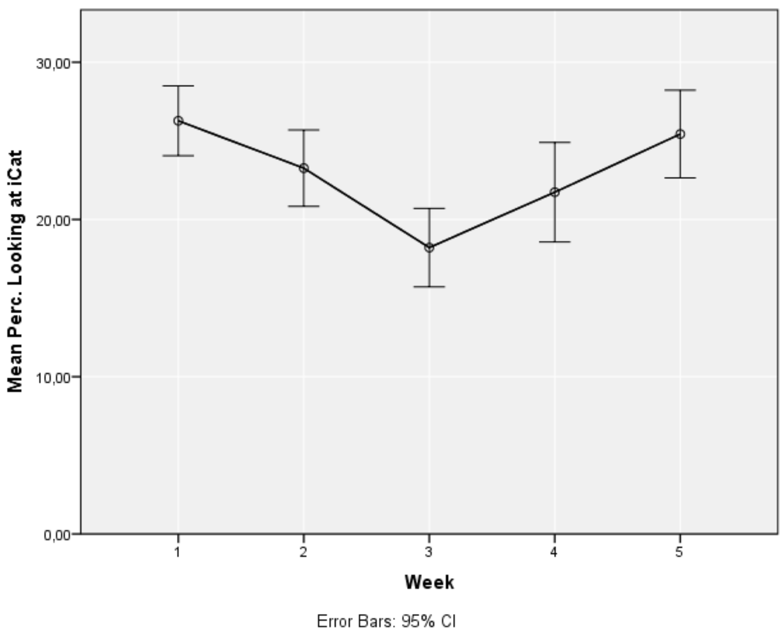


Figure 7.8: Means and 95% confidence intervals of the percentage that children spent looking at the iCat in each interaction session.

	session 1	session 2	session 3	session 4	session 5
Victories	14	6	9	0	12
Draws	2	1	6	10	3
Defeats	0	9	1	6	1

Table 7.5: Total number of victories, draws and losses of children when playing against the iCat for each interaction session.

7.3.3 Interviews

In this section, we present the most interesting findings obtained during the analysis of the open-ended interviews both in the first and last weeks of interaction. Children were interviewed in a different room by another experimenter. The interview questions were separated by themes. For each theme, there was a common set of questions in both interviews and, in the final interview, additional questions were added and a new theme related to changes during interaction was included. Table 7.6 contains the interview questions for the different themes.

The methodology for analysing the interview answers was similar to the one described in section 6.2.2. From now on, we will refer to the different participants of the study as P1, P2,..., P16.

Theme	**Question**	**Response Type**
Perceived role	1. Do you believe iCat is more similar to an opponent, a teacher or a friend? Why?	select option, open
Perception of supportive behaviours	2. How did the iCat behave when you had problems during the game?	open
	3. How did the iCat behave when you played well?	open
	4*. Did you understand the iCat's facial expressions and comments? How so?	yes-no, open
	5*. What was the weirdest thing the iCat said or done?	open
Preferences on supportive behaviours	6*. These cards contain sentences that the iCat said to you while you were playing together. Please order these cards from the one you liked the most to the one you like the least.	ranking
Changes during interaction	7*. Do you think your actions changed the iCat's behaviour? How so?	yes-no, open
	8*. Did you enjoy more the first games, the last ones or was it always the same? Why?	select option, open
Advantages, disadvantages, suggestions	9. Enumerate the 3 things you liked the most and the 3 things you liked the least in the iCat.	open
	10*. If iCat could help you in other tasks, what would you chose?	open
	11. Do you have any suggestions to improve the iCat?	open

Table 7.6: Interview questions divided by themes. The questions marked with * were only asked in the final interview.

Perceived Role

One of first questions that we asked children both in the first and in the final interview was related to the perceived role of the robot, by asking them the following question:

Do you believe iCat is more similar to an opponent, a teacher, a friend, or something else?

As we can see from the chart displayed in Figure 7.9, children's perception of the robot changed over time. While in the first interview most of the children considered the robot as an opponent, in the final interview most of them considered the robot as a friend. When we asked them to justify their answer, children provided answers expected in a regular human-human social interaction. For example, P9's justified her answer in the first week by saying:

> *"For now, I think he is just an opponent... just for today! Because we didn't know each other, and I didn't know well how he was."*

In the final interview, many children found similarities between the iCat's behaviour and the behaviour of their colleagues while playing chess, as we can see from P8's answer:

> *"He is more like a friend. Because it's like playing with Maria (my best friend). Whenever I play a bad move, she says 'no, you can't play that, if you want I can tell you a good move' and then she tells me a good move (...)"*

Perception of Supportive Behaviours

The questions in this theme worked as a manipulation check to understand whether children perceived correctly the behaviours implemented in the robot. In general, children perceived the robot's behaviour as we intended to: most of them answered that the robot helped them when they were experiencing difficulties in the game (Q2), and that the iCat praised them when they played well (Q3). We found no substantial differences to these questions between the answers from the first and the last interview.

The same results were found in the question regarding the expressive behaviour of the iCat in the final interview (Q4). All children answered "yes" when we asked them

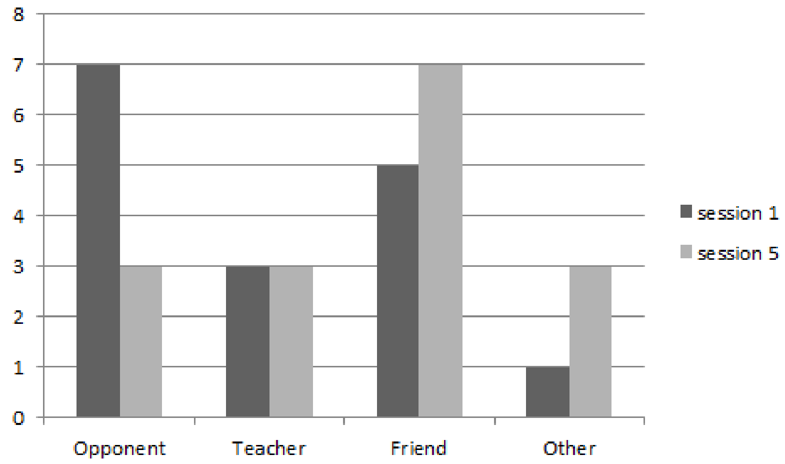


Figure 7.9: Children's answer to the interview question *Do you believe iCat is more similar to an opponent, a teacher, a friend or something else?* in the 1st and 5th weeks.

if they understood the iCat's expressions. However, when asked to elaborate on their answer, only 14 out of the 16 children provided a valid answer. The remaining 2 children could not say why did they understood the expressions or provided an incorrect answer, for example, by saying that the robot got happy when they played bad moves.

As for Q5 (*What was the weirdest thing the iCat said or done?*), 6 participants referred to concrete moments of the game. For example, P10 answered:

> "*He let me capture his Queen, and I didn't understand why!*"

These answers suggest that the supportive behaviour "Play Bad Move" was not completely understood as a deliberative action of the robot, but rather as a mistake. Additionally, 7 participants considered the comment "Lucky in love, unlucky in chess" as the weirdest behaviour of the iCat. The remaining participants did not found anything strange in the robot's behaviour.

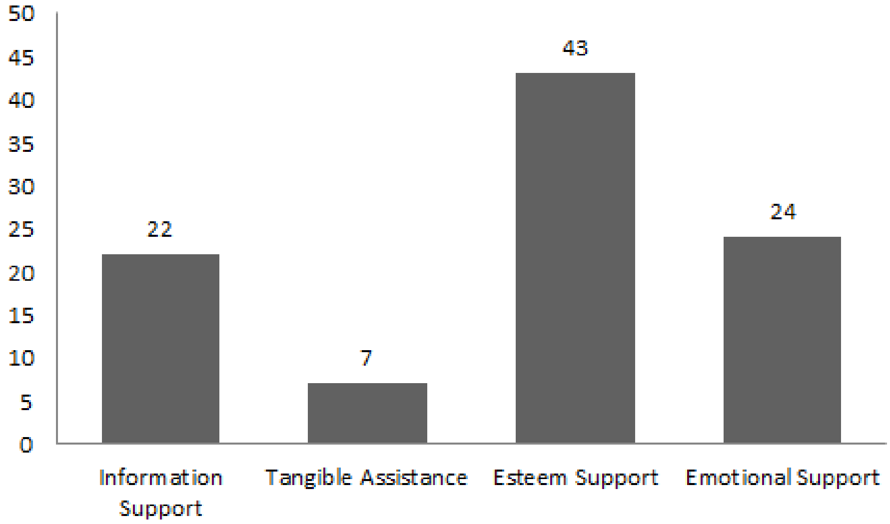


Figure 7.10: Overall rankings of the preferred support behaviour categories.

Preferences on Support Behaviour Types

As mentioned in section 7.2.3, in the final interview we asked children to order a set of four cards containing utterances illustrating the different types of supportive behaviours displayed by the iCat (Q6). To analyse these data, we classified children's rankings as follows: we attributed 3 points to the most preferred support behaviour type, 2 points to the second most preferred, 1 point to the third and 0 points to the least preferred. We summed the points for each category for every child, ending up with the ranking displayed in Figure 7.10. As we can see from the Figure, the preferred supportive behaviour category was esteem support, followed by emotional support, information support, and finally tangible assistance.

Changes during Interaction

This theme contained two different questions that aimed to investigate the main changes in children's perception of the robot throughout the sessions. First, we asked children directly if they considered that their actions change the robot's behaviour (Q7). This question is also related to the dimensions of Perceived Affective and Behavioural Interdependence of Social Presence. 14 participants responded affirmatively to this question. When asked to justify their answer, 7 children stated that the iCat's behaviour changed based on their feelings, and 6 children considered that the robot's behaviour changed due to the moves they played. One of the participants (P4) even mentioned that the iCat changed its behaviour because it was able to learn over time:

> "*He learned about me with every move I made, and then used that to improve...*"

As for Q8 (*Did you enjoy more the first games, the last ones or was it always the same? Why?*), the last games were the ones slightly preferred by most children, as depicted in the chart from Figure 7.11. The novelty effect of playing with a robot for the first time was the main reason why children preferred the first games. For example, P3 said:

> "*Because in the first game I had never seen the iCat, I didn't know what was going to happen...*"

On the other hand, children who preferred the final games provided two main reasons. First, they considered that they played better in the final games. Secondly, they considered the robot's behaviour as more appropriate. P9, for example, valued the robot's advice more in this session:

> "*He gave me advices that he had never given before (...)*"

Finally, one participant (P10) said that he preferred the final games because he had more fun at the end:

> "*It was more complex this time, but I enjoyed more... in the first games I was too nervous!*"

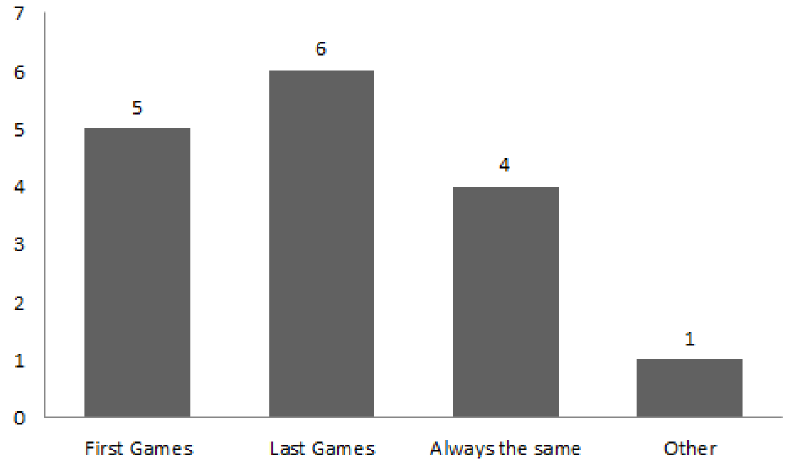


Figure 7.11: Children's answers to the question "Did you enjoy more the first games, the last ones or was it always the same?"

