

### **Tilburg University**

### Fun & face

Shahid, S.

Publication date: 2012

Document Version Publisher's PDF, also known as Version of record

Link to publication in Tilburg University Research Portal

Citation for published version (APA): Shahid, S. (2012). Fun & face: Exploring non-verbal expressions of emotion during playful interactions. TiCC PhD Séries No. 21.

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
  You may not further distribute the material or use it for any profit-making activity or commercial gain
  You may freely distribute the URL identifying the publication in the public portal

**Take down policy**If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 12. May. 2021

### **FUN & FACE**

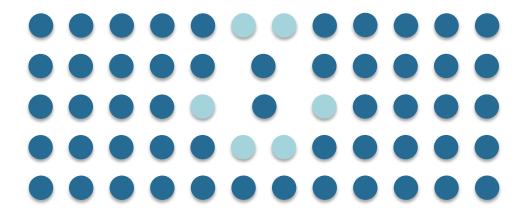
## EXPLORING NON-VERBAL EXPRESSIONS OF EMOTION DURING PLAYFUL INTERACTIONS





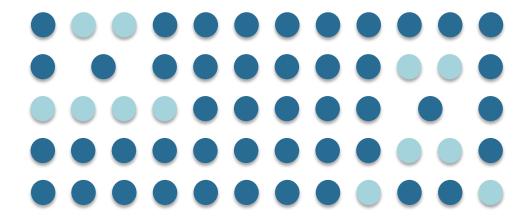


Suleman Shahid



Fun & Face: Exploring non-verbal expressions of emotion during playful interactions

Suleman Shahid



# Fun & Face: Exploring non-verbal expressions of emotion during playful interactions

Suleman Shahid, PhD Thesis Tilburg University, 2012





SIKS-dissertation Series No. 2012-13

The research reported in this thesis has been carried out under the auspices of SIKS, the Dutch Research School for Information and Knowledge Systems.

ISBN / EAN: 978-94-6203-036-7 Print: CPI Wöhrmann print service

Cover design: Lennard van de Laar & Suleman Shahid

### © 2012 Suleman Shahid

No part of this thesis may be reproduced, stored in a retrieval system or transmitted in any form or by any means. Without written permission of the author or, when appropriate, of the publishers of the publications.

## Fun & Face: Exploring non-verbal expressions of emotion during playful interactions

### **PROEFSCHRIFT**

ter verkrijging van de graad van doctor
aan Tilburg University
op gezag van de rector magnificus,
prof. dr. Ph. Eijlander,

in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in zaal DZ1 van de Universiteit op vrijdag 25 mei 2012 on 14:15 uur

door

Suleman Shahid

geboren te Lahore, Pakistan

### Promotors:

Prof. dr. E.J. Krahmer

Prof. dr. M.G.J. Swerts

### Promotie-commissie:

Prof. dr. E. Andre

Dr. M. B. Goudbeek

Prof. dr. P. Markopoulos

Prof. dr. M. Neerinckx

Dr. J. Westerink



Acknowledgements

### Acknowledgements

I never thought that after starting my PhD in a leap year, I would have to wait for the next leap year to finish it:). Writing a PhD dissertation is a long-term process, which demands continuous focus and dedication. Although to most people, doing PhD research generally looks like an independent undertaking and endeavor, it is almost impossible to keep going for four years with a decent pace, high level of dedication and interest without the support and encouragement of many people. Now after almost four years I am writing the last lines of my dissertation, and there is nothing more fulfilling than thanking all those who contributed to my PhD research, supported me and made my stay in Tilburg fun and fruitful. There are many to thank and I would like to start with my copromotors Emiel Krahmer and Marc Swerts.

I would like to express my sincere appreciation to Emiel Krahmer for his guidance and invaluable contributions to my thesis. Without his dedication, encouragement and prompt supervision, this PhD thesis would not have been possible. Emiel not only provided admirable supervision but also went above and beyond his typical responsibilities for bringing comfort in my PhD life. His support was never restricted to the everyday PhD tasks. I am overwhelmed by his support and genuine advices in numerous professional as well as personal matters. In the last four years, Emiel's balanced personality has continuously inspired me and I have been amazed by his remarkable time management skills, optimism, sincere amiability and positive attitude towards life. My interaction with him not only transformed me into a better researcher and teacher but also in a content person.

Thanks to Emiel, I had very excited and busy four years, full of PhD and non-PhD related activities. I particularly thank him for allowing me to do so many different things besides my PhD work e.g. teaching in Tilburg and Eindhoven, student supervision in the Netherlands and Pakistan, different administrative tasks and numerous Friday afternoon projects. He not only supported all these things but also gave tips on how to balance these activities. Sometimes PhD research looked like a part time activity when the other activities took over but he never showed a sign of worry and kept on appreciating me. His appreciation made my PhD experience thoroughly enjoyable and fruitful.

I am also indebted to Marc for his constructive supervision and dedication. Marc taught me 'how to tackle a research problem' and played a very effective role in improving my research skills and shaping my thesis. When I started this PhD as a 'Design Engineer', I had little idea about this kind of 'experiment-based' research. I would not have been able to finish this thesis without Marc's creative input. Marc always believed in me, boosted my confidence and appreciated my creative assignments. Marc's appreciation for creativity is also due to the fact that he himself is a very creative person. During our meetings, he threw out numerous great ideas for designing experiments and played a key role in coming up with a number of cool experiment 'conditions'. I am amazed by the speed with which Marc can propose high quality new ideas in a single meeting. Marc's knowledge of his research field and his passion for research has always inspired me. I greatly benefited from his valuable ideas and insightful knowledge and I hope to learn more in coming years.

Marc and Emiel complement each other in an admirable manner. Both are excellent researchers, wonderful human beings and posses an amazing sense of humor. They create an open and fresh atmosphere for their PhDs and it is a real pleasure to be embraced by their knowledge and friendliness.

Hartelijk bedankt Emiel and Marc, voor alles wat jullie voor me hebben gedaan!

I am privileged to have highly knowledgeable and reputable committee members. I sincerely thank them for their constructive criticism and feedback in the process of finalizing my thesis. I accept that I could not incorporate all of their comments in the final version of my thesis but their comments will be extremely useful in the process of revising journal papers and conducting future research on similar topics.

I do not want to say it too loudly but I guess that I could have finished my PhD a year earlier had I not been teaching the intensive User Interface design course. But the question is, did I really want to finish it early (or have a better version of the PhD thesis than I currently have) at the *cost* of *no teaching*? The answer, without a second thought, is No! Teaching the UID course was one of the best things I did during my stay at Tilburg University and I thoroughly enjoyed it. Two persons, who made it possible for me to balance my teaching and research tasks, are Fons Maes and Anne Vermeer.

I am grateful to Fons for his endless support, wonderful incentives, and optimism. He is one of the best heads of the department I met in my professional life and he has an amazing skill of coming up with creative ideas for solving complex administrative problems. I thank him for using this skill for me on a number of occasions. Moreover, I would like to thank him for coming to all UID final presentations, for distributing prizes, for holding a nice speech about me at the end of the course and for motivating me to do even better.

I also wish to express my thanks to Anne Vermeer for his help and regular feedback on teaching and supervision related matters. His trust, sincere advices, and appreciation of my creative work have helped me to excel in teaching during the last few years. Without his support it was not possible to balance the teaching and supervision load with my PhD research tasks.

My time at Tilburg University was made sociable and fun mostly due to the wonderful colleagues and lovely 'old is gold' PhD squad (Marieke, Lisanne, Ruud, Martijn, Lisette, Jorrig, and Constantijn). My thanks are due to all of them, especially Mandy Visser (for being a very caring, accommodative and friendly office mate), Lauraine de Lima (for being the most helpful secretary ever), Carel van Wijk (for his support in quantitative data analysis, especially during the first year of my PhD), Els van Loon, Monique van Alphen and Leen Jacobs (for their support in Immigration, tax and financial matters; without their support it would not have been easy to handle all the different bureaucratic matters), Ruud Mattheij (for his help in managing the UID course), and Lennard van de Laar (for his creative input in the 'front cover' design process).

I am particularly grateful to Martijn Goudbeek for his help regarding experimental design and statistical analysis. In the last couple of years, I disturbed him 'a hell of a lot of times', but for every kind of query his response was always positive - off course always accompanied by a 'witty joke'. I admire his astounding intelligence and lovely sense of humor. I hope to have a fruitful collaboration with him in coming years.

During the last three years, I had the possibility to supervise many students. I would like to take this opportunity to thank all of them, especially Eva van de Sande (for her help in collecting part of the data for chapter 3), Alyssa Tavadian (for her help in collecting part of the data for chapter 4), Lars van Bommel and Dave Kierkels (for their help in conducting the fifth study 'affective gaming for affect elicitation', which I could not include in this thesis due to time constraints), Anne Bergmans and Nienke de Laat (for their creative input and help in several design related matters).

During my PhD, I established a number of research collaborations with people in TNO Human Factors, Noldus Technologies, Philips Research Eindhoven, Eindhoven University of Technology, KU Leuven, and EPFL Lausanne. I am thankful to all my collaborators specially Hans Theuws (for providing me with every new version of Face Reader and other Noldus software for testing purposes), Mark Neerincx and Willem Melder (for their support during the affective mirror study). I also collaborated a lot with Omar Mubin and Abdullah Al Mahmud (my USI fellows) not only on 'PhD related' projects but also on numerous 'Friday afternoon projects'. I am thankful for their support and I hope to have a productive collaboration with them in future as well.

In the last couple of years, I had the opportunity to teach the 'User Centered design' course in the User System Interaction' (USI) program of Eindhoven University of Technology. I enjoyed every minute of this two week course and I am very thankful to Mia Jelsma, Panos Markopolous and Maddy Janse for giving me this excellent opportunity. USI is one of the best things I did in my life and I am always willing to contribute as much as I can for the betterment of this program.

I would like to thank all my friends for their support and for keeping me wonderful company over the years, to name just a few, Imran, Atiq, Zubair, Shan, Nayab, Maarab, Kashif, Zeeshan, Faisal, Kabir, Ishtiaq, Mubasher, Ahmed, and the many others I have definitely forgotten to mention. I am especially thankful to Imran and Zubair (for providing pleasant company in Germany and the Netherlands) and to Shan and Qazi (for doing everything for me in Pakistan on a single call).

Many family members have helped me and provided an unwavering support in the past many years. I am thankful to all of them, specially my brother Numan Shahid. I thank him for taking care of our parents and for helping them in fulfilling endless family obligations. Without his unconditional support for me and for our parents, I might not have been able to stay in the Netherlands, finish my PhD and start a career as an assistant professor.

From the bottom of my heart I would like to express my sincere gratitude to my lovely parents whose deep love and utmost care I have cherished all my life. Their prayers, concern, guidance have always been a source of inspiration and encouragement to me throughout my life. I thank my father for teaching me the importance of hard work and I thank my mother for being the best mother on earth. I also thank her for persuading and even pushing me towards doing a PhD.

These acknowledgements would remain incomplete if I didn't express my appreciation to my wonderful wife. I thank her for everything - for her care, for her patience (when I was not able to differentiate between week days and weekends, and this happened quite often), for buying my little lies (as when after 5pm I would promise that I would be home in 15 minutes and this almost never happened), for all household activities, for teaching me how to live life in an economical manner, and most importantly for raising Shaheer in a responsible manner.

I would like to finish my acknowledgement by saying *Shukar Alhamdulillah* and with a wish that someday I could make a contribution that is worthy!



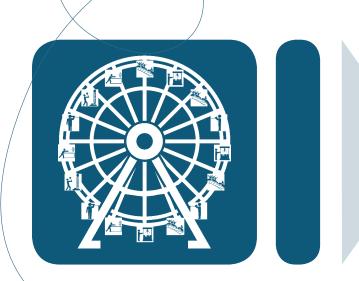


Table of Contents.



- Chapter 1: Introduction
- Chapter 2: The effect of co-presence and culture on the facial expressions of game playing children: An experimental exploration
- Chapter 3: Child-robot Interaction: How does playing a game with a social robot compare to playing a game alone or with a friend?
- Chapter 4: Video-mediated and co-present gameplay: Effects of mutual gaze on game experience, expressiveness and perceived social presence
- Chapter 5: Positive Affective Interactions: The role of repeated exposure and social presence
- 161 Chapter 6: General discussion and conclusion
- Summary
- Publication list
- 183 TiCC Ph.D. series
- 187 SIKS Ph.D. Series





### Introduction

Every year Tilburg hosts a large event known as the "Kermis" (funfair). The Tilburg Kermis is the biggest in the Benelux, as well as one of the oldest (over 440 years old), and it consists of hundreds of amusement rides, interactive fun machines, and a large variety of other games, catered for people of all age groups but especially for children. If one would study the activities of the visitors of the Kermis, it would soon become clear that these various attractions induce a wide spectrum of emotions, ranging from euphoria (for example when young children win a prize in a shooting game) and laughter (for instance when people laugh at their own distorted faces while looking into carnival mirrors) to anxiousness (when adults wait for the results of the lottery) and even frustration (when children cannot control the robotic arm of a Claw crane for grabbing a toy). Most of these emotions are clearly visible on the faces of the visitors. They not only serve as a display of the visitors' internal emotional state, but can also be understood as a sign for other people of how much fun and enjoyment a particular activity triggers. The fascinating thing about these facial expressions is that they are natural, spontaneous, and embedded in a rich social context.