Advantages, Disadvantages, Suggestions

The answers to the questions about suggestions and enumerating positive/negative aspects of the iCat were similar to the ones given by the children of the study reported in section 6.2.2. Moreover, we did not find significant differences between the answers from the first and the last interviews.

However, in this interview, there was an additional question (*If iCat could help you in other tasks, what would you chose?*) that yielded interesting findings. 7 of the 16 children answered that the iCat could help them with their homework, while 4 children would like to receive advice from the iCat on several matters, such as:

> *"He could help me when I was feeling sad."* (P9)
>
> *"He could help me in solving problems that I have in school."* (P12)
>
> *"He could give me his opinion on different subjects (...)"* (P13)

One of the participants (P2) who would like to receive advice from the iCat also expressed desire to do other things with the robot:

> *"Everything! I could play with him at school, and even invite him to sleep over sometimes. but since he is a robot, I don't know... does he sleep at all?"*

Additionally, two subjects wished that the iCat could play other board games with them, and tasks such as football, building other robots and doing housework were also mentioned.

7.3.4 Interaction Logs

By combining the interview data about children's rankings of supportive behaviour types with the final interaction logs, we were able to evaluate whether the robot's adaptive behaviour indeed reflected children's preferences. In other words, analyse if after the 5 interactions, the preferred supportive behaviours of each child were the ones that the robot used more often while interacting with that child. To do so, we compared the children's rankings obtained during the interviews with the final weights of the behaviour categories retrieved from the interaction logs individually for each child. To statistically compare these two variables, we applied Kendall's Tau, a statistical test that measures the correlation between two rankings.

As we can see from the results presented in Table 7.7, there was a positive correlation between the two rankings for 8 of the 16 participants (although not always statistically significant), a negative correlation for 7 participants and a neutral correlation (independent values) for one participant. This means that, at least for half of the participants, the robot was able to capture the users' preferences correctly using the implemented adaptive mechanism, and employed more often the strategies that children preferred the most.

Another interesting finding is that the success of the adaptation seems to be related with children's Extroversion scores on the Personality test. By applying Pearson

Participant ID	Kendall's tau	Sig.	Extroversion
1	-0,33	0,25	2,5
2	-0,33	0,25	2,67
3	-1	0	2,17
4	-0,33	0,25	2,67
5	1*	0	2,83
6	0,67*	0,09	2,67
7	0,67*	0,09	2,83
8	0,33*	0,25	2,5
9	-0,33	0,25	2,67
10	1*	0	2,5
11	0,33*	0,25	2,83
12	-0,33	0,25	2,17
13	-0,33	0,25	2,5
14	0	0,5	2,33
15	0,33*	0,25	2,67
16	0,33*	0,25	1,83

Table 7.7: Correlation coefficient between the supportive behaviour category rankings in the interview and the final values obtained from the adaptation algorithm for each user. The values marked with * are positively correlated. The last column contains children's Extroversion scores obtained in the Personality Questionnaire.

correlation test between Kentall's tau coefficients and Extroversion scores, we found a moderate positive correlation ($r = 0.363$, $p = 0.84$). Even though the value of this correlation was not statistically significant, this result suggests that the adaptive mechanisms are better suited for extrovert children, which are usually more expressive and less shy to display affective reactions. Note that our mechanism for calculating rewards was based on children's non-verbal affective behaviour.

7.4 Discussion

Overall, the obtained results are consistent with our main experimental hypothesis: the developed empathic model had a positive impact in long-term interactions between children and a social robot. In this discussion, we often compare the obtained results with the ones reported in the first exploratory long-term study described in section 4.2, since the study design was nearly the same.

The ratings of Social Presence, Engagement, Help and Self-validation remained similar after five weeks (the 95% confidence intervals overlap in a substantial amount in every dimention of the questionnaire), contrasting with the results obtained in our initial exploratory long-term study where the robot was not endowed with the empathic model. Moreover, the means and confidence intervals of one of the Social Presence dimensions – Perceived Behavioural Interdependence – even increased in the last week (comparing to the first week in this study), which suggests that, over time, users were even more aware that their actions influenced the iCat's behaviour. We also found that children felt supported by the robot in a similar extent to what, in general, children feel supported by their peers. Another interesting finding was that the most valued form of social support was esteem support (reinforcing the other person's sense of competence and self-esteem).

The video analysis yielded results aligned with the ones obtained in the questionnaire data. The amount of time that users look at the robot seems to be more related to the difficulty of the exercise rather than by the number of previous interactions with the robot. The interactions in which children paid less attention to the robot were the ones where they had higher chances of winning the game. In the first and in the last games, the difficulty of the exercises was roughly the same, and so was the average "looking at" time. Therefore, it seems reasonable to conclude that the amount of time that users spent looking at the robot did not decrease over time simply because the novelty effect faded away, in contrast with the results obtained in the first exploratory study.

The results discussed so far were complemented with the analysis of the interview data. In the interviews, we confirmed that the implemented supportive behaviours were well understood and valued by children. In general, the emotion-oriented behaviours – esteem and emotional support – outranked the task-oriented behaviours – information support and tangible assistance. This result can be interpreted in two ways. First, it may be the case that, when playing competitive games, children prefer less tangible ways of support in contrast to being helped directly by the robot, which can reduce the merit of their victory (if they end up winning the game). An alternative explanation is that is that the implemented task-oriented behaviours were not helpful enough. In the social support literature, task-oriented support often includes behaviours such as lending the other person something (e.g., money) or offering to take over of the other person's responsibilities while he/she is under stress, which are behaviours that are not applicable to this scenario (nor in most of the existing HRI scenarios). The esteem support category contained behaviours in which the iCat praised the user. As such, these results are in line with previous findings in HCI, in which computers capable of some forms of flatter are perceived more positively by users [53]. The low rankings of tangible assistance might have been caused by the concrete behaviours implemented in this category – play bad move, start from beginning and use of humour – as already discussed in the previous subsection. Since children from this chess level often commit mistakes in the game (e.g., letting the other player capturing an important piece), during the interviews we noticed that most children associated the robot's bad moves to an involuntary fault that happened not necessarily to help them. Regarding the use of humour (one of the possibilities for tension reduction), they may have become too repetitive over the course of the interactions because the robot only had two different jokes to tell.

Another interesting finding for long-term interaction was that children started considering the robot more as a friend and less as an opponent. Usually, people are more willing to spend time with friends than with strangers. As such, if robots can be seen as "friends", it is expected that they will be more capable of sustaining long-term interactions with users. Moreover, after the five weeks, the role of the robot seemed to be clear on children's minds. When asked about other tasks that they envision the iCat doing, most of them suggested helping with their homework or giving them advice in other subjects. These answers suggest that most children perceived the robot as a peer who can assist them in cognitive tasks.

In the results concerning adaptation, there is some evidence that the implemented mechanisms were effective for most users, especially for extroverts. These results were not totally unexpected, since the reward system was based on user's expressive behaviour.

Nevertheless, they indicate that at least for expressive children (or maybe scenarios where children are more expressive than in chess), children's non-verbal behaviour can be a very informative feedback mechanism for social interaction with robots and other virtual characters.

As with any user study with children, the results need to be interpreted in a cautious way. Children have an intrinsic tendency to please adults, a well studied phenomenon in the field of psychology known as *suggestibility* [29][133], which depends not only on the content and format of the questions, but can also be influenced by other factors such as the age and gender or the interviewer/experimenter [119]. To undermine suggestibility, we tried to combine different data collection methods and analysis (questionnaires, video observation and interview data), but we are aware that there is a lot of pressure for children to behave in school contexts and this factor needs to be taken into account when interpreting the results.

Chapter 8

Conclusions

This thesis addressed the role of empathy in long-term human-robot interaction. More specifically, we investigated the effects of several forms of empathic and supportive behaviour in social robotic companions, as well as different mechanisms to generate such behaviours. An empathic model that allows social robots to generate appropriate empathic responses and supportive behaviours in the long-term was developed. An iterative approach was taken to develop the model, which was continuously evaluated and refined through several user studies.

We started by performing a long-term exploratory user study to understand which aspects change the most when users interact with robots for extended periods of time. The results of this study supported our initial hypotheses that the robot should understand and react appropriately to some affective states of the users, under the risk of being perceived as not "socially present". But *does the presence of empathic behaviour in a social robot affect the way people perceive the robot? Which perception and expression mechanisms are necessary for a social robot to behave in an empathic manner?* To address these questions, we performed a follow-up study in which simple empathic behaviours (facial expressions and verbal comments) were evaluated in a scenario where a social robot comments on a chess match between two human players, empathising with one of the players and behaving neutrally towards the other. From this study, we concluded that the manipulation of simple empathic behaviours in the robot can impact how users perceive the robot in terms of perceived friendship.

These preliminary studies suggested that empathy may play an important role in long-term human-robot interaction. However, one question remained unanswered: *how to select appropriate empathic behaviour based on the affective state of the user?* Moreover, when the robot has several empathic behaviours available to display to the user, *how to chose personalised empathic strategies for a particular user?* To answer these questions, we proposed an empathic model that allows a robot not only to react empathetically, but also to employ the most effective supportive strategies for a particular user. Two different user studies were performed to evaluate this model, both in a scenario where an iCat robot plays chess with children. The first was a short-term study in which we compared three different versions of the robot's behaviour: a neutral version, a version where the supportive behaviours of the robot were selected randomly and, finally, a version with our adaptive model. The last version was the one with more positive results. For example, children who interacted with the adaptive empathic robot reported more positive affective states than children from the remaining two conditions. After some refinements in the case study scenario, namely in the structure of the support behaviour categories and on the algorithm for selecting the supportive behaviours, a final long-term study was conducted with the goal of answering the final research question proposed in this thesis: *what is the effect of the robot's empathic strategies in the long-term relationship established with the user?* In this study, we showed that the developed model had a positive impact in long-term interaction between children and a social robot.

The results of this dissertation raise the possibility of a link between empathy and higher perceived social presence, engagement, help and self-validation in children who interact with social robots for extended periods of time. This suggests that empathic robots will be more likely to engage users in the long-term than robots without such capabilities. However, the research leading to these results was conducted in a scenario where children played with a social robot in a school setting. Therefore, more studies need to be conducted to assess whether the same results apply to other settings and user groups. The same cautions should be taken concerning the validity of the empathic model. Although the model was designed to be generic, the hypothesis of the model applicability to other domains remains unvalidated, since we used the same scenario throughout this work to allow the comparison between different stages of development.

8.1 Design Guidelines

Considering the state of the art and the work presented in this thesis, we developed a set of guidelines for designing artificial companions for long-term interaction. This preliminary set of guidelines aims to help researchers with the goal of developing social robots capable of engaging users for extended periods of time. This proposal should not be understood as final, but rather as an initial specification that requires further investigation and refinements. We would also like to stress that these guidelines were specifically formulated for long-term human-robot interaction, and thus may not represent the entire needs of the HRI field in general (discussed, for example, in the work of Kahn et al. [77]). Additionally, we also present a set of guidelines that can be useful in the design of long-term interaction user studies.