But a closer look at the emotional faces of the funfair visitors might also raise questions. Do all children, for example, express their emotions in a similar way, or are there systematic differences depending on, say, whether a child is younger or older, or from one cultural background or another? And what role do other people, such as friends or family members, play when a child expresses his or her emotions; does this influence the way an emotion is expressed? Does winning a prize (in a shooting game, say) in the presence of a friend instead of alone cause a different, possibly stronger emotional response? And what if the co-present person is not a friend but a stranger? These are some of the questions that are addressed in this thesis.

### Background

As the funfair example illustrates, emotions can both have an intrapersonal function (revealing a person's internal state) as well as an interpersonal one (communicating this emotion to others). Early studies on emotion tended to focus on the intrapersonal aspects (Tomkins, 1962) and emotions were often seen as deeply personal experiences (Simon, 1967). These days researchers also believe that an important function of facial expressions of emotion is to *communicate* these emotions, arguing that they serve as social signals (Oatley, 2000). Indeed, the source of our emotions is often social and they can be seen as typical responses to different social events and entities occurring in a particular context (the joy of sitting in the roller coaster with a friend, the shared pleasure of taking a ride with the whole family in the Ferris Wheel). Modern theories of emotion, in particular, focus on the social facet of emotions and emphasize that for understanding human behaviors and social experiences, it is imperative to understand human emotions in social contexts.

In general, researchers subscribing to the 'sociology of emotions' approach emphasize that both the micro-social context (including individuals' identities and social relationships) and the macro-social contexts (such as cultural norms and societal constraints) play a role in generating and regulating our emotional responses (Parkinson, 1996). In the case of the micro-social context, there is a general consensus that the presence of others, such as friends, family members, even strangers, not only generates emotional responses in us but also influences their expression. A comedy movie,

watched in the presence of a friend, may look more hilarious than watched alone. At the macrosocial level, it has been argued that cultures can have different display rules for different social situations and relationships, which guide humans in suitably showing or suppressing a particular emotion depending on the social context (Schimmack, 1996)

The idea that emotions are an integral part of our social life is not restricted to the archetypical human-human interactions in face-to-face settings. As technology is becoming an essential part of human life, it is also giving a new dimension to emotional transactions. For example, we use modern computing systems to convey our feelings (updating our emotional state using social media) and satisfy our emotional needs (having a Skype session with loved ones, or sharing pictures on Facebook about a recent trip to a funfair and waiting anxiously for comments). In addition, technology may elicit a wide range of emotional reactions, such as frustration when a particular device or program cannot be used in the way we want or happiness after winning an online game from a digital opponent. In recent years, many researchers have worked to make human-computer interaction more natural, allowing computers to recognize and adapt to human emotions, in a strand of research also known as "affective computing" (Picard, 1997). One exciting application in this field is Affective Gaming, where a player's emotional states are monitored and directly fed back into the system, with the purpose of maintaining a state of flow which may improve the overall game experience, for instance making the game more challenging if players appear to be bored and making it easier if they appear to be frustrated (Mandryk and Atkins, 2007).

Studying emotions in naturalistic and authentic social settings is a difficult task and it is therefore not surprising that a fundamental issue in emotion research is how to develop methods for inducing emotions in a naturalistic and ethical way. An important drawback of many existing emotion induction procedures (such as showing validated movie clips with a certain emotional content, or playing emotional music (Coan and Allen, 2007)) is that they are primarily designed for non-social situations, in which individual participants are confronted with these emotional stimuli. Furthermore, such emotion induction methods may be effective in one particular culture or for one age group (typically adults), but not for a variety of other groups and cultures. This is why some have argued that the results gained using these techniques may not easily generalize (Kaiser and Wehrle, 1996).

In this thesis, we suggest a new method for eliciting emotions in natural and ethical manner for different social settings, which we call the GAME (Game As a Method of eliciting Emotions) paradigm (Shahid et al., 2008). Games are usually characterized as interactive, result-oriented, competitive, dynamic and engaging in nature (Salen and Zimmerman, 2003) and precisely these features can be exploited for using them as emotion inducers. The level of engagement in a game can lead to an intense emotional state, which may be positive --when the game is going well-- or negative --when it is not-- (Lazzaro, 2004). The relation between emotions and games is interesting from a double perspective: on the one hand, games appear to present themselves as handy tools to induce natural emotions; on the other hand, the level of emotional involvement while playing games can be used as a metric to measure the fun and engagement in the games.

Using the GAME paradigm, we zoom in on two factors potentially influencing the expression of emotion, namely physical co-presence of an external agent (a factor related to micro-social context) and cultural background of users (a factor related to macro-social context). In particular, we explore to what extent non-verbal expressions of emotion of participants engaging in playful interactions are influenced by their cultural background and by the physical co-presence of a game-partner. To study the impact of cultural background, we compare players from a Western-European culture (The Netherlands) with those from a South-Asian culture (Pakistan), a comparison which so-far has received hardly any attention but one which is interesting nevertheless, since these cultures differ along a number of dimensions which might influence emotional expressiveness. In addition, and orthogonally to the previous point, we experimentally compare a number of social settings, comparing the emotional expressions of participants playing games alone with those who play games with a social robot (the iCat) or with a friend, who could be physically co-present or not.

### Overview

In this dissertation we report on four studies, each one using a game developed under the GAME paradigm. Each of these four studies explores one way in which the social context may shape nonverbal expressions of emotion and social interaction among players. The first study investigates the influence of physical co-presence of a friend as a game partner on the emotional expressions of game playing children across cultures (Pakistani vs. Dutch). We also explore to what extent a simple card game can be used for eliciting rich socio-emotional responses across cultures. If emotional expressions primarily serve to reflect the individual and are not influenced by social context (culture and co-presence), then we can hypothesize that emotional responses will be highly comparable for both cultures while playing alone or with a friend. However, if the expressions also have a considerable social component and are rooted in culture, we expect to find differences in expressiveness across cultures and between individual players and pairs of players.

The second study explores the influence of the presence of a social robot (the iCat) as a game partner on the social and emotional response of children across cultures (Pakistani vs. Dutch). We employ the same collaborative card game and general set-up as used in Study 1 for eliciting emotions and social interaction. This allows us to effectively propose a new method to evaluate children's appreciation of social robots, by asking the child whether playing a game with a social robot is more similar to playing this game alone or with a friend. Ideally, playing with a social robot will be experienced more like playing with a friend (an ideal and truly social experience) than playing alone (an unsocial experience). If children find playing with a social robot comparable to playing with a friend, then we would expect similar socio-emotional response during child-child and child-robot interactions. Similar to the first study, if cultural background shapes the way children emotionally interact with a social robot then we would also expect to find differences in expressiveness and social interaction across cultures.

The third study investigates how variations in natural gaze influence the perceived social presence, emotional expressions, social interaction and game playing behavior of children during video-mediated gameplay. We also compare how similar a video-mediated gameplay experience is compared to a co-present experience (that is: a situation in which children are physically co-present in pairs and interact in a face-to-face manner). In this study, we particularly focus on the role of 'mutual gaze' (i.e. ability to have an eye-contact) during video-mediated gameplay. Mutual gaze plays a key role in regulating human-human communication, suggesting that the absence of mutual

gaze during mediated communication may influence players' game experience and perceived sense of presence in a technologically mediated environment. Furthermore, if the co-present (face-to-face) game playing experience is considered as an ideal form of communication then we would expect differences in the way people express emotions and exchange social cues during video-mediated gameplay and co-present gameplay.

The fourth and final study evaluates if and how an affective interface (a laughter mirror), which distorts a user's face in a manner which might be perceived as funny, and which adapts itself according to the perceived user's state, can be used as a tool for inducing positive emotions. In this study, consistent with the first three studies, we explore the influence of physical co-presence of a friend on the induction and expression of positive emotions and laughter. If emotion induction and positive expressions are influenced by the social context, we again expect a stronger response of participants in the presence of their friends. In this study, we also focus on one additional factor, namely "repeated exposure" and investigate whether the induced positive affect is repeatable after a fixed interval. If the laughing mirror is strong enough in inducing laughter repeatedly then we expect comparable emotional response (both felt and expressed) after a first and a second session.

### References

Coan, J.A., Allen, J.J.B., 2007. Handbook of Emotion Elicitation and Assessment. Oxford University Press.

Kaiser, S., Wehrle, T., 1996. Situated emotional problem solving in interactive computer games. In N. Frijda (Ed.), Proceedings of the International Society for Research on Emotions, 276-280.

Lazzaro, N., 2004. Why We Play Games: Four Keys to More Emotion Without Story. Game Developer Conerence (GDC), Online April. 2012,

http://www.xeodesign.com/whyweplaygames/xeodesign\_whyweplaygames.pdf.

Mandryk, R.L., Atkins, M.S., 2007. A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies. International Journal of Human-Computer Studies, 65(4), 329-347.

Oatley, K., 2000. Emotion and Distributed Cognition. Invited talk to the 2000 Annual Conference of the British Psychological Society, Summarised by Paul Redford in 'The Psychologist', 13(6), 290–229.

Parkinson, B., 1996. Emotions are Social. The Journal of the British Psychological Society, 87, 663-683.

Picard, R., 1997. Affective Computing. The MIT Press, Cambridge, MA.

Salen, K., Zimmerman, E., 2003. Rules of Play, Game Design Fundamentals. The MIT Press, Massachusetts.

Schimmack, U., 1996. Cultural Influences on the Recognition of Emotion by Facial Expressions: Individualistic or Caucasian Cultures? Journal of Cross-Cultural Psychology, 27(1), 37-50.

Shahid, S., Krahmer, E., Swerts, M., 2008. GamE: Game as a method for Eliciting Emotions. Proceedings of the Measuring Behaviour Conference (MB 2008), Maastricht, the Netherlands.

Simon, H.A., 1967. Motivation and emotional controls of cognition. Psychological Review, 74, 29-39.

Tomkins, S., 1962. Affect, inagery, consciousness. Springer, New York.





## CHAPTER 2

The effect of co-presence and culture on the facial expressions of game playing children: An experimental exploration



# The effect of co-presence and culture on the facial expressions of game playing children: An experimental exploration

### ABSTRACT.

In this chapter we study how children of different age groups (8 and 12 years old) and with different cultural backgrounds (Dutch and Pakistani) signal positive and negative emotions while playing a game either alone or together. A simple but effective number guessing game was developed and used as a tool for inducing emotions, which was played by Pakistani and Dutch individual and pairs of children. The data collected was used in a series of cross-cultural perception studies, in which Dutch and Pakistani observers classified the emotional expressions of the Dutch and Pakistani children, where a higher level of expressiveness is assumed to result in better classifications. Results show that classification accuracy is uniformly high for Pakistani children, but drops for older and for winning Dutch children. Furthermore, the correct classification in both cultures is higher for children playing games in pairs than for children playing alone. The implications of these findings for affective computing are discussed.

A journal paper based on this chapter is submitted for publication. Earlier versions of this chapter appeared as: 1) Book chapter: Shahid, S., Krahmer, E., Swerts, M., 2008. Alone or Together: Exploring the effect of physical co-presence on the emotional expressions of game playing children across cultures. In P. Markopoulus et. al. (Ed.), Fun and Games, 94-105. Berlin/Heidelberg: Springer. - **Best Paper Award**, and as 2) Shahid, S., Krahmer, E., Swerts, M., 2008. GAME: Game As a Method for eliciting Emotions, Proceedings of the Measuring Behaviour Conference (MB 2008), Maastricht, the Netherlands.

### 1. Introduction

Emotions are an important ingredient of daily life and play a central role in many of our everyday interactions. It is not surprising, therefore, that in recent years research on emotions has also entered the field of human-computer interaction, based on the assumption that this can make interaction between humans and computers more natural and satisfying (Pantic and Rothkrantz, 2003) (Picard, 2003). There are already a number of application areas where integrating emotions into the system design is both realistic and valuable. One exciting example is Affective Gaming, where a player's affective states are monitored and directly fed back into the system, with the purpose of maintaining a state of flow which improves the overall game experience (Mandryk and Atkins, 2007).

A key challenge faced in the design of many affective system is that of recognizing natural human emotions. Researchers have explored two alternative strategies: performing physiological measurements on users, or focussing on their facial expressions and other forms of non-verbal behavior. Physiological measurements, such as skin conductance, blood pressure, heart rate, and brain scanning (Picard et al., 2001), have been argued to be reliable (a person may smile without feeling happy, but physiological signals are harder to fake) (Picard and Klein, 2002). This is an area of active research, and recent technological advancements are helping researchers to gain valuable insights into the affective states of users. Still, this approach also has its limitations. In particular, it is not always obvious what physiological measures are actually assessing and measurements are sometimes difficult to interpret. Furthermore, in real world applications (including but not limited to gaming), measuring the physiological response of a user can be invasive and practically unfeasible, even with the availability of good sensing technologies.

An alternative approach, and the one we explore in this chapter, is to automatically observe the non-verbal behaviour of users. It is well known that humans have a natural tendency to communicate their emotional states both consciously and unconsciously via facial expressions, body postures, gestures and tone of voice. Such non-verbal expressions can often be measured in a non-intrusive way (more and more computers and mobile devices contain build-in camera's). Moreover, there is a long research tradition (at least going back to (Darwin, 1972)) of analysing non-verbal expressions of emotion, and the insights from this tradition can offer an important source of inspiration for developing affective interfaces.

However, there is a significant gap between the way non-verbal expressions of emotion have often been studied in the past, and the expressions of users that are encountered by affective interfaces such as games. In many older studies, emotions were studied based on posed, static expressions of adult individuals, whereas computer games are often played by children of different ages, who show spontaneous, dynamic expressions, and who may play a game alone or with a friend. In addition, these children may come from different cultural backgrounds, which is another understudied aspect in older research on emotion. In general, it is fair to say that little is known about how children of different age groups and belonging to different cultures show their emotions in a natural, social environment.

Clearly, this is a potential problem for many affective applications. It has been found, for instance, that an emotion recognizer, which was trained on acted, spoken expressions, performs poorly when tested on spontaneous expressions in a natural setting (Vogt and André, 2005). In a

similar vein, it does not seem farfetched to assume that such recognizers would show a comparable drop in performance when trained on data from one cultural group, and tested on users from a different cultural background. However, the extent to which emotional facial expressions of gameplaying children differ as a function of the setting in which games are played (alone or together, in one culture or another) remains largely unexplored

In this chapter, we report on a series of experiments that focus on the facial expressions of 8 and 12 year old children, with either a West-European (Dutch) or South-Asian (Pakistan) cultural background, playing a computer game either alone or with a friend. Using perception experiments, we investigate if and to what extent their emotional facial expressions change in these varying circumstances.

### 1.1 Affective Interfaces

In recent years, interface designers have started including emotional expressions in user interfaces using different modalities such as speech (Clavel et al., 2008), textual contents (Lu et al., 2009), video (Baenziger et al., 2009) and synthetic facial expressions (Sattar et al., 2009). In some cases, combinations of these modalities can be used as well, such as in the design of Embodied Conversational Agents. In such agents, recognizing and showing appropriate emotions is often an integral part of the agent's design (Beale and Creed, 2009), to increase the credibility of the agent (Cowell and Stanney, 2005). Applications of these agents include interactive teaching, and developing a social bond between (virtual) teacher and students (Sarrafzadeh et al., 2008), persuading users to maintain a healthy lifestyle (Bickmore and Picard, 2005), and virtual training and simulation (Gratch and Marsella, 2004). Affective Gaming is another important application area, where the recognition of user emotions and the adequate response to them can be integrated into the system design (Mandryk and Atkins, 2007), which may help students to learn complex concepts in a playful manner (Annetta et al., 2009), and may even induce emotions in players (Shahid et al., 2009).

Affective computing is a relatively new area of research and that is why researchers are also focusing on developing new methods for collecting emotional data (Isomursu et al., 2007), for recognizing affective states of users (Bailenson et al., 2008) and for evaluating the use of emotions in interfaces (Isbister et al., 2007). However, designing an affective interface is not straightforward (Tractinsky, 2004), not least because users may differ along numerous dimensions (including age, social co-presence and cultural background) and how such differences should impact the design of affective interfaces is still relatively understudied.

### 1.2 Social and cultural aspects of human emotion

While there can be little doubt that emotions are at least partially a social phenomenon, early studies on emotion were inclined to focus on the intrapersonal aspects of emotions (Tomkins, 1962) and emotions were often seen as profoundly personal experiences (Simon, 1967). Many researchers believed that an important function of emotional facial expressions was to display the internal state of the individual (Ekman, 1994). More recently, researchers also began to investigate the interpersonal aspect of emotions, emphasizing that emotional expressions may serve as social signals (Fridlund, 1994, Rime et al., 1991). (Oatley, 2000) claimed that one of the major roles of emotion is "the management of co-operation between self-and-other". According to this perspective, the expression of emotion helps guide and motivate social partners towards the completion of shared objectives (Ward, 2005).

In general, researchers subscribing to both perspectives accept that non-verbal expressions of emotions, either via voice or via facial expressions, play a role in social communication, but the general trend in research has been to study the expression of emotions by looking at individuals, mostly without taking the social context into account (Manstead, 2005). The field of human-computer interaction (HCI) is no exception in this regard where the effect of affective interfaces has mainly been investigated in a non-social context, albeit with a few notable exceptions such as (Kleinsmith et al., 2006)

Naturally, many different factors may contribute to the social context, but in this study we limit ourselves to two important ones, namely *physical co-presence* and *culture*. If emotional expressions primarily serve to reflect the individual state of an individual, then it could be hypothesized that the emotional expressions of users will remain the same across various contexts (when they do an activity alone or together with someone else). Although some researchers have investigated the effect of the social context on the expression of emotions (Masuda et al., 2008), there are relatively few studies looking at the direct effect of *physical co-presence* on the expression of emotion (Wagner and Lee, 1999).

Even though certain expressions of emotion have been claimed to be universal, culture arguably also plays an important role in shaping the emotional reactions and perceptions of people (Elfenbein et al., 2002, Mesquita, 2001). What is natural in a particular social setting of one culture, might be impolite or otherwise marked in a similar social setting of another culture (Mesquita and Frijda, 1992, Mesquita and Leu, 2007). While various studies have compared the expression of emotions across cultures in different social settings (Elfenbein and Ambady, 2002), ours is, to the best our knowledge, the first study looking at game-playing children of different age groups from a cross-cultural perspective.

### 1.3 The current study

In this study, we study the emotional expressions of children from two different cultures (Pakistan and Dutch) who either play a game alone or in pairs (together with a friend). If the children's emotional expressions have a considerable social and cultural component then we would expect to find differences in expressivity not only between individual players and pairs of players but also between Pakistani players and Dutch players. In addition to these contextual factors, we also include one individual factor in our studies, namely age. An interesting albeit largely unexplored question is how children learn to "use" their non-verbal expressions in a social context. Generally speaking, adults are more stable in their expression of emotions than children, who are overall more expressive than adults e.g., (Fabes and Martin, 1991). In particular, young children use emotions as a communication aid because of their limited verbal abilities. This raises the question how children "grow" towards more adult-like expressiveness, and whether this developmental process is the same for different cultures and contextual settings. In general, we know relatively little about how children of different age groups express their emotions audio-visually, and how sensitive observers are to these emotional expressions. For our studies, we concentrate on facial expressions during game play,

not only because this is highly relevant for HCI and affective gaming, but also because games offer an attractive and natural way of inducing emotions.

An important drawback of many existing emotion induction procedures (such as showing movie clips with an emotional content, or playing emotional music) (Coan and Allen, 2007) is that they are primarily designed for non-social situations, in which individual participants are confronted with emotional stimuli. Furthermore, such emotion induction methods may be effective in one particular culture or for one age group (adults), but not for a variety of groups and cultures. This is why some have argued that the results gained using these techniques may not easily generalize (Kaiser and Wehrle, 1996).

Games, by contrast, may be characterized as interactive, result-oriented, competitive, dynamic and engaging in nature (Salen and Zimmerman, 2003) and precisely these features can be exploited for using them as emotion inducers. The level of engagement in a game can lead to an intense emotional state, which may be positive (when the game is going well) or negative (when it is not) (Lazzaro, 2004). There are very few examples where researchers used games as a tool for inducing emotions (Wang and Marsella, 2006), although as far as we know, not in a social or cross-cultural setting.

The rest of this chapter is organized as follows. In section 2, we describe our data collection procedure, which was based on a simple card-guessing game. This game was played by Dutch (a western-European culture) and Pakistani (a south-Asian culture) children of two age groups, playing the game either on their own or together with a friend. The data thus collected was used in two perception studies. Study 1, described in section 3, zoomed in on the effect of culture on emotional expressions, by asking both Dutch and Pakistani adult observers to guess based solely on the facial expressions of individual children whether they had just lost or won their most recent game. Study 2, which is the topic of section 4, looks at the effect of co-presence, using a similar setup. Finally, in section 5 we discuss the collected data in a more qualitative manner and draw conclusions for affective game design.

### 2. Data Collection

### 2.1 GAME paradigm

To induce positive and negative emotions in children in a natural way, we developed the GAME (Game As a Method of eliciting Emotions) paradigm. Under the GAME paradigm, the objective is to design simple games which can be used as tools for inducing emotions in a natural way (Shahid et al., 2008).

### The card game

We developed a simple card game as a tool for inducing emotions in children. In the card game, children have to guess whether an upcoming number is higher or lower than the previous (reference) number. When the game starts, players see a row of six cards on the screen where the number of the first card is visible ('3' in the case of the example in Figure 1) and the other five cards are placed upside down so the numbers are hidden. The numbers on the cards range from 1 to 10, and a number displayed once is not repeated in a particular game. Players have to guess whether the number on the next card will be higher or lower than the previous number.

Once players have guessed the number, the relevant card is turned around and the correct answer is shown on the screen. Players are also informed about the correctness or incorrectness of their answer via a characteristic non-speech feedback sound. If players make a wrong guess, the game is finished and they move to the next game. If players guess the number correctly, they are asked to guess the next number. Players only win a game if they guess all numbers in the row correctly. In the case of winning a game, they receive a coin. At the end of the experiment, which consisted of multiple games, players could exchange the coins they won for a prize.

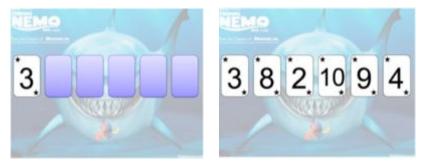


Figure 1: Winning Scenario - Start of a game (left) and end of the game (right)

The card games were developed using Microsoft\* PowerPoint\*. Appropriate colourful images were chosen for the game background and different animations were used to turn cards around in an attempt to make the game attractive for children. During the experiment, children played six games, and could in theory win six coins. However, unbeknownst to the children, each game was completely deterministic, and two different game variants were used. In the first alternative, a rational decision procedure would result in winning the game, and in the second alternative, being rational would result in losing the game. Figure 2 is an example of the latter: the most probable outcome for the final card would be a "lower" number than 9, but guessing "lower" would make this a losing game. The two other losing games were: 8-3-9-7-2-1 and 1-9-4-8-3-2. Winning scenarios were 9-2-8-1-10-7, and 3-8-2-10-9-2 and 1-3-9-2-7-6. Winning and losing games were mixed in the sequence, starting and ending with a variant in which children were likely to win.



Figure 2: Losing Scenario - Start of a game (left) and end of the game (right)

### 2.3 Participants

In total, 144 children participated in the experiment of which 96 children played the games in pairs, 48 Dutch children (24 pairs) and 48 Pakistani ones (24 pairs). The remaining 48 children played the game individually, 24 Dutch and 24 Pakistani ones. In both conditions (individual and pairs), half of the Dutch and Pakistani children were around 8 years old (group 4 in the Dutch elementary school system, second grade) and the other half were around 12 years old (Dutch group 8, sixth grade). Pairs were always self-selected and always consisted of children from the same grade. Parents and teachers gave prior written permission for their child to participate, and signed a consent form stating that the recordings could be used for research purposes.

### 2.4 Procedure

The experiment was conducted in four elementary schools, two in Tilburg (the Netherlands) with Dutch participants and two in Lahore (Pakistan) with Pakistani participants. The procedure for both conditions (pairs and individuals) in all four schools was essentially the same. A separate room was chosen, where, depending on the condition, pairs of children or individual children were invited and asked to sit on the chair(s) placed in front of a desk on which a laptop computer was placed. On the top of the laptop, a camcorder was positioned in such a way that it could record the children's faces and the upper part of their body. Another computer was attached to the laptop to facilitate the experiment leader in controlling the game. The experimenter was always outside of the visual field of the game-playing children. Once the children were in the room and had chosen the appropriate seat(s) for sitting, the experimenter welcomed the children and started a small talk discussion by asking a few questions to break the ice (e.g., "How old are you? Do you like to play games?"). After this introductory phase, the experimenter gave spoken instructions, telling the children about the game and the coins they could win. All the game rules outlined in the previous section were explained to the children (in Urdu and Dutch for Pakistani and Dutch children, respectively), and when they seemed to understand the rules, the experimenter started a practice game ("So you only have to say whether the next card is higher or lower. This is just an exercise and it doesn't really count"). After this practice session, the experimenter asked the children whether they had any questions, and if not the experimenter left the children's field of vision and started the first experimental game.