8.1.1 Building Artificial Companions

1. Appearance depends on application domain and should reflect companion's capabilities. Embodiment can play an important role in the first impressions and future expectations that people create about a robot. In a study using static images of robots, DiSalvo and colleagues [44] found that the facial features of several robotic heads influenced significantly the perception of humanness in those robots. A study by Hayashi et al. [64]

also concludes that "if a robot's appearance and behaviour is less human-like, people would expect less cognitive human-likeness", which at least for now is desirable given the limited autonomous capabilities of robots. Human-like appearance is not always desirable, as suggested in Mori's theory of the uncanny valley [107]. Moreover, as argued by Dautenhahn [39], "anthropomorphism might raise false expectations regarding the cognitive and social abilities that the robot cannot fulfil". This is especially true in long-term interactions, since it is more likely that these "flaws" become visible over time. For example, in the "roboceptionist" study reported in the related work section [58], the human-like face of the robot (displayed in a computer screen) created very high expectations in users, particularly in terms of the robot's dialogue capabilities. Since these expectations were not met, the amount of time that users spent interacting with the robot decreased day after day.

Therefore, the choice of the robot's appearance must take into account not only its behavioural and social capabilities, but also the application domain where the robot will operate and its function. For example, while animal-inspired embodiments are well suited for health-care related domains, as they elicit care-taking behaviours from humans, functional embodiments are more appropriate for work environments or domestic settings, where the ways in which the robot can assist users are a major determinant.

2. Companions should perceive and react appropriately to the affective states of users. Several researchers have investigated the display of emotions and other non-verbal behaviours by social robots in long-term interaction studies [90, 95, 74, 86], but the effects of empathy (that is, understanding the user's affective state and displaying appropriate affective reactions) in long-term interaction studies have not been investigated extensively so far. However, some researchers even consider the awareness of the user's affective state more important than the actual display of emotions by the robot. While discussing the desired features of future artificial companions, Pulman [117] argues that "a companion which behaved in the same way whatever our emotional state would be thought of as insufficiently aware of us. But this may not mean that the companion itself has to express emotions: all that is necessary to achieve this is the ability to recognise our own displays of emotion". In this dissertation, we presented empirical evidence suggesting that empathy plays an important role in long-term human-robot interaction.

In humans, it is believed that empathy facilitates the creation and development of human social relationships [3, 42]. Moreover, many authors also highlight the central role of empathy in learning [5, 126], which makes this capability very important for social robots developed for educational domains. For these reasons, we believe that the awareness of the user's affective states and, more importantly, the capacity to correctly respond to those states, are extremely relevant features for future artificial companions.

3. Companions should display continuity and incremental novel behaviours.
As outlined in section 2.1, to maintain a social relationship, humans perform a series of activities to keep the relationship in a satisfactory condition. Greetings and farewells, or performing everyday tasks together, are examples of routine behaviours relevant for the maintenance of social relationships. Examples of strategic behaviours include relational communication (e.g., talking about the relationship or recalling past activities together), performing planned activities together and self-disclosure (revealing personal information as a mechanism to give and receive advice that increases trust and intimacy) [46].

Some of these relationship maintenance strategies were successfully applied in our application scenario and also in the works presented in the related work section, namely greetings and farewells [84, 128, 58] and self-disclosure [78, 80]. For example, in our scenario, the fact that the iCat used children's names when performing the greeting behaviour at the beginning of the interactions always elicited positive feelings on children. These mechanisms are important for keeping users interested and with a sense that they are "unique" for the robot. As such, they are of particular relevance in application domains such as Education or Public Spaces, where users usually have more freedom to abandon the interaction if they wish.

4. Companions should remember past interactions with the user and use that infor-

mation to personalise the interaction.

The ability to memorise past interactions will definitely make social robots more flexible and personalised to particular users, regardless of the application domain in which they are supposed to operate. As stated by Dautenhahn, “rather than relying on an inbuilt fixed repertoire of social behaviours, a robot should be able to learn and adapt to the social manners, routines and personal preferences of the people it is living with” [39].

To achieve this, researchers have been investigating biological-inspired computational models of memory to be included in social robots or virtual agents that will interact with users for extended periods of time [68, 99], but the actual benefits that memory can bring are still unclear. Nevertheless, they anticipate that memory will give users the impression of behavioural coherence and plausibility, and therefore it might positively influence the perception of intelligence and quality of the interaction with the robot [99]. However, even simple memory mechanisms can make a difference. For example, in our scenario, the robot was able to memorise the result of previous games with a user. Being able to recall those moments was often regarded by the children as one of the positive aspects of interacting with the robot.

But memory can also bring other benefits, such as the ability to adapt or personalise the interactions. In our case, the robot “memorised” the most effective supportive behaviours for a particular user, and was able to use that information to better respond to the user in future interactions. Similarly, in the study by Koay et al. [89], simple preferences regarding proxemics and type of the responses employed by the robot (e.g., the amount of interruptions, how users like to be addressed, etc.) contributed to increasing trust and sense of control towards the robot.

8.1.2 Long-term Interaction Studies

1. Sample size and number of interaction sessions should be reasonable to allow complete data analysis.

One of the most prominent challenges when conducting long-term interaction studies is the time and effort required to analyse large amounts of data. Not only because one needs to collect data from several interaction sessions, but also because, since usually the number of subjects is limited (also due to time restrictions and because of the difficulties in recruiting participants for long-term studies), qualitative methods are more valuable, yet again more time consuming.

As such, the number of participants should be reasonable considering the number of interaction sessions, the employed data collection methods and the number of human resources allocated to analyse the data. For example, if only one person is assigned to work on the data analysis, there is no point in collecting videos of 10 sessions from 100 participants, since it is not feasible for a single person to analyse such large amounts of data. For the same reasons, the number of interaction sessions should also be planed carefully. Evidently, the number of sessions needs to capture users' repeated exposure to the companion, but having too many sessions beyond that will increase the amount of data to analyse, possibly without different results.

2. Use control conditions carefully.
The use of one (or more) control conditions should be done only if extremely necessary. Most of the existing long-term studies performed so far do not have a control condition. This happens not only because the studies are exploratory, and having a control condition usually duplicates the amount of data to be analysed, but also because user's experience over time can already be considered a strong independent variable.

3. Data collection methods should take into account the environment and types of users of the study.
The data collection methods selected for the experiment should be adequate to both the environment where the study is taking place and the type of users that will interact with the robot. While video recordings may be appropriate for studies in public spaces, they are not very suitable for domestic settings. Nevertheless, qualitative mea-

sures are often preferred to quantitative measures, mainly due to the issues of sample size. Ganster et al. [56] stressed the importance of using appropriate research methods for analysing long-term interactions, stating that methods and instruments that have been applied successfully in short-term studies might not be appropriate for long-term interactions. They provide a set of objective and subjective measurements that could be used successfully in this case (e.g., keeping diaries, task-accomplishments and change of performance, physiological methods, eye tracking or the analysis of audio and video material), arguing that a combination of different types of methods would bring value both in terms of data quality and reliability.

Additionally, it is important to collect data through all the interaction sessions – and not only from the first and last sessions. This can help to determine at which point of the interaction the user's attitudes towards the robot has changed (if there was a change). In sum, when selecting measures for longitudinal studies, there are several aspects that need to be balanced. These aspects include not only the advantages and disadvantages of the type of measure in terms of data quality and reliability, but also the amount of time and costs required to collect and analyse such data.

8.2 Future Work

There are several research directions that can be taken in the future as a result of the work presented in this thesis. First, we proposed and evaluated an empathic model with several components (empathic appraisal, supportive behaviour, model of the user and adaptation). However, since we evaluated the entire model, we were unable to understand which specific parts contributed the most for improving users' perception of the robot in long-term interactions. As such, future long-term studies with parts of the model being tested individually need to be carried out.

The proposed model is generic and thus could be applied to other social robots in different application scenarios, but in this thesis only one case study was evaluated. Nevertheless, the implementation of this model in other social robots should be quite straightforward. Moreover, we believe that the empathic model would also be relevant in other scenarios different than turn-based games, for example, in learning environments or health-care related applications.

Another aspect that could be further explored is related to the adaptation mechanism. We proposed a non-intrusive RL approach in which the robot can learn and adapt to the user without the user having to explicitly do nothing more than engaging in a natural interaction with the robot. Using this approach, social robots in other contexts could adapt to different user preferences in real time. For example, a social robot assisting users in their homes could learn the best ways to approach to the user (e.g., speaking or waving), or even the accepted social distance that it should keep from the user during the interaction. A robot in a playground environment could learn which games certain children prefer the most. In these examples, the reward function could be based on user's non-verbal behaviour as in our case study, but other rewards can also be considered. For example, in learning scenarios, a robot acting as a tutor could adapt its teaching strategies to maximise the number of children's correct answers in a test.

8.3 Selected Publications

The work presented in this dissertation resulted in several publications. Below we provide a list of the most relevant publications that contributed to the dissemination of this work.

Leite, I., Pereira, A., Mascarenhas, S., Martinho, C., Prada, R., Paiva, A. (2013). The Influence of Empathy in Human-Robot Relations. *International Journal of Human-Computer Studies*, 71(3), Pages 250-260, ISSN 1071-5819, doi 10.1016/j.ijhcs.2012.09.005.

Leite, I., Martinho, C., Paiva, A. (2013). Social Robots for Long-Term Interaction: a Survey, *International Journal of Social Robotics*, ISSN 1875-4791, doi 10.1007/s12369-

013-0178-y, Springer.

Leite, I., Castellano, G., Pereira, A., Martinho, C., Paiva, A. (2012). Long-term Interactions with Empathic Robots: Evaluating Perceived Support in Children. *Proc. of International Conference on Social Robotics*, Chengdu, China, October 2012 [Best Student Paper Award].

Leite, I., Castellano, G., Pereira, A., Martinho, C., Paiva, A. (2012). Modelling empathic behaviour in a robotic game companion for children: an ethnographic study in real-world settings. *Proc. of ACM/IEEE International Conference on Human-Robot Interaction*, Boston, MA, USA, March 2012, pp. 367-374.

Leite, I., Pereira, A., Castellano, G., Mascarenhas, S., Martinho, C., Paiva, A. (2012). Modelling Empathy in Social Robotic Companions, *Advances in User Modeling: Selected papers from UMAP 2011 Workshops (Revised Selected Papers)*, LNCS, No. 7138, Springer Berlin/Heidelberg.

Leite, I. (2011). Using Adaptive Empathic Responses to Improve Long-term Interaction with Social Robots. *Proc. of the International Conference on User Modeling (Doctoral Consortium), Adaptation and Personalization 2011*, Girona, Spain, Springer Berlin/Heidelberg.

Leite, I., Pereira, A., Castellano, G., Mascarenhas, S., Martinho, C., Paiva, A. (2011). Social Robots in Learning Environments: a Case Study of an Empathic Chess Companion. *Proc. of the International Workshop on Personalization Approaches in Learning Environments (PALE)*, Girona, Spain, CEUR Workshop Proceedings (ISSN 1613-0073).

Leite, I., Mascarenhas, S., Pereira, A., Martinho, C., Prada, R., Paiva, A. (2010). "Why Can't We Be Friends?" An Empathic Game Companion for Long-Term Interaction. *Proc. of the 10th International Conference on Intelligent Virtual Agents*, Philadelphia, USA, Springer Berlin/Heidelberg.