The experiment did not have a fixed duration because the (pairs of) children varied in the time they needed to make a decision, and in the case of pairs this sometimes involved substantial discussions between the children, which also varied substantially from pair to pair. On average each experimental session lasted for approximately 10 to 15 minutes. After the sixth and final game, the experimenter congratulated the children and they could trade in their coins for an individual gift (a whistle or small plastic ball in the case of the Dutch participants and big size chocolates and key rings for the Pakistani participants).

Overall, the game worked well, in that all individuals and pairs of participants indeed made the rational choices they were expected to make in most of the games, so that each individual child and pair of children lost at least two games and won at least two games. Furthermore, not a single child noticed and reported that the game was rigged, i.e., a deterministic simulation. The data gathered in this way is rich and constitutes a valuable collection of non-verbal emotional responses to winning or losing a game. Informal observations reveal clear differences in the way emotions are expressed non-verbally between individual players, and between individual players and players in pairs. There also seem to be clear differences among the different age groups and cultures. We attempt to quantify all these differences in two cross-cultural perception studies described below.

### 3. Study 1: The effect of culture on emotional responses

### 3.1 Stimuli

From the child-pairs (92 children) that participated in the game playing sessions, we selected the responses to the first two of their winning games (in which they made a correct prediction for the last card) and the first two of their losing games (in which the final guess turned out to be incorrect). Some of the Pakistani and Dutch children could not be used in the perception test because of the poor video quality. The stimuli were cut from the moment the final card was turned until the primary response of the game playing children was finished. This resulted in 88 Dutch stimuli [11 (8 year old pairs) + 11 (12 year old pairs) x 2 won x 2 lost] and 80 Pakistani stimuli [10 (8 year old pairs) + 10 (12 year old pairs) x 2 won x 2 lost]. The selected stimuli were presented in a perception test to observers for a rating experiment. In this experiment, stimuli were presented in a random order, in a vision-only format to avoid participants from relying on auditory cues (e.g. "Yahooo!").

### 3.2 Participants

131 adults participated in the cross-cultural perception experiments, 71 Dutch (M = 21 years, SD = 3.7 years) and 60 Pakistani (M = 20 years, SD = 2.3 years), with a roughly equal number of men and women.

### 3.3 Procedure

Four experiments were conducted: Dutch viewers judging Dutch children [31 participants], Dutch viewers judging Pakistani children [40 participants], Pakistani viewers judging Dutch children [30 participants] and Pakistani viewers judging Pakistani children [30 participants] with essentially the same procedure for all four experiments. For every experiment, groups of participants were invited into a quiet classroom where the computer screen was projected on the classroom wall using a beamer. Participants were told that they would see 88 stimuli in the case of Dutch children or 80 stimuli in the case of Pakistani children. In addition, they were instructed that the children were showing their emotions after winning or losing a game and that their task was to determine whether the children had just won or lost their game. Each stimulus was preceded by a number displayed on the screen indicating the upcoming stimulus, and followed by a six second pause during which participants could fill in their score on the answer form. The actual experiment was preceded by a short training session in which 3 clips were shown (different from the ones shown in the actual experiment) to make participants familiar with the stimuli and the experimental task. If everything was clear, the actual experiment started, which lasted for approximately 16 minutes. During the experiment there was no interaction between participants and experimenter.



Figure 3: Representative stills for (L-R): Pakistani 8 year old winning and loosing, (1st row), Dutch 8 year old winning and loosing, (2<sup>nd</sup> row), Pakistani 12 year old winning and loosing, (3<sup>rd</sup> row), and Dutch12 year old winning and loosing, (4th row)

# 3.4 Statistical Analysis

All tests for significance were performed using a repeated measures analysis of variance (ANOVA) with two within-subjects factors: Age group (levels: 8 years old, 12 years old), Game status (levels: Win, Lose) and two between-subjects factors: Game players (levels: Dutch, Pakistani), Observers (levels: Dutch, Pakistani) and with percentage of correct classification as the dependent variable.

# 3.5 Results

Table 1 summarizes the results. Overall, classification accuracy is high, which indicates that participants were able to correctly determine the status of the game based on the emotional response of the children. In most cases there are no differences between Dutch and Pakistani observers. Quite substantial differences are found between stimuli of Dutch and Pakistani children, however; the percentage of correct classifications for the Pakistani children are uniformly high (always more than 80% correct), whereas for the Dutch children it can be seen that the percentage of correct classification drops for the 12 years olds. Interestingly, when 12 year old Dutch children win the game, this is rather difficult to see; most observers, especially the Dutch ones, perceive this as losing.

Game Players	Age Group	Game Status	Observers	
			Dutch	Pakistani
Dutch	8	Lost	.87 (.02)	.83 (.02)
		Won	.70 (.02)	.74 (.02)
	12	Lost	.76 (.02)	.76 (.02)
		Won	.37 (.02)	.46 (.02)
Pakistani	8	Lost	.86 (.01)	.83 (.02)
		Won	.92 (.02)	.89 (.02)
	12	Lost	.93 (.02)	.86 (.02)
		Won	.91 (.02)	.87 (.02)

Table 1 - (Experiment 1): Mean correct classification results (with standard errors between brackets) for Dutch and Pakistani observers judging stimuli from Dutch and Pakistani game playing children as a function of Age Group and Game Status

The first thing to note in Table 1 is that the percentage of correct classifications is higher for Pakistani children (M = .89) than for Dutch Game playing children (M = .69), F(1,127) = 380.86, p < .001,  $\eta_p^2 = .75$ . In addition, overall the percentage of correct classifications is higher for 8 year olds (M = .83) than it is for 12 year olds (M = .74), F(1,127) = 226.80, p < .001,  $\eta_p^2 = .64$ . Similarly, overall the percentage of correct classifications is higher for losing games (M = .84) than it is for winning games (M = .73), F(1,127) = 62.07, P < .001,  $\eta_p^2 = .33$ . No main effect of Observer was found (F < 1).

Given the differences in how Dutch and Pakistani children respond to winning and losing, the interaction effects are especially interesting. To begin, we found significant interactions between Age group and Game players (F (1,127) = 342.14, p < .001,  $\eta$ 2p = .73) and between Game status and Game players (F (1,127) =103.63, p < .001,  $\eta$ 2p = .45), which can be explained by observing that the percentage of correct classifications drops for Dutch 12 year olds and for Dutch children that lose their game. These two interactions are visualized in Figures 4 and 5, respectively.

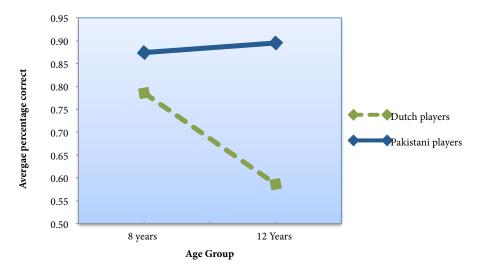


Figure 4: Average percentage correct as a function of Age Group and Game Players

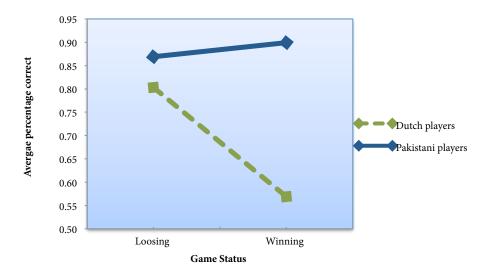


Figure 5: Average percentage correct as a function of Game Status and Game Players

We also found a significant interaction between Age group and Game Status (F (1,127) = 127.81, p < .001,  $\eta^2_{P} = .50$ ). This interaction can be explained by the fact that overall for 8 years old, both winning and losing are mostly recognized correctly, while for 12 years old (mainly for Dutch 12 years old), losing is more often classified correctly than winning (Figure 6). The related three-way interaction between Age group, Game status and Game players was significant as well (F (1,127) = 40.12, p < .001,  $\eta^2_p = .24$ ).

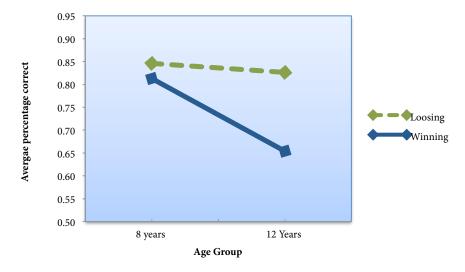


Figure 6: Average percentage correct as a function of Age Group and Game Status

# 3.6 Summary

The results from study 1 showed that the overall number of correct classifications was high, indicating that people could usually decide, based on the visual responses of the children, whether they had just lost or won a game. Interestingly, for Dutch children it appears that losing is easier to detect than winning; this is true for both 8 and 12 year olds, but the differences are especially clear for the older children, where both Dutch and Pakistani observers misclassify the majority of the games that were won as games that were lost. It is clear from study 1 that *culture* plays an important role in shaping the expression of emotions. We not only found cross-cultural differences but also *age* differences in the expression of emotions. Furthermore, we also learnt that there was no in-group advantage in judging the emotions of game playing children. Dutch were no poorer than Pakistanis in judging Pakistani children, and vice versa. In study 1, the stimuli always consisted of *pairs* of children. To find out to what extent the *physical co-presence* effects the emotional response of children and how this emotional response varies across cultures, we conducted a second study.

# 4. Study 2: The effect of physical presence on emotional responses

### 4.1 Stimuli

From all of the individuals and pairs that participated in a game playing session, we selected the first response of their winning game (in which they made a correct prediction for the last card) and the first response of their losing games (in which the final guess turned out to be incorrect). In addition, from the clips of child-pairs, we randomly selected one child from each pair by zooming in on his/her face as shown in figure 7. In this selection, half of the children sitting on the right chair and half of the children sitting on the left chair were selected. We selected equal numbers of boys and girls. As in Study 1, the stimuli were cut from the moment the final card was turned until the main response of the child was finished. This resulted in 96 Dutch stimuli 2[alone/together] x 2[win/lost]

x 2[boy/girl] x 2[8/12 years old] x 6[instances]. In a similar fashion, the 96 Pakistani stimuli were developed. Stimuli were presented to participants in a random order, in a vision-only format to avoid participants from relying on auditory cues.



Figure 7: Representative stills of children playing the game alone (top row) or in pairs (bottom row). Pakistani children are on the Left, Dutch ones on the Right.

# 4.2 Participants

72 Dutch adults (M = 21 years, SD = 2.1 years), with a roughly equal number of men and women, participated in the perception experiments. None had participated in Study 1.

# 4.3 Procedure

Two group experiments were conducted: Dutch viewers judging Pakistani children [36 participants] and Dutch viewers judging Dutch children [36 participants] with essentially the same procedure for both experiments. Each experiment lasted for approximately 18 minutes. The procedure for both experiments was similar to those conducted in Study 1 and described in section 3.4.

# 4.4 Statistical Analysis

All tests for significance were performed using a repeated measures analysis of variance (ANOVA) with three within-subjects factors: Co-presence (levels: individual, individual from pairs), Age (levels: 8 years old, 12 years old), Sex (levels: boy, girl), and one between-subjects factor: Culture (levels: Dutch, Pakistani) and with percentage of correct classification as the dependent variable.

# 4.5 Results

Table 2 summarizes the results. Again it is clear that most classifications are correct, which indicates that participants were able to correctly determine the status of the game based on the emotional response of the children.

Co-presence	Sex	Age Group	Mean (SE)
Individuals	Boy	8	.62 (.02)
		12	.46 (.02)
	Girl	8	.56 (.02)
		12	.55 (.02)
Individuals taken	Boy	8	.76 (.01)
from pairs		12	.68 (.01)
	Girl	8	.71 (.02)
		12	.72 (.01)
Individuals	Boy	8	.62 (.02)
		12	.78 (.02)
	Girl	8	.59 (.02)
		12	.64 (.02)
Individuals taken	Boy	8	.92 (.01)
from pairs		12	.87 (.01)
	Girl	8	.89 (.02)
		12	.94 (.01)
	Individuals taken from pairs  Individuals	Individuals taken from pairs  Girl  Individuals  Boy  Girl  Individuals  Boy  Girl  Individuals taken from pairs	12

Table 2 - (Experiment 2): Mean of correct classification results (with standard errors between brackets) for Dutch observers judging stimuli from Dutch and Pakistani game playing children (individual and individual taken from pairs) as a function of Co-presence, Sex and Age Group

First of all, it is interesting to observe about these results that the average percentage of correct classifications is significantly higher for individuals taken from pairs (M = .81) than it is for individuals playing alone (M = .60), F(1,70) = 716.75, p < .001,  $\eta^2_p = .91$ . Furthermore, the average percentage of correct classifications for Pakistani children (M = .78) is significantly higher than for Dutch children (M = .64), F(1,70) = 168.9, p < .001,  $\eta_p^2 = .71$ ). No significant main effect of sex is found.

Once again, given the differences between individuals and pairs in their emotional expressions, it is interesting to look at the significant two-way interactions. We found a significant interaction between culture and co-presence (F(1,70) = 24.18, p < .001,  $\eta_p^2 = .26$ ). This interaction is visualized in Figure 8 and can be explained as follows: even though Pakistani children are overall more expressive than Dutch children (higher percentage correct), the difference between Pakistani individuals (M = .66) and Pakistani individuals from pairs (M = .90) is larger than the difference between Dutch individual (M = .55) and Dutch individuals from pairs (M = .72).

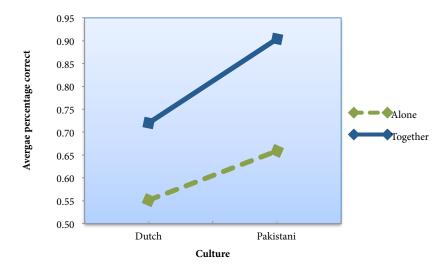


Figure 8: Average percentage correct as a function of Culture and Co-presence

We also found a significant interaction between Culture and Sex (F(1,70) = 7.77, p < .01,  $\eta_p^2 =$ .11). This interaction can be explained by observing that while both Pakistani boys and girls are more expressive than Dutch boys and girl, this difference is overall somewhat higher for the boys (Figure 9).

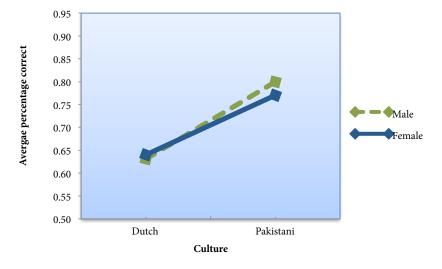


Figure 9: Average percentage correct as a function of Culture and Sex

Another interesting significant interaction was found between Culture and Age (F (1,70) = 46.78, p < .001,  $\eta^2_p$  = .40). This interaction reveals that even though the Pakistani children aged 8 and 12 are highly expressive, the Dutch 8-year olds are overall clearly more expressive than the Dutch 12 year olds (as visualized in Figure 10).

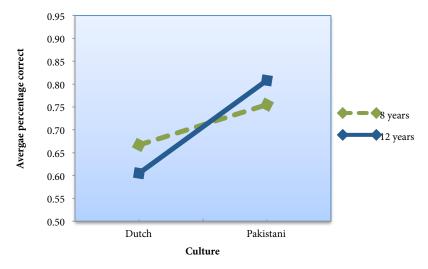


Figure 10: Average percentage correct as a function of culture and age group

We also found a significant interaction between Co-presence and Sex (F(1,70) = 9.12, p < .01,  $\eta_p^2 = .11$ ). This interaction can be explained by the fact that while playing a game together, both boys and girls are comparable in their expressiveness, but when playing alone, boys were overall perceived as somewhat more expressive than girls (Figure 11).

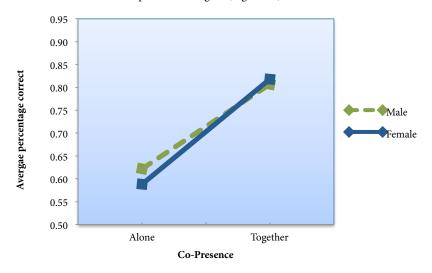


Figure 11: Average percentage correct as a function of Co-presence and Sex

Finally, a significant interaction was found between Sex and Age (F(1,70) = 11.28, p < .01,  $\eta^2_p = .14$ ). This interaction is caused by the fact that 8 year old boys are overall more expressive than 12 year old boys, while the opposite is true for girls, as can be seen in Figure 12.

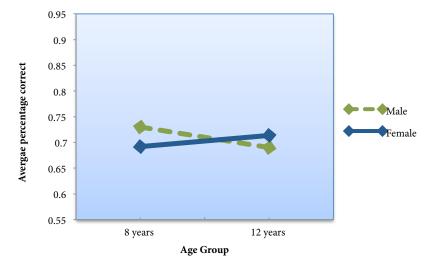


Figure 12: Average percentage correct as a function of Age Group and Sex

# 4.6 Summary

We found that judges' classifications were more often correct for children taken from pairs than for children playing alone, even though in both cases only the face of a single game playing child was visible. This suggests that children playing the game in pairs were more expressive than those playing the game individually, irrespective of the cultural background of the children.

# 5. General Conclusion and Discussion

In this chapter, we investigated how children of different age groups (8 and 12 years old) belonging to different cultures (Pakistani and Dutch) express their emotions in different social settings (playing a game alone or together with their friends). For this purpose, we devised a simple but effective and culture-independent computer game, in which individual children or pairs of children have to guess whether a card will contain a higher or lower number than a reference card. The recordings collected in this way were used in two perception experiments. In the first, both Dutch and Pakistani adult viewers saw fragments (without sound) of Pakistani and Dutch children; they were instructed to guess for each fragment whether the children that were shown had just won or lost their game.

We found that overall the number of correct classifications was high, indicating that our judges could usually decide whether the children had just lost or won, based on their facial expressions. Interestingly, scores were high overall for Pakistani children, whereas for the Dutch children it was found that losing was easier to detect than winning, especially for the 12 year old children. This

indicates that expressions of disappointment were more pronounced in the Dutch children than expressions of joy, as Figure 13 illustrates.





Figure 13: 8 year old Dutch children loosing (left) and 12 year old Dutch children winning (right)

This is different for the Pakistani children, where both winning and losing are easier to detect (though the classification scores are slightly higher for winning games), both for 8 and for 12 year old children, as shown in figure 14





Figure 14: 8 year old Pakistani children losing (left) and 12 years old Pakistani children winning (right).

Experiment 2 presented judges with clips of individual, "zoomed in" children who either played the games alone or together with a friend. Besides confirming many of the findings from Experiment 1, this study also revealed that judges were more often correct about children who played in pairs than about children who played alone, indicating the children are more expressive when a friend is present. This finding, which was true for both cultures, confirms our major hypothesis that physical co-presence influences the emotional expressions of children, and is in line with previous research (e.g. (Wagner and Lee, 1999)) arguing that the presence of other people influences the behaviour of participants.

The findings for the Dutch children are in accordance with the 'internalization of emotion' framework, which says that younger children express their emotions more clearly than older ones (Thompson, 1994), but this pattern was not found for the Pakistani children. An interesting question is what causes the difference between cultures that we found. While it is conceivable that differences in gaming experience partially explain these findings, we believe they reflect real differences between the two cultures. One argument for this is that the exact same pattern of our cross-cultural findings (Dutch signal negative emotions more clearly, Pakistani signal positive emotions more clearly) was also encountered in an analysis of how adult speakers signal positive and negative emotions, using a completely different setting, namely producing emotionally laden sentences rather than playing a game (Shahid et al., 2008). Systematic comparisons of western European and south Asian expression and recognition of emotion are scarce, with the notable exception of (Elfenbeinet al., 2002) who compare the emotion recognition patterns of American, Japanese and Indian adults.

Comparing Figure 13 and 14 suggests that Pakistani children respond more enthusiastically to winning than Dutch children, and this is indeed one noticeable difference between the two cultures. A qualitative analysis of the recordings revealed three further behavioural differences. First of all, it appeared that the Pakistani children gesture more than the Dutch children. In particular, Pakistani children frequently produce a number of culture specific gestures, most notably a praying gesture, as shown in figure 15. Pakistani children prayed quite often, especially before the last card of a game was shown. Dutch children, by contrast, never prayed. This, incidentally, might be a further explanation for the observed cross-cultural difference; if one prays for a positive outcome (winning), it would be inappropriate to show too much disappointment when losing.



Figure 15: Pakistani 12 years old praying before the final card is shown in the left picture, red square highlights the praying gesture

Another noteworthy finding is related to the fact that the Pakistani 12 years old girls (especially the individuals taken from pairs) are not only more expressive than their Dutch counterparts but also more expressive than the 8 years old Pakistani girls and boys, an illustration of which can be found in Figure 16.



Figure 16: 12 year old Pakistani girls winning

This is an interesting finding, which, at first sight, seems to be incompatible with expectations based on the age theory (Thompson, 1994) that younger children are more expressive than older ones, though it is consistent with the outcome of the work by (Chapman, 1973) that girls are more expressive in showing their emotions than boys. This result is also in line with another study where it was shown that females were more expressive in the presence of a friend than in the presence of stranger (Wagner and Lee, 1999). We conjecture that our findings might be due to culture-specific display rules, which dictate that Pakistani girls from a certain fairly young age on are expected to behave in a mature and calm way when expressing their emotions in a public setting. They may compensate for this in game situations like the ones we analysed.

A final noteworthy difference is that the social and physical bonding while playing together appears to be stronger in Pakistani pairs than in Dutch pairs. Pakistani children took more time in coming to a decision than their Dutch counterparts, and also frequently tried to sit closer to each other (they tended to move their chair closer to that of their partner, even when the sitting arrangement was deliberately fixed). Irrespective of age, Pakistani children touched, hugged and even hit (mostly in a polite manner) each other more often than Dutch children (as shown in figure 17).





Figure 17: 8 year old Pakistani girls hugging each other after loosing a game (left), 12 year old Pakistani boy hitting his friend after loosing a game (right).

In fact, the Dutch children never hit each other and hardly ever hugged after winning or losing a game. One important thing to note is that the Pakistani hugging and hitting only happened among same sex pairs. A minority of the 12 year old pairs consisted of members of both sexes, and in these cases a certain physical distance was maintained, albeit without effecting the overall expressiveness of children. In the case of 8 year olds, opposite sex pairs occasionally touched each other but not as common as in the 8 year old same sex pairs. This is most probably due to cultural-specific social norms of Pakistani society. At a certain age, children (and especially girls) are encouraged to maintain some distance to the opposite sex, and close physical contact, such as hugging, would be considered odd. The detailed annotations of different visual behaviours will be presented in the Chapter 3.

Our findings are relevant for Human-Computer Interaction in general, and Affective Computing in particular, in two different ways. First, we have shown that computer games offer a natural and highly efficient way to collect affective data. The card game we developed is a straightforward implementation of an extremely simple game, certainly when compared to games currently on the market. Still, as this study has shown, children responded strongly, and with a variety of emotional behaviours, after winning or losing this game. Second, our results indicate that there may be considerable differences in how positive and negative emotions (as a response to winning or losing a game) are expressed, depending on whether they are produced by younger children or older ones, by boys or girls, by western European or south Asian children and by children who play alone or who play together with a friend. Since we know from earlier research that emotion recognition systems perform best when the training data matches the real, application data, it seems important to take factors such as age, culture and co-presence into account when designing affective applications.

# References

Annetta, L.A., Minogue, J., Holmes, S.Y., Cheng, M.-T., 2009. Investigating the impact of video games on high school students' engagement and learning about genetics. Computers & Education, 53(1), 74-85.

Baenziger, T., Grandjean, D., Scherer, K.R., 2009. Emotion Recognition From Expressions in Face, Voice, and Body: The Multimodal Emotion Recognition Test (MERT). Emotion, 9(5), 691-704.

Bailenson, J.N., Pontikakis, E.D., Mauss, I.B., Gross, J.J., Jabon, M.E., Hutcherson, C.A.C., Nass, C., John, O., 2008. Real-time classification of evoked emotions using facial feature tracking and physiological responses. International Journal of Human-Computer Studies, 66(5), 303-317.

Beale, R., Creed, C., 2009. Affective interaction: How emotional agents affect users. International Journal of Human-Computer Studies, 67(9), 755-776.

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

Chapman, A.J., 1973. Social facilitation of laughter in children. Journal of Experimental Social Psychology(9), 528-541.

Clavel, C., Vasilescu, I., Devillers, L., Richard, G., Ehrette, T., 2008. Fear-type emotion recognition for future audio-based surveillance systems. Speech Communication, 50(6), 487-503.

Coan, J.A., Allen, J.J.B., 2007. Handbook of Emotion Elicitation and Assessment. Oxford University Press.

Cowell, A.J., Stanney, K.M., 2005. Manipulating of Non-verbal Interaction Style and Demographic Embodiment to Increase Anthropomorphic Computer Chracter Credibility. International Journal of Human Computer Studies, 62, 281-306.

Darwin, C., 1972. The expression of the emotions in man and animals, 3rd ed. Oxford University Press, Oxford, UK.