Leite, I., Pereira, A., Mascarenhas, S., Castellano, G., Martinho, C., Prada, R.,

Paiva, A. (2010). Closing the Loop: from Affect Recognition to Empathic Interaction. *Proc. of the 3rd International Workshop on Affect Interaction in Natural Environments (AFFINE'10)*, ACM Multimedia 2010, Florence, Italy, ACM.

Leite, I., Castellano, G., Pereira, A., Martinho, C., Paiva, A., McOwan, P. (2009). Designing a Game Companion for Long-Term Social Interaction, *Workshop on Affective Interaction in Natural Environments (AFFINE'09)*, Boston MA, USA, ACM Proceedings.

Leite, I., Pereira, A., Martinho, C., Paiva, A., McOwan, P., Castellano, G. (2009). Towards and Empathic Chess Companion. *Workshop on Empathic Agents*, AAMAS'09, Budapest, Hungary.

Leite, I., Martinho, C., Pereira, A., Paiva, A. (2009). As Time Goes by: Long-term Evaluation of Social Presence in Robotic Companions. *Proc. of the 18th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2009)*, Toyama, Japan, pp.669-674, IEEE Computer Society.

Leite, I., Pereira, A., Martinho, C., Paiva, A. (2008). Are emotional robots more fun to play with? *Proc. of the 17th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2008)*, Munich, Germany, pp.77-82, IEEE Computer Society.

Leite, I., Martinho, C., Paiva, A, Pereira, A. (2008). Social Presence in Long-Term Human-Computer Relationships. *Fourth International Workshop on Human-Computer Conversation*, Bellagio, Italy.

Leite, I., Pereira, A., Martinho, C., Paiva, A. (2008). iCat: an Affective Game Buddy Based on Anticipatory Mechanisms. *Proc of the 7th international conference on Autonomous Agents and Multiagent Systems (AAMAS 2008)*, May 2008, pp. 1253-1256, IFAAMAS.

Appendix A

Final Long-term Study Questionnaire

This appendix contains the questionnaire employed in the final long-term study with the iCat. Before presenting the actual questionnaire that children filled in during the study (section A.2), we present the complete list of questionnaire items both in English and in Portuguese. After the English version of each questionnaire item, we also indicate the measure that this item is related to: Engagement, Friendship (F), Social Presence (SP) or Social Support.

In the final long-term study, the Social Support items (22 to 30) were only present in the questionnaire of the last interaction session. The Social Presence items were also used in the first long-term study reported in section 4.2, the Friendship measures were previously used in the study of section 4.3, and in the short-term interaction study described in chapter 6 the same questions for Engagement and Friendship were employed.

A.1 Questionnaire Items

1. iCat made me participate more in the game. (Engagement)
   O iCat fez-me participar mais no jogo.

2. It was fun playing with iCat. (Engagement)
   Jogar com o iCat foi animado.

3. Playing with iCat caused me real feelings and emotions. (Engagement)
   Senti emoções ao jogar com o iCat.

4. I lost track of time while playing with iCat. (Engagement)
   O tempo passou depressa ao jogar com o iCat.

5. iCat helped me during the game. (F – Help)
   O iCat ajudou-me durante a partida.

6. iCat's comments were useful to me. (F – Help)
   Os conselhos do iCat foram úteis para mim.

7. iCat's comments were helpful when I needed them. (F – Help)
   O iCat ajudou-me quando precisei.

8. I felt that I could play better in the presence of iCat. (F – Self-Validation)
   O iCat fez com que eu sentisse que podia fazer bem as coisas.

9. iCat praised me when I played well. (F – Self-Validation)
   O iCat elogiou-me quando fiz algo bem feito.

10. I noticed iCat. (SP – Co-presence)
    Reparei no iCat.

11. iCat noticed me. (SP – Co-presence)
    O iCat reparou em mim.

12. I remained focused on iCat. (SP – Attentional Allocation)
    Estive atento ao iCat durante o jogo.

13. iCat remained focused on me. (SP – Attentional Allocation)
    O iCat esteve atento a mim durante o jogo.

14. My thoughts were clear to iCat. (SP – Perceived Message Understanding)
    As minhas ideias eram claras para o iCat.

15. iCat's thoughts were clear to me. (SP – Perceived Message Understanding)
    As ideias do iCat eram claras para mim.

16. I could tell how iCat felt. (SP – Perceived Affective Understanding)
    Consegui perceber o que o iCat estava a sentir.

17. iCat could tell how I felt. (SP – Perceived Affective Understanding)
    O iCat percebeu o que eu estava a sentir.

18. iCat was influenced by my mood. (SP – Perceived Affective Interdependence)
    A disposição do iCat mudou a minha disposição.

19. I was influenced by iCat's mood. (SP – Perceived Affective Interdependence)
    Eu mudei a disposição do iCat.

20. iCat's behaviour was tied to mine. (SP – Perceived Behavioural Interdependence)
    O comportamento do iCat mudou o meu.

21. My behaviour was tied to iCat's behaviour. (SP – Perceived Behavioural Interdependence)
    O meu comportamento mudou o comportamento do iCat.

22. iCat comforts me when I am upset. (Social Support)
    O iCat deu-me força quando eu estava chateado/a.

23. iCat cares about me. (Social Support)
    O iCat preocupou-se comigo.

24. iCat gives me good advice. (Social Support)
    O iCat deu-me bons conselhos.

25. iCat accepts me for who I am. (Social Support)
    O iCat aceita quem eu sou.

26. iCat supports my decisions. (Social Support)
    O iCat apoiou as minhas decisões.

27. I can count on iCat. (Social Support)
    Posso confiar no iCat.

28. iCat encourages me. (Social Support)
    O iCat encorajou-me durante o jogo.

29. iCat understands me. (Social Support)
    O iCat consegue perceber-me.

30. iCat praises me when I've done something well. (Social Support)
    O iCat elogia-me quando faço algo bem feito.

31. I would like to play again with iCat... a) everyday; b) once a week; c) once a month; d) never again.
    Gostaria de poder jogar com o iCat... a) todos os dias; b) uma vez por semana ; c) uma vez por mês; d) nunca mais.

A.2 The Questionnaire

Questionário

Nome:______________________________________ **Idade:** _____

Indica o quanto concordas ou discordas com as seguintes afirmações, assinalando o quadrado que melhor expressa a tua opinião.

Exemplo:
Gosto de jogar xadrez.

Discordo totalmente	Discordo um pouco	Não concordo nem discordo	Concordo um pouco	Concordo totalmente
			X	

1) O iCat fez-me participar mais no jogo.

Discordo totalmente	Discordo um pouco	Não concordo nem discordo	Concordo um pouco	Concordo totalmente

2) Jogar com o iCat foi animado.

Discordo totalmente	Discordo um pouco	Não concordo nem discordo	Concordo um pouco	Concordo totalmente

3) Senti emoções ao jogar com o iCat.

Discordo totalmente	Discordo um pouco	Não concordo nem discordo	Concordo um pouco	Concordo totalmente

4) O tempo passou depressa ao jogar com o iCat.

Discordo totalmente	Discordo um pouco	Não concordo nem discordo	Concordo um pouco	Concordo totalmente

5) O iCat ajudou-me durante a partida.

Discordo totalmente	Discordo um pouco	Não concordo nem discordo	Concordo um pouco	Concordo totalmente

6) Os conselhos do iCat foram úteis para mim.

Discordo totalmente	Discordo um pouco	Não concordo nem discordo	Concordo um pouco	Concordo totalmente

7) O iCat ajudou-me quando precisei.

Discordo totalmente	Discordo um pouco	Não concordo nem discordo	Concordo um pouco	Concordo totalmente

8) O iCat fez com que eu sentisse que podia fazer bem as coisas.

Discordo totalmente	Discordo um pouco	Não concordo nem discordo	Concordo um pouco	Concordo totalmente

9) O iCat elogiou-me quando fiz algo bem feito.

Discordo totalmente	Discordo um pouco	Não concordo nem discordo	Concordo um pouco	Concordo totalmente

10) Reparei no iCat.

Discordo totalmente	Discordo um pouco	Não concordo nem discordo	Concordo um pouco	Concordo totalmente

11) O iCat reparou em mim.

Discordo totalmente	Discordo um pouco	Não concordo nem discordo	Concordo um pouco	Concordo totalmente

12) Estive atento ao iCat durante o jogo.

Discordo totalmente	Discordo um pouco	Não concordo nem discordo	Concordo um pouco	Concordo totalmente

13) O iCat esteve atento a mim durante o jogo.

Discordo totalmente	Discordo um pouco	Não concordo nem discordo	Concordo um pouco	Concordo totalmente

14) As minhas ideias eram claras para o iCat.

Discordo totalmente	Discordo um pouco	Não concordo nem discordo	Concordo um pouco	Concordo totalmente

15) As ideias do iCat eram claras para mim.

Discordo totalmente	Discordo um pouco	Não concordo nem discordo	Concordo um pouco	Concordo totalmente

16) Consegui perceber o que o iCat estava a sentir.

Discordo totalmente	Discordo um pouco	Não concordo nem discordo	Concordo um pouco	Concordo totalmente

17) O iCat percebeu o que eu estava a sentir.

Discordo totalmente	Discordo um pouco	Não concordo nem discordo	Concordo um pouco	Concordo totalmente

18) A disposição do iCat mudou a minha disposição.

Discordo totalmente	Discordo um pouco	Não concordo nem discordo	Concordo um pouco	Concordo totalmente

19) Eu mudei a disposição do iCat.

Discordo totalmente	Discordo um pouco	Não concordo nem discordo	Concordo um pouco	Concordo totalmente

20) O comportamento do iCat mudou o meu.

Discordo totalmente	Discordo um pouco	Não concordo nem discordo	Concordo um pouco	Concordo totalmente

21)O meu comportamento mudou o comportamento do iCat.

Discordo totalmente	Discordo um pouco	Não concordo nem discordo	Concordo um pouco	Concordo totalmente

22)O iCat deu-me força quando eu estava chateado/a.

Discordo totalmente	Discordo um pouco	Concordo um pouco	Concordo totalmente

23)O iCat preocupou-se comigo.

Discordo totalmente	Discordo um pouco	Concordo um pouco	Concordo totalmente

24)O iCat deu-me bons conselhos.

Discordo totalmente	Discordo um pouco	Concordo um pouco	Concordo totalmente

25)O iCat aceita quem eu sou.

Discordo totalmente	Discordo um pouco	Concordo um pouco	Concordo totalmente

26)O iCat apoiou as minhas decisões.

Discordo totalmente	Discordo um pouco	Concordo um pouco	Concordo totalmente

27)Posso confiar no iCat.

Discordo totalmente	Discordo um pouco	Concordo um pouco	Concordo totalmente

28)O iCat encorajou-me durante o jogo.

Discordo totalmente	Discordo um pouco	Concordo um pouco	Concordo totalmente

29)O iCat consegue perceber-me.

Discordo totalmente	Discordo um pouco	Concordo um pouco	Concordo totalmente

30)O iCat elogia-me quando faço algo bem feito.

Discordo totalmente	Discordo um pouco	Concordo um pouco	Concordo totalmente

31) Gostaria de poder jogar com o iCat ...

a. Todos os dias ………....☐
b. Uma vez por semana…☐
c. Uma vez por mês……...☐
d. Nunca mais...…………..☐

Muito obrigada pela tua participação!