Ekman, P., 1994. Strong evidence for universals in facial expressions: a reply to Russell's mistaken critique. Psychological Bulletin(115), 268-287.

Elfenbein, H., Ambady, N., 2002. On the universality and cultural specificity of emotion recognition. Psychological Bulletin(128), 203-235.

Elfenbein, H., Ambady, N., Mandal, M., Harizuka, S., 2002. Cross-cultural Patterns in Emotion Recognition: Highlighting Design and Analytical Techniques. Emotion(2), 75-84.

Fabes, R., Martin, C., 1991. Gender and age stereotypes of emotionality. Personality and Social Psychology Bulletin(17), 532-540.

Fridlund, A., 1994. Human facial expression: an evolutionary view. New York: Academic Press.

Gratch, J., Marsella, S., 2004. A domain-independent framework for modeling emotion. Journal of Cognitive Systems Research, 5, 269-306.

Isbister, K., Hook, K., Laaksolahti, J., Sharp, M., 2007. The sensual evaluation instrument: Developing a trans-cultural self-report measure of affect. International Journal of Human-Computer Studies, 65(4), 315-328.

Isomursu, M., Tahti, M., Vainamo, S., Kuutti, K., 2007. Experimental evaluation of five methods for collecting emotions in field settings with mobile applications. International Journal of Human-Computer Studies, 65(4), 404-418.

Kaiser, S., Wehrle, T., 1996. Situated emotional problem solving in interactive computer games. In N. Frijda (Ed.), Proceedings of the International Society for Research on Emotions, 276-280.

Kleinsmith, A., De Silva, P.R., Bianchi-Berthouze, N., 2006. Cross-cultural differences in recognizing affect from body posture. Interacting with Computers, 18(6), 1371-1389.

Lazzaro, N., 2004. Why We Play Games: Four Keys to More Emotion Without Story. Game Developer Conerence (GDC).

Lu, C.-Y., Lin, S.-H., Liu, J.-C., Cruz-Lara, S., Hong, J.-S., 2009. Automatic event-level textual emotion sensing using mutual action histogram between entities. Expert Systems with Applications, 37(2), 1643-1653.

Mandryk, R.L., Atkins, M.S., 2007. A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies. International Journal of Human-Computer Studies, 65(4), 329-347.

Manstead, A.S.R., 2005. The social dimension of emotions. The Psychologist, 18(8), 484-487.

Masuda, T., Ellsworth, P.C., Mesquita, B., Leu, J., Tanida, S., Veerdonk, E.D., 2008. Placing the face in context: Cultural differences in the perception of facial emotion. The Journal of Personality and Social Psychology, 94(3), 365-381.

Mesquita, B., 2001. Emotions in collectivist and individualist contexts. Journal of Personality and Social Psychology, 80(1), 68-74.

Mesquita, B., Frijda, N.H., 1992. Cultural variations in emotions: A review. Psychological Bulletin, 112, 179-204.

Mesquita, B., Leu, J., 2007. The cultural psychology of emotions, in: Kitayama, S., Cohen, D. (Eds.), Handbook for cultural psychology. New York: Guilford Press.

Oatley, K., 2000. Emotion and Distributed Cognition. Invited talk to the 2000 Annual Conference of the British Psychological Society, Summarised by Paul Redford in 'The Psychologist', 13(6), 290-229.

Pantic, M., Rothkrantz, L.J.M., 2003. Toward an affect-sensitive multimodal human-computer interaction. Proceedings of the IEEE, IEEE Press, 1370 - 1390.

Picard, R., Vyzas, E., Healy, J., 2001. Towards Machine Emotional Intelligence: Analysis of Affective and Physiological State. IEEE Transactions on Pattern Analysisand Machine Intelligence, 23(10), 1175-1191.

Picard, R.W., 2003. Affective computing: challenges. International Journal of Human-Computer Studies, 59(2), 55-64.

Picard, R.W., Klein, J., 2002. Computers that recognise and respond to user emotion: theoretical and practical implications. Interacting with Computers, 14(2), 141-169.

Rime, B., Mesquita, B., Philippot, P., Boca, S., 1991. Beyond the emotional event: Six studies on the social sharing of emotion. Cognition & Emotion, 5, 435-465.

Salen, K., Zimmerman, E., 2003. Rules of Play, Game Design Fundamentals. The MIT Press, Massachusetts.

Sarrafzadeh, A., Alexander, S., Dadgostar, F., Fan, C., Bigdeli, A., 2008. "How do you know that I don't understand?" A look at the future of intelligent tutoring systems. Computers in Human Behavior, 24(4), 1342-1363.

Sattar, A., Stoiber, N., Seguier, R., Breton, G., 2009. Gamer's Facial Cloning for Online Interactive Games. International Journal of Computer Games Technology.

Shahid, S., Krahmer, E., Swerts, M., 2008a. GamE: Game as a method for Eliciting Emotions. Proceedings of the Measuring Behaviour Conference (MB 2008), Maastricht, the Netherlands.

Shahid, S., Krahmer, E., Swerts, M., 2008b. Real vs. acted emotional speech: Comparing Caucasian and South-Asian speakers and observers. Proceedings of the Speech Prosody, Brazil, 669-672.

Shahid, S., Krahmer, E., Swerts, M., Melder, W.A., Neerincx, M.A., 2009. Exploring social and temporal dimensions of emotion induction using an adaptive affective mirror. Proceedings of the 27th international conference extended abstracts on Human factors in computing systems, Boston, USA, ACM, 3727-3732

Simon, H.A., 1967. Motivation and emotional controls of cognition. Psychological Review, 74, 29-39.

Thompson, R., 1994. Emotion Regulation: A Theme in Search of Definition. The Development of Emotion Regulation.

Tomkins, S., 1962. Affect, inagery, consciousness. Springer, New York.

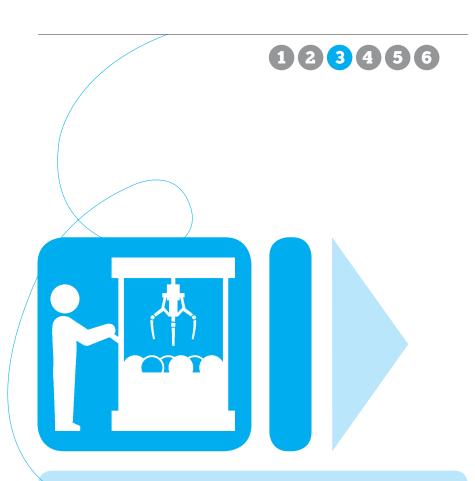
Tractinsky, N., 2004. Tools over solutions? comments on Interacting with Computers special issue on affective computing. Interacting with Computers, 16(4), 751-757.

Vogt, T., André, E., 2005. Comparing feature sets for acted and spontaneous speech in view of automatic emotion recognition. Proceedings of IEEE International Conference on Multimedia & Expo (ICME 2005), Trento, Italy.

Wagner, H., Lee, V., 1999. Facial behavior alone and in the presence of others. In P. Philippott et al. (Eds.), The social context of nonverbal behavior, New York: Cambridge University Press.

Wang, N., Marsella, S., 2006. Introducing EVG: An emotion evoking game. the Proceedings of the Intellignet Virtual Agents (IVA), Lecture Notes in Computer Science, Springer, 282-291.

Ward, R., 2005. What is affective sensing for, In the Workshop on the Role of Emotion in HCI (HCI 2005). BCS.



# CHAPTER 3

Child-robot Interaction: How does playing a game with a social robot compare to playing a game alone or with a friend?



# Child-robot Interaction: How does playing a game with a social robot compare to playing a game alone or with a friend?

# ABSTRACT.

In this chapter we investigate how children in different age groups (8 and 12 year olds) and belonging to different cultural backgrounds (Pakistani vs. Dutch) experience interacting with a social robot (iCat) during collaborative game play. We propose a new method to evaluate children's interaction with such a robot, by asking whether playing a game with a social robot is more similar to playing this game alone or with a friend. A combination of self-report scores, perception test results and behavioral analyses indicate that child-robot interaction in game playing situations is highly appreciated by children, although more by Pakistani and younger children than by Dutch and older children. Results also suggest that children had more fun playing with the robot than playing alone, but had more fun still when playing with a friend. Children were more expressive when playing with the robot than they were when playing alone, but less expressive than when playing with a friend. Our results not only stress the importance of using new benchmarks for evaluating child-robot interaction but also highlight the significance of cultural differences for the design of social robots.

A journal paper based on this chapter is submitted for publication. Earlier versions of this chapter appeared as: Shahid, S., Krahmer, E., and Swerts, M., 2011. Child-robot Interaction: Playing Alone or Together? Proceedings of the ACM annual conference on Human factors in computing systems (CHI EA 11), 1399-1404, as Shahid, S., Krahmer, E., Swerts, M., and Mubin, O., 2011. Who is More Expressive During Child-Robot Interaction: Pakistani or Dutch Children?, Proceedings of the IEEE/ACM annual conference on Human Robot Interaction (HRI 11), and as Shahid, S., Krahmer, E.J., Swerts, M.J.G., and Mubin, O. (2010). Child-Robot Interaction during Collaborative Game Play: Effects of Age and Gender on Emotion and Experience. Proceedings of the ACM OzCHI 2010

# 1. Introduction

In the last decade, there has been an increased interest in designing robots that have the ability to build an interpersonal relationship with humans by exchanging social and emotional cues (Breazeal, 2002). Unlike traditional research on service robotics where humans were seen as obstacles and robots were supposed to avoid them while performing their tasks, the new breed of robots is deliberately designed to interact and cooperate with humans both for serious and entertaining purposes (Fong et al., 2003; Lee et al., 2004). These robots are called social robots and are used increasingly not only in entertainment and education, but also in rehabilitation and therapy.

Social robots may have a particularly strong impact for specific user groups. In particular, children are likely to have the potential to get substantial benefits from robotic devices in many ways (Woods, 2006). Child-robot interaction (CRI) in itself is emerging as a research area and researchers are particularly motivated by the possibilities of letting children interact and collaborate with a robot in a social and intuitive way, which ideally could be highly similar to how they interact with their peers. The relationship and type of interaction between children and robots is fundamentally social and children probably do not use robots merely as a tool in a task-oriented manner (Salter et al., 2010).

Keeping in mind the social nature of child-robot interaction, researchers have started asking key questions about the societal consequences and acceptance of robots, especially when focusing on younger users. How do children perceive and interact with these robots? How do they establish a social bond with them? Can these social robots provide children with similar psychological outcomes and levels of interaction as human partners do (Salter et al., 2008)? A complication is that norms of social behavior differ from culture to culture, and as a consequence it is conceivable that different forms and social behaviors of robots may generate different responses across cultures (Nomura et al., 2007). Now, when the world is turning into a global village and the market for robots is rapidly expanding and they are starting to reach diverse cultures and user groups, it is crucial to investigate how children with different cultural backgrounds respond to a social robot.

Existing research on how children perceive and interact with robots, particularly across cultures, is inconclusive. A few studies have shown that children like robots as companions and that certain elements of human-human communication are replicated in human-robot communication, but others have shown that there are differences in how people collaborate with human and human-like (i.e., robotic) partners (Castellano et al., 2010). Only a very limited number of studies have looked at human-robot interaction from a cross-cultural perspective, and these studies suggest that the type and richness of social relations that people have with robots vary considerably across cultures. In general, it seems fair to say that in the area of human-robot interaction, and CRI in particular, most of the research is conducted in developed countries, with hardly any studies paying attention to the cultures of developing countries (with Bartneck et al., 2005, as a notable exception). Finally, CRI as a research area is still in its infancy and previously most of the research in HRI has focused on adult participants who mainly interacted with robots in a task-oriented manner. Children are different from adults in many ways. In fact, even children of different age groups are different from each other, and little research exists on how children of different age groups (and different cultural backgrounds) interpret and interact with robots in playful settings.

To address these various issues, we developed a new experimental paradigm in which children of different age groups (8 vs. 12 years) and cultures (Pakistani vs. Dutch) play the same game either alone, with a friend or with a social robot as their game partner. We are interested in whether playing a game with a modern social robot is more similar to playing this game alone or with a friend, and in how children of different cultures establish a social bond with a robot and exchange (non)verbal cues during playful interactions.

### 1.1 Child-Robot Interaction

Child-robot interaction is an increasingly important application area for robotic research, and the number of robots and range of applications specifically designed for children is rapidly increasing. A key research question addressed in this field is whether and how a robot can be used as a medium for social interaction. The fundamental challenge is to identify means by which children interact with robots and design the overall child-robot interaction experience in such a way that it closely resembles the human-human interaction experience (Francois, 2009).

Robot technology is being used for children, for example, in the education sector where robots not only act as assistants to teachers but also as autonomous agents in a school environment (Kanda et al., 2004). Various studies have looked at robots in education, and the results of different studies are mixed (Han et al., 2008; Oh and Kim, 2010). In general, it is shown that children are able to develop a social bond with educational robots over time and also report a positive learning experience, but factors like participants' age, cultural backgrounds and duration of interaction can have a large influence on the results. Another important application area for robot technology is in child development, rehabilitation and healthcare (Baxter et al., 2011; Kozima et al., 2007; Robins et al., 2004). Studies in this area show that robot technology can assist children with physical disabilities to move and play, and can help children with autism learn about communication and emotions (Conn et al., 2008). Studies such as these also confirm that the success of learning and rehabilitation technologies largely depends on individual differences (including age and mental capacity) and the context in which these technologies are used (Francois, 2009; Salter et al., 2008).

Robots are also being designed for purely entertaining and playful purposes, and different studies have investigated children's attitudes towards this application of robots. Sony's AIBO is an example of an autonomous entertaining home robot, which is often used in HRI and CRI studies (Bartneck et al., 2007; Kerepesi et al., 2006). It is designed to elicit socio-emotional responses and shows learning and growth abilities. The iCat is another animal-like robot frequently used in HRI and CRI studies, developed for in-home usage (Castellano et al., 2010). It is designed to show six basic emotions and elicit rich socio-emotional responses (Verhaegh et al., 2007).

Evaluating how social or entertaining robots are is a challenging issue in human-robot interaction research and it is even more difficult when special user groups, such as children are involved. A well-known finding from the HCI literature is that, in many situations, humans treat computers as social actors (Reeves and Nass, 1996). Unlike traditional computers, which exhibit relatively few cues that could be seen as socially intelligent, social robots are "borderline objects" that explicitly mimic many properties of living beings. It seems plausible that the findings of Reeves and Nass apply even stronger to social robots. If this is the case, then an interesting question is to what extent this applies to children and how children of different age groups treat a social robot during an

interactive game play.

For investigating the social aspects of child-robot interaction, different evaluation methodologies have been used. In one type of experimental setup, child-robot interaction is compared with child-animal and child-toy interactions (Pepe et al., 2008; Ribi et al., 2008). Such investigations are uncommon but indeed help in identifying where on the "social" continuum these 'borderline objects' fall in terms of interaction. For our study, we relied on a similar paradigm, where the novelty is that we focus on a human partner instead of an animal partner or a toy. A human partner is, obviously quite different from an animal partner or a pet (Mitchell and Hamm, 1997). Earlier research on child-animal relationship shows that children associate different attributes and expect a different kind of a response from a pet than from a peer (Bursie, 2004).

In one study, (Kahn et al., 2006) looked at the differences in children's responses to a stuffed dog and a robotic dog. They reported that the children engaged in imaginary play with the robotic dog in a similar fashion to how they engaged with the stuffed dog. However, based on a behavioral analysis, they reported that the quality and type of interactions with these two artifacts varied. Children showed more reciprocal and apprehensive behavior towards the robotic dog, and displayed less cautious behavior to the stuffed dog, which also included more mistreatment. In another study, (Melson et al., 2005) investigated the interactions of children with a robotic dog (AIBO) in comparison to a live dog (an Australian shepherd). The results showed that children preferred to stay in closer proximity to the live dog than to the robotic dog, and were more inclined to associate sociality, morality and mental states to the live dog. They also touched and engaged in physical activity more with the live dog than with the robotic dog. Another study (Kerepesi et al., 2006), compared the interaction of children and adults with a real and with a robotic dog and showed that the robot was appreciated as an equally affective playing partner as the dog puppy.

In another type of experimental setup, researchers compared different types of robots (e.g., iCat vs. Nao, a full-body robot) (Baxter et al., 2011), robots with virtual agents (Looije et al., 2008), or different attributes of the same robot (e.g., aggressive vs. polite) (Fussell et al., 2008). In one such study, (Vink, 2011) compared a socio-cognitive variant of the iCat (the robot behaved socially and took the mood of the child into account) with an ego-reactive iCat (the robot did not take the mood of the child into account) with children. The results showed that children were happier when interacting with the socio-cognitive iCat than with the ego-reactive iCat.

Taken together, the results of these various studies present an unclear picture. On the one hand children are able to differentiate a robot from a toy and a pet, but on the other hand, they have mixed responses and attitudes towards robots. Furthermore, in surveys and questionnaires children indicate that a robot is like a companion or a family member, but behavioural analyses show that children clearly respond differently to living objects, such as a pet dog, than to a robot. In general, differences between the various studies in methodology and set-up make it difficult to draw general conclusions.

Therefore, in this chapter, we focus on young children of different age groups in a game-playing context and investigate how children play collaborative games and establish a bond with a social robot. We compare the child-robot interactions to situations where children either play the same game alone or together with a friend. This experimental setup helps to understand where on the

continuum CRI falls; is it more like a truly non-social situation (playing alone) or more like a normal social situation (playing with a friend). In addition, we will study these issues from a cross-cultural perspective, comparing Dutch and Pakistani children for reasons discussed in the next section. This comparison is interesting, since it contrasts a developed and a developing country, which may have an impact on how participants from these cultures appreciate the interaction with a robot. In general, there may be differences in the way such different cultures respond to robotic technology due to the their exposure to robots in everyday life, social norms, religious beliefs, and the way their media portrays robots.

# 1.2 Effect of Culture

"Culture is a fuzzy set of attitudes, beliefs, behavioural conventions, and basic assumptions and values that are shared by a group of people, and that influence each member's behaviour and each member's interpretations of the 'meaning' of other people's behaviour" (Spencer-Oatey, 2000). What may be logical to do in a particular social setting of one culture might be impolite or even rude in a similar social setting of another culture. Different cultures have different social norms for responding to different social situations and for showing emotional and other non-verbal expressions (Elfenbein et al., 2002). This suggests that it is important to look at different cultures when studying child-robot interaction.

Cross-cultural differences have been investigated across many disciplines and the field of HCI is no exception. In the HCI literature, it has been shown that people belonging to different cultural backgrounds respond to computers and technology in different ways (Oshlyansk, 2007). In one cross-cultural study, (Weil and Rosen, 1995) showed cultural differences on "technophobia," especially when it comes to advanced technology. In another cross-cultural study, (Shinozawa et al., 2003) investigated social responses to lifelike characters, showing that Japanese participants had stronger emotional responses (both negative and positive) to an onscreen robot than to a physical, 3D robot, whereas the opposite held for American participants.

As shown in the previous chapter, we conducted a comparable cross-cultural study where children of two age groups and belonging to two cultures (Pakistan and Dutch) played computer games in different social settings (Shahid et al., 2008a). The results showed that the Pakistani children were more expressive than their Dutch counterparts. Furthermore, we also found that Pakistani 8 and 12 year olds behave differently from Dutch 8 and 12 years olds during interactive gameplay. Here an important question arises: do these cross-cultural differences between children of different cultures (developed and developing countries) and age groups exhibited during human-computer interactions also hold for human-robot interaction? This question is addressed in this chapter.

In the area of human-robot interaction (HRI), a number of cross-cultural studies have been conducted which show that different cultures appreciate robots differently (Shibata et al., 2003, 2004), (Bartneck et al., 2005), (Nomura et al., 2007). Most of these studies are conducted with adults and there are only few examples where CRI is investigated across cultures, and even then the focus is sometimes on the parents rather than the children. In a couple of studies, (Cho et al., 2008) used an educational robot and compared child-robot interaction in a school environment across cultures. The results suggested that European parents seem to have a more negative view on educational robots than Asian ones. Korean parents have a strong tendency to regard these robots as 'friends of

children'. In this study, the opinion of children was not directly asked, as the focus was on the parent's attitudes. In an extension of this study, (Han et al., 2009) showed that Spanish and Japanese parents hold more conservative attitudes towards tutoring robots than Korean parents. Findings such as these findings suggest that even within developed countries people respond to robots in a different manner.

Children from western cultures and cultures of developed countries are more likely to have been exposed to robots through toys, television shows, cartoon, and games. Japan is an extreme example, where robotic culture has been embedded into regular culture, and Japanese children are more likely to see robots as friends than children from other developed cultures (Han et al., 2009). Compared to developed countries, children from developing countries have had only limited exposure to robotic devices and it seems likely that this will influence their expectations about robots and the way they treat them. What this influence will be remains unclear; it is conceivable that they would appreciate interacting with a social robot more, because it is newer for them, but it could also be that they will be more restrained in their interactions, precisely because of this novelty.

In some studies, a number of previously found cross-cultural differences in human-human interaction are tested during human-robot interaction. In one such study, (Evers et al., 2008) showed that differences in cultural backgrounds can lead to varied responses to autonomous robots. They showed that when a robot was presented as an in-group member (forming a team with the participant), Chinese participants were more satisfied with the interaction and reported a stronger sense of control compared to US participants. In another study, (Wang et al., 2010) reported that Chinese participants were likely to trust a robot and show compliance with its recommendations if the robot communicated more implicitly, while the reverse was true for American participants. There are only a handful of HRI studies where researchers compared participants of developed and developing countries. The results of one such study showed that respondents from the USA (a developed country) were more positive about robots than respondents from Mexico (a developing country). They also showed that, contrary to the popular belief that Japanese love everything about robots, Japanese participants were actually concerned about the emotional aspects of interacting with robots (Bartneck et al., 2007).

In general, cross-cultural studies such as those discussed above show the diversity of responses to robots across different cultures, and that findings from cross-cultural psychology about the differences in human-human interaction may also hold for human-robot interaction. One key limitation of existing research on robotics is that most of this research is conducted in developed or western countries. In the case of Asian countries, most researchers still focused on developed countries. Results of different studies conducted in few developed Asian countries cannot be generalized for 'Asia' as a whole, given that differences between Asian cultures can be huge (Aubrey, 2009). Furthermore, many of the existing cross-cultural HRI studies are either based on surveys or on self-report questionnaires and only few studies used behavioral analyses for identifying the complex differences across cultures.

In this study, we ask whether the cultural background of children has an influence on how they interact with a social robot, focusing on a comparison of Dutch and Pakistani children of different age groups during an interactive gameplay.

# 1.3 The Current Study

In this chapter, we study child-robot interaction using an experimental paradigm in which children (both boys and girls, either 8 or 12 years old) played a simple card game in one of three conditions: alone, with a friend, or with a social robot. After each game, subjective fun and overall game experience ratings were collected from all children (Experiment 1, Section 2). In addition, all game playing interactions were videotaped and randomly selected fragments were subsequently presented in a perception test to human judges, allowing us to objectively compare the non-verbal behavior of the children in the various conditions, and see whether these were in accordance with the selfreports (Experiment 2, Section 3). Finally, we did an observational analysis (Section 4) in which we looked at different visual (positive and negative) features and social cues shown during different game playing conditions.

# 2. Experiment 1: Game Playing

We used a card guessing game as a tool for evaluating child-robot interaction and for investigating the influence of a robot's presence as a game partner on the emotional response and social experience of children. The card guessing game was developed under the GAME (Game As a Method of eliciting Emotions) paradigm. Under this paradigm, the objective is to design simple games which can be used as tools for inducing emotions and eliciting rich social interactions in a natural and ethical way (Shahid et al., 2008b). Games are a popular medium to investigate human-robot interaction (Castellano et al., 2010; Short et al., 2010). Compared to traditional experimental settings of eliciting rich socio-emotional behaviour, games are more interactive in nature, but still allow control (to both players and experimenters) to create a natural ambience (Kaiser and Wehrle, 1996).

### 2.1 Method

# 2.1.1 Material

In this study, we use the same game materials (the card guessing game) as in the previous chapter. In the card game, children have to guess whether an upcoming number is higher or lower than the previous (reference) number. When the game starts, players see a row of six cards on the screen where the number of the first card is visible ('3' in the case of the example in Figure 1) and the other five cards are placed upside down so the numbers are hidden. The numbers on the cards range from 1 to 10, and a number displayed once is not repeated in a particular game. Players have to guess whether the number on the next card will be higher or lower than the previous number. Once players have guessed the number, the relevant card is turned around and the correct answer is shown on the screen. Players are also informed about the correctness or incorrectness of their answer via a characteristic non-speech feedback sound. If players make a wrong guess, the game is finished and they move to the next game. If players guess the number correctly, they are asked to guess the next number. Players only win a game if they guess all numbers in the row correctly. In the case of winning a game, they receive a coin. At the end of the experiment, which consisted of multiple games, players could exchange the coins they won for a prize.