Appendix B

Personality Questionnaire

In this appendix, we present the Personality Questionnaire that children filled in the final long-term study with the iCat. The complete list of questionnaire items is first presented both in English and in Portuguese. After the English version of each item, we also indicate the related personality trait: Extroversion, Neuroticism, Openness to Experience, Conscientiousness or Agreeableness.

B.1 Questionnaire Items

1. Usually I play... [on my own/with others]. (Extroversion)
   Normalmente brinco... [sozinho/acompanhado].

2. When I go to school... [I'm afraid of something/I'm not afraid of anything]. (Neuroticism)
   Quando vou para a escola... [tenho medo de alguma coisa/não tenho medo de nada].

3. When I see flying birds... [it doesn't impress me/it impresses me a lot]. (Openness)
   Quando vejo pássaros... [não fico impressionado/fico impressionado].

4. I usually do my housework... [willingly/unwillingly]. (Conscientiousness)
   Faço os meus deveres em casa... [com vontade/sem vontade].

5. When my classmate needs something... [I don't notice/I notice]. (Agreeableness)
   Quando um colega precisa de alguma coisa... [não reparo/reparo].

6. When others play... [I join them/I don't join them]. (Extroversion)
   Quando vejo outros a brincar... [junto-me a eles/fico de fora].

7. When something goes wrong... [I stay calm/I quickly get angry]. (Neuroticism)
   Quando algo corre mal... [mantenho-me calmo/fico chateado rapidamente].

8. On a trip, I mostly enjoy... [exploring new things/resting]. (Openness)
   Num passeio, gosto mais de... [explorar coisas novas/descansar].

9. My room is... [messy/tidy]. (Conscientiousness)
   O meu quarto ... [desarrumado/arrumado].

10. When I can help somebody... [I help/I don't help]. (Agreeableness)
    Quando posso ajudar alguém... [Ajudo/Não ajudo].

11. When someone jokes... [it doesn't make me laugh/I laugh with the others]. (Extroversion)
    Quando alguém conta piadas... [não me fazem rir/rio-me com os outros].

12. Usually... [I'm worried/I'm calm]. (Neuroticism)
    Normalmente... [estou preocupado/estou calmo].

13. Learning new, difficult things... [is not pleasant for me/is pleasant for me]. (Openness)
    Aprender coisas novas e diferentes... [não é agradável/é agradável].

14. When I get money from someone... [I save it for later/I buy something immeadiately]. (Conscientiousness)
    Quando recebo dinheiro de alguém... [guardo para depois/compro logo alguma coisa].

15. When I've got a new toy... [I don't lend it/I lend it]. (Agreeableness)
    Quando tenho um brinquedo novo... [não empresto/empresto].

16. With a person I just got to know... [I like talking/I don't like talking]. (Extroversion)
    Com alguém que acabei de conhecer... [gosto de falar/não gosto de falar].

17. When I get hurt... [I quickly stop thinking about it/it bothers me for a long time]. (Neuroticism)
    Quando me aleijo... [rapidamente deixo de pensar nisso/fico incomodado por algum tempo].

18. In my free time I like to... [read something new/play my favourite game]. (Openness)
    Nos tempos livres, gosto de... [ler coisas novas/brincar às minhas brincadeiras favoritas].

19. When I come back from school, at first... [I play/I do my homework]. (Conscientiousness)
    Quando chego da escola, primeiro... [brinco/faço os trabalhos de casa].

20. When someone apologises to me for something... [I forgive quickly/I don't want to forgive]. (Agreeableness)
    Quando alguém me pede desculpa... [aceito imediatamente/não quero aceitar].

21. When I get a gift... [it's OK for me/I'm so happy that I want to jump]. (Extroversion)
    Quando recebo um presente... [acho normal/fico contente e aos saltos].

22. When I usually wait for something... [I'm impatient/I'm patient]. (Neuroticism)
    Quando estou à espera de alguma coisa... [sou impaciente/tenho paciência].

23. On TV I like to watch... [cartoons or movies only/also other programmes]. (Openness)
    Na televisão, gosto de ver... [apenas filmes e desenhos animados/também outros programas].

24. When a task I do starts to get difficult... [I try to finish it/I start doing something different]. (Conscientiousness)
    Quando uma tarefa começa a ficar difícil... [tento terminá-la/começo a fazer outra coisa diferente].

25. When I have broken something... [I don't admit it/I admit it]. (Agreeableness)
    Quando estrago alguma coisa... [não admito que fui eu/admito que fui eu].

26. In a group... [I usually choose a game for us to play/someone else usually chooses a game for us to play]. (Extroversion)
    Num grupo... [normalmente escolho eu as brincadeiras/outra pessoa costuma escolher as brincadeiras].

27. When someone tells a story... [I just listen/I listen and imagine everything]. (Openness)
    Quando alguém conta uma história... [apenas oiço/oiço e imagino tudo].

28. When I act in front of others... [I'm nervous/I'm not nervous]. (Neuroticism)
    Quando tenho de actuar para outras pessoas... [fico nervoso/não fico nervoso].

B.2 The Questionnaire

Que tipo de pessoas és? O que normalmente fazes?

PBPS-C v2, Marta Maćkiewicz, Jan Cieciuch, 2011

Nas páginas seguintes, vais encontrar diferentes situações.
Por exemplo:

„Quando chove ….”

Estas situações são apresentadas em figuras.
Em cada figura, a personagem principal tem um cachecol às riscas.

Este personagem está a fazer duas coisas completamente diferentes, por exemplo:

Jogo computador

Quando chove…

depende

Vejo televisão

Pensa um pouco, o que fazes **mais vezes** nesta situação?
És mais parecido/a com a pessoa de cachecol

nesta figura ou nesta figura

A seguir, assinala com uma cruz a figura que escolheste.

Muito importante!

Em cada situação, podes escolher apenas uma opção
Se por vezes te comportas de uma maneira
e outras vezes de outra, escolhe a opção do centro („DEPENDE”).

MAS por favor evita escolher a caixinha do meio.
Pensa sempre no que fazes **mais vezes**!

Antes de começar, escreve o teu nome e indica se é rapaz ou rapariga!

Nome: ………………………………………………………………

Idade : ………………….

Rapaz/Rapariga

O que normalmente fazes?

1. Normalmente brinco...

sozinho/a

☐ depende

acompanhado

2. Quando vou para a escola...

tenho medo
de alguma coisa

☐ depende

não tenho medo de nada

3. Quando vejo pássaros...

não fico impressionado/a

☐ depende

fico impressionado/a

4. Faço os meus deveres em casa...

com vontade

☐ depende

sem vontade

O que normalmente fazes?

5. Quando um colega precisa de alguma coisa...

Eu não reparo

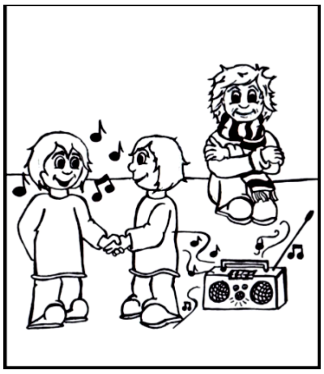
depende

Eu reparo

6. Quando vejo outros a brincar...

junto-me a eles

depende



fico de fora

7. Quando algo corre mal...

mantenho-me calmo/a

depende

fico chateado/a rapidamente

8. Num passeio, gosto mais de...

explorar coisas novas

depende

descansar

O que normalmente fazes?

9. O meu quarto é…

desarrumado

depende

arrumado

10. Quando posso ajudar alguém…

ajudo

depende

não ajudo

11. Quando alguém conta piadas…

não me fazem rir

depende

rio-me com os outros

12. Normalmente…

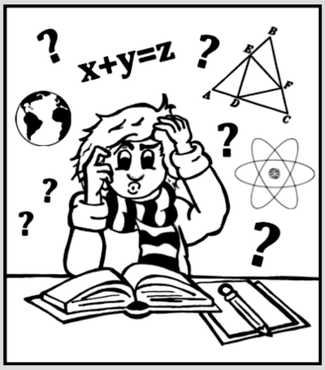
estou preocupado/a

depende

estou calmo/a

O que normalmente fazes?

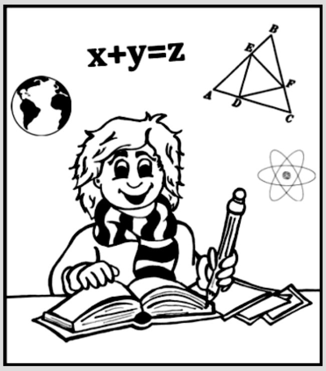
13. Aprender coisas novas e diferentes...



nao é agradável

☐ ☐ ☐

depende



é agradável

14. Quando recebo dinheiro de alguém...

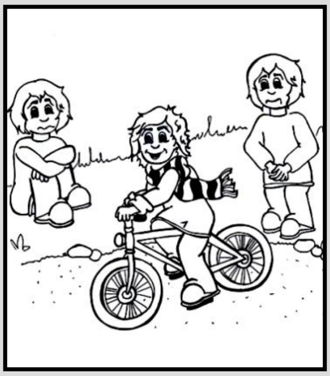
guardo para depois

☐ ☐ ☐

depende

compro logo alguma coisa

15. Quando tenho um brinquedo novo...



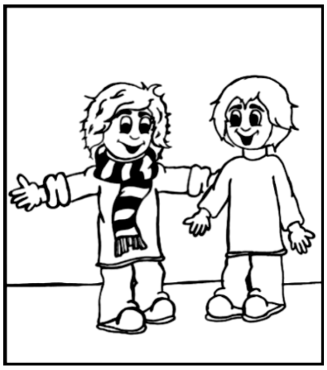
não empresto

☐ ☐ ☐

depende

empresto

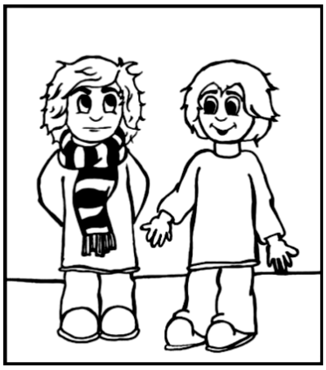
16. Com alguém que acabei de conhecer...



gosto de falar

☐ ☐ ☐

depende



não falo muito

O que normalmente fazes?

17. Quando me aleijo…

rapidamente deixo de pensar nisso

depende

fico incomodado/a por algum tempo

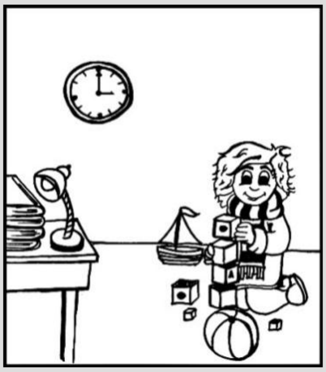
18. Nos tempos livres, gosto de…

ler coisas novas

depende

brincar às minhas brincadeiras favoritas

19. Quando chego da escola, primeiro…

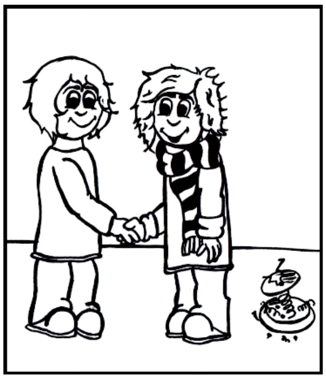


brinco

depende

faço os trabalhos de casa

20. Quando alguém me pede desculpa…



aceito imediatamente

depende

não quero aceitar

O que normalmente fazes?

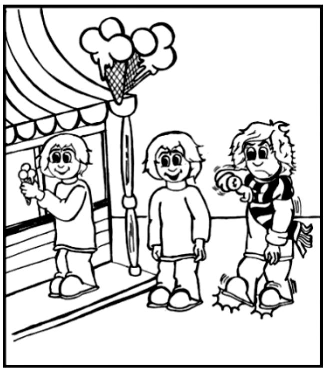
21. Quando recebo um presente…

acho normal

depende

fico contente
e aos saltos

22. Quando estou à espera de alguma coisa…



sou impaciente

depende

tenho paciência

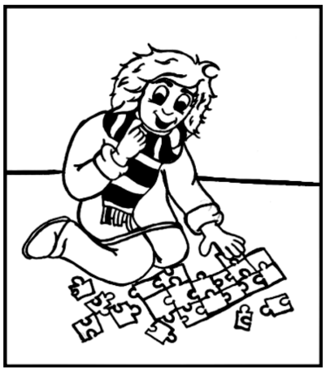
23. Na televisão, gosto de ver…

apenas filmes
e desenhos animados

depende

também outros
programas

24. Quando uma tarefa começa a ficar difícil…



tento terminá-la

depende

começo a fazer
outra coisa diferente

O que normalmente fazes?

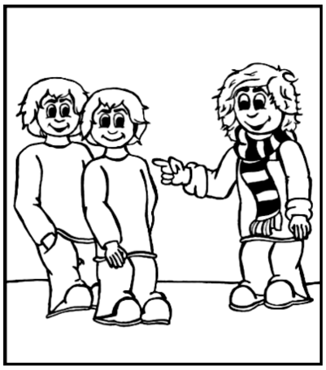
25. Quando estrago alguma coisa...

não admito que fui eu

depende

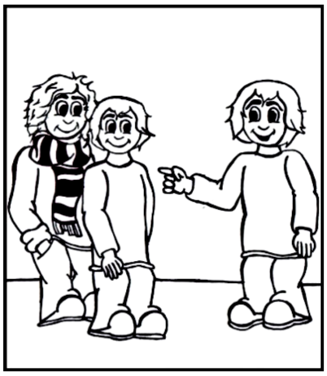
admito que fui eu

26. Num grupo...



normalmente escolho eu as brincadeiras

depende



outra pessoa costuma escolher as brincadeiras

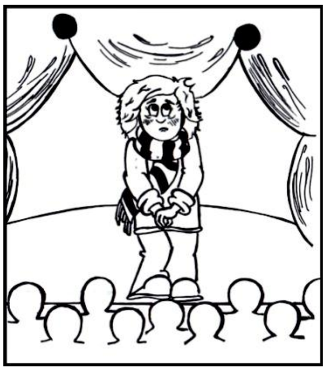
27. Quando alguém conta uma história...

apenas oiço

depende

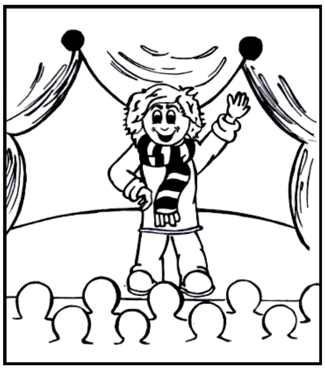
oiço e imagino tudo

28. Quando tenho de actuar para outras pessoas...



fico nervoso/a

depende



não fico nervoso/a

Bibliography

[1] R.G. Adams, R. Blieszner, and B. De Vries. Definitions of friendship in the third age: Age, gender, and study location effects. *Journal of Aging Studies*, 14:117–133, 2000.

[2] P. Aghaei Pour, M. Hussain, O. AlZoubi, S. D'Mello, and R. Calvo. The impact of system feedback on learners affective and physiological states. In *Intelligent Tutoring Systems*, pages 264–273. Springer, 2010.

[3] C. Anderson and D Keltner. The role of empathy in the formation and maintenance of social bonds. *Behavioural and Brain Sciences*, 25(1):21–22, 2002.

[4] J. Archer. Why do people love their pets? *Evolution and Human Behavior*, 18(4):237 – 259, 1997.

[5] D.N. Aspy. *Toward a Technology for Humanizing Education.* Research Press Company, Champaign, Illinois, 1974.

[6] P. Auer, N. Cesa-Bianchi, and P. Fischer. Finite-time analysis of the multiarmed bandit problem. *Machine learning*, 47(2):235–256, 2002.

[7] Peter Auer, Nicolò Cesa-Bianchi, Yoav Freund, and Robert E. Schapire. The nonstochastic multiarmed bandit problem. *SIAM Journal on Computing*, 32(1):48–77, January 2003.

[8] M. R. Banks, L. M. Willoughby, and W. A. Banks. Animal-assisted therapy and loneliness in nursing homes: Use of robotic versus living dogs. *Journal of the American Medical Directors Association*, 9(3):173 – 177, 2008.

[9] S.B. Barker and K.S. Dawson. The effects of animal-assisted therapy on anxiety ratings of hospitalized psychiatric patients. *Psychiatric Services*, 49(6):797, 1998.

[10] C. Bartneck, J. Reichenbach, and A. Breemen. In your face, robot! the influence of a characters embodiment on how users perceive its emotional expressions. In *Proceedings of the Design and Emotion*, 2004.

[11] T. N. Beran and A. Ramirez-Serrano. Can children have a relationship with a robot? In *Human-Robot Personal Relationships*, volume 59 of *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, pages 49–56. Springer Berlin Heidelberg, 2011.

[12] E. Berscheid. Interpersonal relationships. *Annual Review of Psychology*, 45(1):79–129, 1994.

[13] E. Berscheid and H.T. Reis. Attraction and close relationships. *The handbook of social psychology*, 2:193–281, 1998.

[14] CindyL. Bethel and RobinR. Murphy. Review of human studies methods in hri and recommendations. *International Journal of Social Robotics*, 2(4):347–359, 2010.

[15] S. Bhamjee, F. Griffiths, and J. Palmer. Children's perception and interpretation of robots and robot behaviour. In *Human-Robot Personal Relationships*, volume 59 of *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, pages 42–48. Springer Berlin Heidelberg, 2011.

[16] T. Bickmore and D. Schulman. Practical approaches to comforting users with relational agents. In *CHI'07 extended abstracts on Human factors in computing systems*, pages 2291–2296. ACM, 2007.

[17] T. Bickmore, D. Schulman, and L. Yin. Maintaining engagement in long-term interventions with relational agents. *Applied Artificial Intelligence*, 24(6):648–666, 2010.

[18] T.W. Bickmore. *Relational agents: Effecting change through human-computer relationships*. PhD thesis, Massachusetts Institute of Technology PhD Dissertation, 2003.

[19] T.W. Bickmore and R.W. Picard. Establishing and maintaining long-term human-computer relationships. *ACM Transactions on Computer-Human Interaction*, 12(2):327, 2005.

[20] F. Biocca. The cyborg's dilemma: Embodiment in virtual environments. In *Cognitive Technology, 1997.'Humanizing the Information Age'. Proceedings., Second International Conference on*, pages 12–26. IEEE, 1997.

[21] H. Boukricha and I. Wachsmuth. Empathy-based emotional alignment for a virtual human: A three-step approach. *KI - Knstliche Intelligenz*, pages 1–10, 2011.

[22] S. Brave, C. Nass, and K. Hutchinson. Computers that care: investigating the effects of orientation of emotion exhibited by an embodied computer agent. *International Journal of Human-Computer Studies*, 62(2):161–178, 2005.

[23] C. Breazeal. *Designing sociable robots*. The MIT Press, 2004.

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

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

[26] G. Castellano, R. Aylett, K. Dautenhahn, A. Paiva, P.W. McOwan, and S. Ho. Long-term affect sensitive and socially interactive companions. In *Proceedings of the 4th International Workshop on Human-Computer Conversation*, 2008.

[27] G. Castellano, I. Leite, A. Pereira, C. Martinho, A. Paiva, and P. W. McOwan. Inter-ACT: An affective and contextually rich multimodal video corpus for studying interaction with robots. In *Proceedings of the ACM International Conference on Multimedia*, pages 1031–1034. ACM, 2010.

[28] G. Castellano, I. Leite, A. Pereira, C. Martinho, A. Paiva, and P.W. McOwan. It's all in the game: Towards an affect sensitive and context aware game companion. In *Affective Computing and Intelligent Interaction, 2009. ACII 2009. 3rd International Conference on*, pages 1 –8, sept. 2009.

[29] S.J. Ceci and M. Bruck. Suggestibility of the child witness: A historical review and synthesis. *Psychological bulletin*, 113(3):403, 1993.

[30] S. Chaffar and C. Frasson. Using emotional coping strategies in intelligent tutoring systems. In *Intelligent Tutoring Systems*, pages 285–287. Springer, 2010.

[31] T. Chesney and S. Lawson. The illusion of love: Does a virtual pet provide the same companionship as a real one? *Interaction Studies*, 8(2):337–342, 2007.

[32] M. Chi, K. VanLehn, D. Litman, and P. W. Jordan. Inducing effective pedagogical strategies using learning context features. In Paul De Bra, Alfred Kobsa, and David N. Chin, editors, *UMAP*, volume 6075 of *Lecture Notes in Computer Science*, pages 147–158. Springer, 2010.

[33] M.T.H. Chi. Quantifying qualitative analyses of verbal data: A practical guide. *Journal of the Learning sciences*, 6(3):271–315, 1997.

[34] B. Cooper, P. Brna, and A. Martins. Effective affective in intelligent systems – building on evidence of empathy in teaching and learning. In A. Paiva, editor, *Affective Interactions*, volume 1814 of *Lecture Notes in Computer Science*, pages 21–34. Springer Berlin / Heidelberg, 2000.

[35] R. Cowie. Companionship is an emotional business. In Y. Wilks, editor, *Close Engagements with Artificial Companions. Key social, psychological, ethical and design issues*, pages 3–10. John Benjamins Publishing Company, 2010.

[36] H. Cramer, J. Goddijn, B. Wielinga, and V. Evers. Effects of (in)accurate empathy and situational valence on attitudes towards robots. In *HRI '10: Proceeding of the 5th ACM/IEEE international conference on Human-robot interaction*, pages 141–142, New York, NY, USA, 2010. ACM.

[37] C. Croarkin and P. Tobias. Nist/sematech e-handbook of statistical methods. *National Institute of Standards and Technology, URL http://www. itl. nist. gov/div898/handbook*, 2006.

[38] C.E. Cutrona, J.A. Suhr, and R. MacFarlane. Interpersonal transactions and the psychological sense of support. *Personal relationships and social support*, pages 30–45, 1990.

[39] K. Dautenhahn. Robots we like to live with?! a developmental perspective on a personalized, life-long robot companion. In *Robot and Human Interactive Communication, 2004. ROMAN 2004. 13th IEEE International Workshop on*, pages 17–22. IEEE, 2004.

[40] K. Dautenhahn and I. Werry. Towards interactive robots in autism therapy: Background, motivation and challenges. *Pragmatics & Cognition*, 12(1):1–35, 2004.

[41] M. H. Davis. The effects of dispositional empathy on emotional reactions and helping: A multidimensional approach. *Journal of Personality*, 51(2):167–184, 1983.

[42] F. De Vignemont and T. Singer. The empathic brain: how, when and why? *Trends in Cognitive Sciences*, 10(10):435–441, 2006.

[43] S. A. Denham, J. Mitchell-Copeland, K. Strandberg, S. Auerbach, and K. Blair. Parental contributions to preschoolers' emotional competence: Direct and indirect effects. *Motivation and Emotion*, 21:65–86, 1997.

[44] C.F. DiSalvo, F. Gemperle, J. Forlizzi, and S. Kiesler. All robots are not created equal: the design and perception of humanoid robot heads. In *Proceedings of the 4th conference on Designing interactive systems: processes, practices, methods, and techniques*, pages 321–326. ACM, 2002.

[45] S. D'Mello, S. Craig, K. Fike, and A. Graesser. Responding to learners' cognitive-affective states with supportive and shakeup dialogues. *Human-Computer Interaction. Ambient, Ubiquitous and Intelligent Interaction*, pages 595–604, 2009.

[46] S. Duck. *Human relationships*. Sage Publications, Inc, 1998.

[47] N. Eisenberg and R. A. Fabes. Empathy: Conceptualization, measurement, and relation to prosocial behavior. *Motivation and Emotion*, 14:131–149, 1990.

[48] N. Eisenberg, S. Losoya, and I.K. Guthrie. Social cognition and prosocial development. *The development of social cognition*, pages 329–363, 1997.

[49] N. Endenburg, H. Hart, and J. Bouw. Motives for acquiring companion animals. *Journal of Economic Psychology*, 15(1):191–206, 1994.

[50] A. Feingold. Good-looking people are not what we think. *Psychological Bulletin*, 111(2):304–341, March 1992.

[51] Y. Fernaeus, M. Håkansson, M. Jacobsson, and S. Ljungblad. How do you play with a robotic toy animal?: a long-term study of pleo. In *Proceedings of the 9th International Conference on Interaction Design and Children*, pages 39–48. ACM, 2010.

[52] N. D. Feshbach. Studies of empathic behavior in children. *Progress in experimental personality research*, 8:1–47, 1978.

[53] B.J. Fogg and C. Nass. Silicon sycophants: the effects of computers that flatter. *International Journal of Human-Computer Studies*, 46(5):551 – 561, 1997.

[54] D. François, S. Powell, and K. Dautenhahn. A long-term study of children with autism playing with a robotic pet: Taking inspirations from non-directive play therapy to encourage children's proactivity and initiative-taking. *Interaction Studies*, 10(3):324–373, 2009.

[55] B. Friedman, P. H. Kahn, and J. Hagman. Hardware companions?: what online aibo discussion forums reveal about the human-robotic relationship. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '03, pages 273–280, New York, NY, USA, 2003. ACM.

[56] T. Ganster, S. Eimler, A. von der Pütten, L. Hoffmann, and N.C. Krämer. Methodological considerations for long-term experience with robots and agents. *Proceedings of EMCSR 2010*, 2010.

[57] L. Giusti and P. Marti. Interpretative dynamics in human robot interaction. In *Robot and Human Interactive Communication, 2006. ROMAN 2006. The 15th IEEE International Symposium on*, pages 111–116. IEEE, 2006.

[58] R. Gockley, A. Bruce, J. Forlizzi, M. Michalowski, A. Mundell, S. Rosenthal, B. Sellner, R. Simmons, K. Snipes, A.C. Schultz, and Jue Wang. Designing robots for long-term social interaction. In *Intelligent Robots and Systems, 2005. (IROS 2005). 2005 IEEE/RSJ International Conference on*, pages 1338 – 1343, aug. 2005.

[59] D.J. Goldsmith and L.A. Baxter. Constituting Relationships in Talk – A Taxonomy of Speech Events in Social and Personal Relationships. *Human Communication Research*, 23(1):87–114, 1996.

[60] A.T. Gordon. *Assessing Social Support in Children: Development and Initial Validation of the Social Support Questionnaire for Children*. PhD thesis, Faculty of Louisiana State University and Agriculture and Mechanical College, University of Houston, 2011.

[61] K.F Halamandaris and K.G Power. Individual differences, social support and coping with the examination stress: a study of the psychosocial and academic adjustment of first year home students. *Personality and Individual Differences*, 26(4):665 – 685, 1999.

[62] C. Harms and F. Biocca. Internal consistency and reliability of the networked minds - measure of social presence. In *Seventh Annual International Workshop: Presence 2004*, 2004.

[63] L.A. Hart. Dogs as human companions: a review of the relationship. *The domestic dog: its evolution, behaviour, and interactions with people*, page 161, 1995.

[64] K. Hayashi, M. Shiomi, T. Kanda, and N. Hagita. Who is appropriate? a robot, human and mascot perform three troublesome tasks. In *RO-MAN, 2010 IEEE*, pages 348 –354, sept. 2010.

[65] Marcel Heerink, Ben J. A. Kröse, Bob J. Wielinga, and Vanessa Evers. Enjoyment intention to use and actual use of a conversational robot by elderly people. In Terry Fong, Kerstin Dautenhahn, Matthias Scheutz, and Yiannis Demiris, editors, *HRI*, pages 113–120. ACM, 2008.

[66] C. Heeter. Being there: The subjective experience of presence. *Presence: Teleoperators and virtual environments*, 1(2):262–271, 1992.

[67] F. Hegel, T. Spexard, T. Vogt, G. Horstmann, and B. Wrede. Playing a different imitation game: Interaction with an Empathic Android Robot. In *Proc. 2006 IEEE-RAS International Conference on Humanoid Robots (Humanoids06)*, pages 56–61, 2006.

[68] W. C. Ho, K. Dautenhahn, M. Y. Lim, P.A. Vargas, R. Aylett, and S. Enz. 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*, pages 277 –284, 27 2009-oct. 2 2009.

[69] M.L. Hoffman. Development of prosocial motivation: Empathy and guilt. *The development of prosocial behavior*, 281:313, 1982.

[70] M.L. Hoffman. *Empathy and moral development: Implications for caring and justice*. Cambridge Univ Press, 2001.

[71] K. Hone. Empathic agents to reduce user frustration: The effects of varying agent characteristics. *Interacting with Computers*, 18(2):227–245, 2006.

[72] A.O. Horvath and L.S. Greenberg. Development and validation of the working alliance inventory. *Journal of counseling psychology*, 36(2):223, 1989.

[73] C. Houston-Price and S. Nakai. Distinguishing novelty and familiarity effects in infant preference procedures. *Infant and Child Development*, 13(4):341–348, 2004.

[74] E. Hyun, H. Yoon, and S. Son. Relationships between user experiences and children's perceptions of the education robot. In *Proceeding of the 5th ACM/IEEE international conference on Human-robot interaction*, pages 199–200. ACM, 2010.

[75] M. Jacobsson. Play, belief and stories about robots: A case study of a pleo blogging community. In *Robot and Human Interactive Communication, 2009. RO-MAN 2009. The 18th IEEE International Symposium on*, pages 232–237. IEEE, 2009.

[76] W.L. Johnson, J.W. Rickel, and J.C. Lester. Animated pedagogical agents: Face-to-face interaction in interactive learning environments. *International Journal of Artificial intelligence in education*, 11(1):47–78, 2000.

[77] P.H. Kahn, H. Ishiguro, B. Friedman, T. Kanda, N.G. Freier, R.L. Severson, and J. Miller. What is a human? toward psychological benchmarks in the field of humanrobot interaction. *Interaction Studies*, 8(3):363–390, 2007.

[78] T. Kanda, T. Hirano, D. Eaton, and H. Ishiguro. Interactive robots as social partners and peer tutors for children: A field trial. *Human-Computer Interaction*, 19(1):61–84, 2004.

[79] T. Kanda, R. Sato, N. Saiwaki, and H. Ishiguro. A two-month field trial in an elementary school for long-term human–robot interaction. *Robotics, IEEE Transactions on*, 23(5):962–971, 2007.

[80] T. Kanda, M. Shiomi, Z. Miyashita, H. Ishiguro, and N. Hagita. A communication robot in a shopping mall. *Transactions on Robotics*, 26(5):897–913, October 2010.

[81] A. Kapoor, W. Burleson, and R. W. Picard. Automatic prediction of frustration. *International Journal of Human-Computer Studies*, 65(8):724–736, 2007.

[82] E. Karapanos, J. Zimmerman, J. Forlizzi, and J.B. Martens. User experience over time: an initial framework. In *Proceedings of the 27th international conference on Human factors in computing systems*, pages 729–738. ACM, 2009.

[83] A. Katcher and G. Wilkins. Dialogue with animals: Its nature and culture. *The biophilia hypothesis*, pages 173–200, 1993.

[84] C.D. Kidd and C. Breazeal. Robots at home: Understanding long-term human-robot interaction. In *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*, pages 3230–3235. IEEE, 2008.

[85] M. Kipp. Anvil: The video annotation research tool. *Handbook of Corpus Phonology. Oxford University Press, Oxford (to appear, 2011)*, 2007.

[86] R. Kirby, J. Forlizzi, and R. Simmons. Affective social robots. *Robotics and Autonomous Systems*, 58(3):322 – 332, 2010. Towards Autonomous Robotic Systems 2009: Intelligent, Autonomous Robotics in the UK.

[87] T. Klamer, S. Ben Allouch, and D. Heylen. "adventures of harvey" - use, acceptance of and relationship building with a social robot in a domestic environment. In *Human-Robot Personal Relationships*, volume 59 of *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, pages 74–82. Springer Berlin Heidelberg, 2011.

[88] J. Klein, Y. Moon, and R.W. Picard. This computer responds to user frustration::: Theory, design, and results. *Interacting with computers*, 14(2):119–140, 2002.

[89] K.L. Koay, D.S. Syrdal, M.L. Walters, and K. Dautenhahn. Living with robots: Investigating the habituation effect in participants' preferences during a longitudinal human-robot interaction study. In *Robot and Human interactive Communication, 2007. RO-MAN 2007. The 16th IEEE International Symposium on*, pages 564–569. IEEE, 2007.

[90] H. Kozima, M.P. Michalowski, and C. Nakagawa. A playful robot for research, therapy, and entertainment. *Int J Soc Robot*, 1:3–18, 2009.

[91] N.C. Krämer and G. Bente. Personalizing e-learning. the social effects of pedagogical agents. *Educational Psychology Review*, 22(1):71–87, 2010.

[92] D.L. Krebs. Altruism: An examination of the concept and a review of the literature. *Psychological Bulletin*, 73(4):258–302, 1970.

[93] K.M. Lee and C. Nass. Social-psychological origins of feelings of presence: Creating social presence with machine-generated voices. *Media Psychology*, 7(1):31–45, 2005.

[94] I. Leite, C. Martinho, A. Pereira, and A. Paiva. icat: an affective game buddy based on anticipatory mechanisms. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Volume 3*, pages 1229–1232. International Foundation for Autonomous Agents and Multiagent Systems, 2008.

[95] I. Leite, C. Martinho, A. Pereira, and A. Paiva. As time goes by: Long-term evaluation of social presence in robotic companions. In *Robot and Human Interactive Communication, 2009. RO-MAN 2009. The 18th IEEE International Symposium on*, pages 669–674. IEEE, 2009.

[96] I. Leite, A. Pereira, C. Martinho, and A. Paiva. Are emotional robots more fun to play with? In *Robot and Human Interactive Communication, 2008. RO-MAN 2008. The 17th IEEE International Symposium on*, pages 77 –82, aug. 2008.

[97] G. Levinger. Development and change. In *Close relationships*, pages 315–359. New York: WH Freeman, 1983.

[98] B. M. Levinson. Human/companion animal therapy. *Journal of Contemporary Psychotherapy*, 14:131–144, 1984. 10.1007/BF00946311.

[99] M.Y. Lim, R. Aylett, W.C. Ho, J. Dias, and P.A. Vargas. Human-like memory retrieval mechanisms for social companions. In *Proc. of 10th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2011)*, pages 1117–1118, 2011.

[100] M. Maćkiewicz and J. Cieciuch. Picture based personality survey for children - new instrument for measuring big five in childhood. In *Proceedings of the 11th European Conf. on Psychological Assessment.*, 2011.

[101] C. Martinho and A. Paiva. Using anticipation to create believable behaviour. In *Proceedings of the 21st national conference on Artificial intelligence - Volume 1*, pages 175–180. AAAI Press, 2006.

[102] M.J. Matarić, J. Eriksson, D.J. Feil-Seifer, and C.J. Winstein. Socially assistive robotics for post-stroke rehabilitation. *Journal of NeuroEngineering and Rehabilitation*, 4(1):5, 2007.

[103] R.R. McCrae and P.T. Costa Jr. Personality trait structure as a human universal. *American Psychologist; American Psychologist*, 52(5):509, 1997.

[104] S. W. McQuiggan and J. C. Lester. Modeling and evaluating empathy in embodied companion agents. *International Journal of Human-Computer Studies*, 65(4):348 – 360, 2007. Evaluating affective interactions.

[105] M. J. Mendelson and F. E. Aboud. Measuring friendship quality in late adolescents and young adults: Mcgill friendship questionnaires. *Canadian Journal of Behavioural Science*, 31(1):130–132, Apr 1999.

[106] F. Michaud, T. Salter, A. Duquette, H. Mercier, M. Lauria, H. Larouche, and F. Larose. Assistive technologies and child-robot interaction. In *AAAI Spring Symposium on Multidisciplinary Collaboration for Socially Assistive Robotics*, 2007.

[107] M. Mori. The uncanny valley. *Energy*, 7(4):33–35, 1970.

[108] H. Nguyen and J. Masthoff. Designing empathic computers: the effect of multimodal empathic feedback using animated agent. In *Proceedings of the 4th International Conference on Persuasive Technology*, Persuasive '09, pages 7:1–7:9, New York, NY, USA, 2009. ACM.

[109] M. Ochs, C. Pelachaud, and D. Sadek. An empathic virtual dialog agent to improve human-machine interaction. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Volume 1*, pages 89–96. International Foundation for Autonomous Agents and Multiagent Systems, 2008.

[110] T. Orbuch and S. Sprecher. Attraction and interpersonal relationships. In John Delamater, editor, *Handbook of Social Psychology*, Handbooks of Sociology and Social Research, pages 339–362. Springer US, 2006.

[111] A. Ortony, G.L. Clore, and A. Collins. *The cognitive structure of emotions.* Cambridge Univ Press, 1990.

[112] A. Paiva, J. Dias, D. Sobral, R. Aylett, P. Sobreperez, S. Woods, C. Zoll, and L. Hall. Caring for agents and agents that care: Building empathic relations with synthetic agents. In *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems - Volume 1*, AAMAS '04, pages 194–201, Washington, DC, USA, 2004. IEEE Computer Society.

[113] M.L. Patterson. Functions of non-verbal behaviour in social interaction. *Handbook of language and social psychology*, pages 101–120, 1990.

[114] H. Prendinger, C. Becker-Asano, and M. Ishizuka. A study in users' physiological response to an empathic interface agent. *I. J. Humanoid Robotics*, 3(3):371–391, 2006.

[115] H. Prendinger and M. Ishizuka. The empathic companion: a character-based interface that addresses users' affective states. *Applied Artificial Intelligence*, 19(3):267–285, 2005.

[116] S. D. Preston and F. B. M. de Waal. Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, 25(01):1–20, 2002.

[117] S. Pulman. Conditions for companionhood. In Y. Wilks, editor, *Close Engagements with Artificial Companions. Key social, psychological, ethical and design issues*, pages 29–34. John Benjamins Publishing Company, 2010.

[118] J. C. Read and S. MacFarlane. Using the fun toolkit and other survey methods to gather opinions in child computer interaction. In *Proceedings of the 2006 conference on Interaction design and children*, IDC '06, pages 81–88, New York, NY, USA, 2006. ACM.

[119] J.C. Read and S. MacFarlane. Using the fun toolkit and other survey methods to gather opinions in child computer interaction. In *Proceedings of the 2006 conference on Interaction design and children*, pages 81–88. ACM, 2006.

[120] B. Reeves and C. Nass. *The media equation: how people treat computers, television, and new media like real people and places.* Cambridge University Press, New York, NY, USA, 1996.

[121] G. M. Reevy and C. Maslach. Use of social support: Gender and personality differences. *Sex Roles*, 44:437–459, 2001. 10.1023/A:1011930128829.

[122] E. Rich. Users are individuals: individualizing user models. *International Journal of Man-Machine Studies*, 18(3):199–214, 1983.

[123] L. D. Riek, P. C. Paul, and P. Robinson. When my robot smiles at me: Enabling human-robot rapport via real-time head gesture mimicry. *Journal on Multimodal User Interfaces*, 3(1-2):99–108, 2010.

[124] J. Robison, S. McQuiggan, and J. Lester. Evaluating the consequences of affective feedback in intelligent tutoring systems. In *Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference on*, pages 1–6. IEEE, 2009.

[125] S.H. Rodrigues, S. Mascarenhas, J. Dias, and A. Paiva. "i can feel it too!": Emergent empathic reactions between synthetic characters. In *Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference on*, pages 1–7. IEEE, 2009.

[126] C.R. Rogers. Empathic: An unappreciated way of being. *The counseling psychologist*, 5(2):2–10, 1975.

[127] S. Rosenthal, J. Biswas, and M. Veloso. An effective personal mobile robot agent through symbiotic human-robot interaction. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1-Volume 1*, pages 915–922. International Foundation for Autonomous Agents and Multiagent Systems, 2010.

[128] A.M. Sabelli, T. Kanda, and N. Hagita. A conversational robot in an elderly care center: an ethnographic study. In *Proceedings of the 6th international conference on Human-robot interaction*, pages 37–44. ACM, 2011.

[129] M. Saerbeck, T. Schut, C. Bartneck, and M.D. Janse. Expressive robots in education: varying the degree of social supportive behavior of a robotic tutor. In *Proceedings of the 28th international conference on Human factors in computing systems*, pages 1613–1622. ACM, 2010.

[130] T. Salter, K. Dautenhahn, and R. Bockhorst. 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*, pages 563 – 568, sept. 2004.

[131] K.R. Scherer. Psychological models of emotion. *The neuropsychology of emotion*, 2000.

[132] P. Schermerhorn, M. Scheutz, and C.R. Crowell. Robot social presence and gender: Do females view robots differently than males? In *Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction*, pages 263–270. ACM, 2008.

[133] Matthew H Scullin and Stephen J Ceci. A suggestibility scale for children. *Personality and Individual Differences*, 30(5):843 – 856, 2001.

[134] K. Severinson-Eklundh, A. Green, and H. Hüttenrauch. Social and collaborative aspects of interaction with a service robot. *Robotics and Autonomous Systems*, 42(3-4):223–234, 2003.

[135] C.L. Sidner, C.D. Kidd, C. Lee, and N. Lesh. Where to look: a study of human-robot engagement. In *Proceedings of the 9th international conference on Intelligent user interfaces*, pages 78–84. ACM, 2004.

[136] L. K. Souza. *Amizade em adultos: adaptação e validação dos questionários McGill e um estudo de diferenças de génere.* PhD thesis, Universidade Federal do Rio Grande do Sul, 2006.

[137] L. Stafford, M. Dainton, and S. Haas. Measuring routine and strategic relational maintenance: Scale revision, sex versus gender roles, and the prediction of relational characteristics. *Communication Monographs*, 67(3):306–323, 2000.

[138] Aaron Steinfeld, Terrence Fong, David Kaber, Michael Lewis, Jean Scholtz, Alan Schultz, and Michael Goodrich. Common metrics for human-robot interaction. In

Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction, HRI '06, pages 33–40, New York, NY, USA, 2006. ACM.

[139] E. Stotland. Exploratory investigations of empathy. *Advances in Experimental Social Psychology*, 4:271–314, 1969.

[140] K. Stubbs, D. Bernstein, K. Crowley, and I. Nourbakhsh. Long-term human-robot interaction: The personal exploration rover and museum docents. In *Proceeding of the 2005 conference on Artificial Intelligence in Education: Supporting Learning through Intelligent and Socially Informed Technology*, pages 621–628. IOS Press, 2005.

[141] J. Sung, H.I. Christensen, and R.E. Grinter. Robots in the wild: understanding long-term use. In *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction*, pages 45–52. ACM, 2009.

[142] J. Sung, R. E. Grinter, and H. I. Christensen. Domestic Robot Ecology. *International Journal of Social Robotics*, 2(4):417–429, July 2010.

[143] L. Takayama, W. Ju, and C. Nass. Beyond dirty, dangerous and dull: what everyday people think robots should do. In *Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction*, HRI '08, pages 25–32, New York, NY, USA, 2008. ACM.

[144] F. Tanaka, A. Cicourel, and J.R. Movellan. Socialization between toddlers and robots at an early childhood education center. *Proceedings of the National Academy of Sciences*, 104(46):17954, 2007.

[145] R.E. Thayer. *The biopsychology of mood and arousal.* Oxford University Press, USA, 1990.

[146] S. Turkle. In good company? In Y. Wilks, editor, *Close Engagements with Artificial Companions. Key social, psychological, ethical and design issues*, pages 3–10. John Benjamins Publishing Company, 2010.

[147] S. Turkle, W. Taggart, C.D. Kidd, and O. Dasté. Relational artifacts with children and elders: the complexities of cybercompanionship. *Connection Science*, 18(4):347–362, 2006.

[148] A. van Breemen, X. Yan, and B. Meerbeek. icat: an animated user-interface robot with personality. In *Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems*, pages 143–144. ACM, 2005.

[149] J.M. Vigil and D.C. Geary. A preliminary investigation of family coping styles and psychological well-being among adolescent survivors of hurricane katrina. *Journal of Family Psychology*, 22(1):176, 2008.

[150] K. Wada and T. Shibata. Robot therapy in a care house-its sociopsychological and physiological effects on the residents. In *Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on*, pages 3966–3971. IEEE, 2006.

[151] K. Wada and T. Shibata. Robot therapy in a care house-results of case studies. In *Robot and Human Interactive Communication, 2006. ROMAN 2006. The 15th IEEE International Symposium on*, pages 581–586. IEEE, 2006.

[152] K. Wada and T. Shibata. Living with seal robots - its sociopsychological and physiological influences on the elderly at a care house. *Robotics, IEEE Transactions on*, 23(5):972–980, 2007.

[153] S. Wellek. *Testing statistical hypotheses of equivalence*. CRC Press, 2003.

[154] Y. Wilks and M. Peltu. Summary and discussion of the issues. In Y. Wilks, editor, *Close Engagements with Artificial Companions. Key social, psychological, ethical and design issues*, pages 3–10. John Benjamins Publishing Company, 2010.

[155] T.A. WILLS. Social support and interpersonal relationships. *Prosocial behavior*, 12:265, 1991.

[156] Z. Zeng, M Pantic, G. I. Roisman, and T. S Huang. A survey of affect recognition methods: Audio, visual and spontaneous expressions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(1):39–58, 2009.