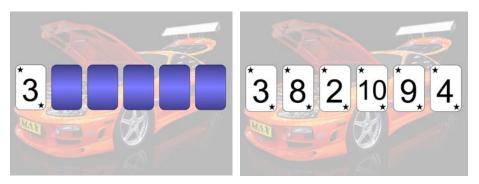


Figure 1: Winning Scenario - Start of a game (left) and end of the game (right)

The card games were developed using Microsoft\* PowerPoint\*. Appropriate colourful images were chosen for the game background and different animations were used to turn cards around in an attempt to make the game attractive for children. During the experiment, children played six games, and could in theory win six coins. However, unknown to the children, each game was completely deterministic, and two different game variants were used. In the first alternative (Figure 1), a rational decision procedure would result in winning the game, and in the second alternative, being rational would result in losing the game. Figure 2 is an example of the latter: the most probable outcome for the final card would be a "lower" number than 9, but guessing "lower" would make this a losing game. The two other losing games were: 8-3-9-7-2-1 and 1-9-4-8-3-2. Winning scenarios were 9-2-8-1-10-7, and 3-8-2-10-9-2 and 1-3-9-2-7-6. Winning and losing games were mixed in the sequence, starting and ending with a variant in which children were likely to win.

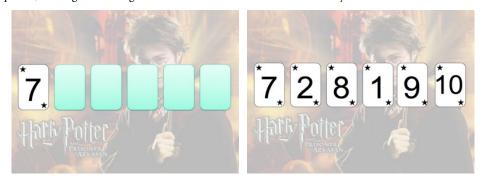


Figure 2: Losing Scenario - Start of a game (left) and end of the game (right)

### **Participants**

112 children, 52 Dutch and 60 Pakistani, participated in this study and played the game with the iCat. For comparisons we also used data from our earlier study as described in chapter 2 (Shahid et al., 2011). In this earlier study, 144 children participated, of which 48 children played the games individually (24 Dutch and 24 Pakistani ones), and 96 children played the games in pairs (24 pairs of Dutch children and 24 pairs of Pakistani children). All together this resulted in 256 participants spread over three conditions (alone, iCat, friend). In all conditions, half of the children were around

8 years old (group 4 in the Dutch elementary school system, second grade in the Pakistani school system) and the other half were around 12 years old (Dutch group 8, Pakistani sixth grade). Parents and teachers gave prior written permission for their child to participate, and signed a consent form stating that the recordings could be used for research purposes.

### 2.1.3 Procedure

The experiment was conducted in six elementary schools, four in Tilburg (the Netherlands) with Dutch participants and two in Lahore (Pakistan) with Pakistani participants. The procedure for the iCat condition was essentially similar to the previous conditions (playing individually or playing in pairs). In all conditions, a separate room was chosen in each school, where depending on the condition, pairs of children (self-selected pairs of the same age group) or individual children (playing game alone or with iCat) were invited and asked to sit on the chair(s) placed in front of a desk on which a large computer screen was placed. In the iCat condition, children sat next to the iCat to keep the sitting position similar to the 'pairs' condition. The iCat was positioned in such a way that by turning its head it could look both at the screen and the child allowing for eye contact (see Figure 3). Half of the children were positioned on the right side of the iCat, and half on the left. The game was displayed on the computer screen placed in front of the participant(s). On top of the computer screen, a camcorder was positioned in such a way that it could record the children's faces and the upper part of their body. A laptop was attached to the screen to facilitate the experiment leader in controlling the game. In the case of the iCat condition, an additional computer was also used for controlling the verbal and non-verbal response of the iCat. The experimenter was always outside the visual field of the game-playing children. Once the children were in the room and had chosen the appropriate seat(s) for sitting, the experimenter welcomed the children and started a small talk by asking a few questions to break the ice (e.g., "How old are you? Do you like to play games?"). In the iCat condition, the children were first informally introduced to the iCat and the role of the iCat as a game partner was explained to them.





Figure 3: Experimental setup for the 'child-robot interaction' condition

After this introductory phase, the experimenter gave spoken instructions, telling the children about the game and the coins they could win. All the game rules outlined in the previous section were explained to the children (in Urdu and Dutch for Pakistani and Dutch children, respectively), and when they seemed to understand the rules, the experimenter started a practice game ("So you only have to say whether the next card is higher or lower. This is just an exercise and it doesn't really

count"). In the case of the iCat condition (as in the pairs condition), participants were encouraged to discuss the outcome with the iCat before giving the final answer. When the practice session was finished, the experimenter asked the children whether they had any questions, and if not the experimenter left the children's field of vision and started the first experimental game.

The experiment did not have a fixed duration because children varied in the time they needed to make a decision, and some took more time to negotiate with their partners or iCat on the decisions. On average each experimental session lasted between approximately 10 and 15 minutes. After the sixth and final game, the experimenter guided the children to another room where they had to fill in a post-experiment questionnaire. After this, the experimenter asked the children some open questions in an interview. At the end of the whole session, the experimenter congratulated the children and informed them that they could trade in their coins for an individual gift (a whistle, small plastic ball, or chocolates in the case of the Dutch participants and chocolates/juices, geometry boxes or key rings for the Pakistani participants).



Figure 4: Children's reactions after winning or losing a game in playing alone (top), playing with iCat (middle) and playing with friend (bottom) conditions

Figure 4 shows children's reactions after winning or losing a game in different game playing conditions. Overall, the game worked well, in that all children made the logical choices they were expected to make in most of the games, so that each child and pair of children lost at least two games and won at least two games. Furthermore, not a single child reported that the game was in fact a fake, i.e., a deterministic simulation.

### **Child-robot Interaction** 2.1.4

We used a Wizard of Oz method to simulate both the verbal and non-verbal behavior of the iCat. The wizard was located outside of the child's field of vision and relied on simple pre-programmed behaviors of the iCat in a contextually appropriate way. A text-to-speech engine was also employed in order to synthesize the vocal responses of the iCat. We designed a number of game-specific scripts, which were controlled by the wizard as well.

The regular game scripts included realistic utterances from the iCat, in such a way that for every higher or lower suggestion of the child ("I think the next card will be higher"), the wizard had possibilities to start a dialogue about the next guess ("I think lower") or to agree with the child ("I think so too"). In all cases, we made sure that the iCat would initiate the discussion in half of the cases ("I think the next card will be higher, what do you think?") and waited for the child in the other half. In addition, we made sure that the robot made correct and wrong suggestions an equal number of times. The iCat also showed certain and uncertain behaviour ("I think bigger but I am not sure") while guessing the next number. In our script and design of interactions with the iCat, we also used some design patterns (e.g. reciprocal turn taking in game context) for social human-robot interaction proposed by (Kahn et al., 2008). In general, developing natural and representative scripts was facilitated by (a) the deterministic nature of the game, which made it possible to make "educated guesses" about what children are likely to say and (b) the fact that we used this same game in our earlier studies, from which we learned how children were likely to express themselves during different phases of the game. This formed the basis for the scripting of the iCat's verbal and nonverbal behavior (see Figure 5).



Figure 5: iCat showing frown expression (with head down, lower eyelids, curved lips) and the child showing sympathy to iCat after losing a game

### 2.1.5 Measurements

For measuring the experience of fun after playing games in three conditions (alone, iCat, friend), we used different variants of the 'funometer' to assess children's experiences (Read, 2008). For analysis, scores of all fun measurement were mapped on a 1-7 point likert scale where 1 was no fun at all and 7 represents a lot of fun. We also videotaped all interactions for further analysis.

In the iCat condition, we used various additional measurements. We used the pictorial card sorting method for ranking children's preference about playing alone, playing with friends, playing together with a stranger and playing with iCat (Markopoulos et al., 2008). For evaluating the children's social experience, we adapted the Game Experience Questionnaire (GEQ) (IJsselsteijn et al., 2007). The adapted questionnaire contains sixteen questions distributed over four categories, 1) endurability (e.g., "I would like to play this game again"), 2) engagement (e.g., "the game was exciting, I was fully involved in the game"), 3) iCat likeability (e.g., "I liked playing with iCat, the iCat was friendly") and 4) iCat smartness (e.g., "I think the iCat was smart", "iCat couldn't help me in this game as good as my real friend would do") all on 5 point Likert scales. All questionnaires were translated in Urdu for Pakistani children and in Dutch for Dutch children.

# 2.2 Data Analysis and Results

To gain some first insight into how children had experienced the game, we compare the reported fun scores across all three conditions. To test for significance we conducted a univariate analysis of variance (ANOVA) with three between-subject factors, namely Partner (levels: Alone, iCat, Friend), Culture (levels: Pakistani, Dutch) and Age (levels: 8 years, 12 years) and with the mean fun score as the dependent variable. The Bonferroni correction was used for pairwise comparisons.

Culture	Age	Conditions		
		Alone	With iCat	With Friend
Dutch	8 years	5.00 (.74)	6.04 (.89)	6.22 (.61)
	12 years	4.63 (.64)	5.88 (.67)	6.04 (.66)
Pakistani	8 years	5.60 (.51)	6.32 (.62)	6.50 (.51)
	12 years	5.09 (.53)	6.12 (.52)	6.67 (.48)

Table 1: Average fun scores with standard deviation (*SD*) on a 7-point Likert scale (1= very negative, 7= very positive)

Table 1 summarizes the results. The ANOVA showed a main effect of partner (F (2,220) = 61.26, p < .001,  $\eta^2_p = .42$ ): the children reported having the most fun while playing with friends (M = 6.40, SD = .61), least fun while playing alone (M = 5.10, SD = .70) and their rating of fun while playing with the iCat was in between (M = 6.01, SD = .70). All pairwise comparisons are statistically

significant (p < .001). Note that the difference in means between the iCat and Alone conditions ( $M_d$ = 0.81) is about twice as large as that between the iCat and Friend conditions ( $M_d = 0.39$ ). The Pakistani children reported having overall more fun (M = 6.10, SD = .56) than Dutch children (M = 6.10, SD = .56) than Dutch children (M = 6.10). 5.63, SD = .73), F(1,220) = 21.66, p < .001,  $\eta^2_p = .09$ . Furthermore, the 8 years old children reported having more fun (M = 5.94, SD = .68) than the 12 years old (M = 5.74, SD = .59), F(1,220) = 5.336, p< .05,  $\eta_p^2$  = .02. No significant interactions were found.

Next we analyzed the GEQ questionnaire by zooming in on the iCat data. In this case, four separate univariate analyses of variance were performed for the four different Game Experience Questionnaire categories (engagement, endurability, likeability, iCat smartness) with two between factors, namely Culture (levels: Pakistani, Dutch), and Age (levels: 8 years, 12 years). The mean score for each GEQ category was the dependent variable. For the purpose of the current analyses, the values of negatively formulated questions were reversed, so that high numbers always represented a positive score.

Table 2 summarizes the results of all four categories. The two-way ANOVA for the 'engagement' dimension revealed that overall Pakistani children were more engaged in the game play while playing with the iCat (M = 4.46, SD = .67) than Dutch children (M = 4.10, SD = .78), F(1,96) =6.906, p < .05,  $\eta^2 p = .067$ . Furthermore, 8 year old children found playing with iCat more engaging (M = 4.56, SD = .50) than 12 year old children (M = 4.00, SD = .86), F(1,96) = 16.710, p < .001,  $\eta^2 p = .001$ .148.

The two-way ANOVA for the 'endurability' dimension revealed a similar pattern. Overall Pakistani children showed a stronger interest in playing again with iCat (M = 4.30, SD = .75) than Dutch children (M = 3.90, SD = .79), F(1,96) = 6.894, p < .05,  $\eta^2 p = .067$ . Furthermore, 8 year old children were more unhappy when the game was over (M = 4.44, SD = .58) than 12 year old children (M =3.72, SD = .81), F(1.96) = 27.574, p < .001,  $\eta^2 p = .223$ .

Culture	Age	GEQ Dimensions			
	-	Engagement	Endurability	Likeability	Smartness
Dutch	8 years	4.40 (.70)	4.28 (.61)	4.32 (.69)	4.36 (.63)
	12 years	3.80 (.91)	3.52 (.77)	3.96 (.45)	4.76 (.43)
Pakistani	8 years	4.72 (.45)	4.60 (.50)	4.56 (.51)	4.12 (.73)
	12 years	4.21 (.76)	3.92 (.77)	4.08 (.70)	4.48 (.59)

Table 2: Average Engagement, Endurability, Likeability and Smartness scores (with SD) on a 5-point Likert (1= very negative, 5= very positive) for Dutch and Pakistani children.

Regarding likeability of iCat, 8 years old children liked the iCat more (M = 4.54, SD = .61) than 12 year old children (M = 4.02, SD = .58), F(1,96) = 12.307, p < .001,  $\eta^2_p = .114$ . Interestingly there was no effect of culture (F<1), which implies that Pakistani (M = 4.32, SD = .65) and Dutch (M = 4.19, SD = .60) equally liked the iCat. Also, compared to 8 year old children (M = 4.24, SD = .69), 12 year old children found iCat more intelligent (M = 4.62, SD = .53), F(1.96) = 9.845, p < .01,  $\eta^2_p =$ .093. Furthermore, Dutch children rated the smartness of iCat higher (M = 4.56, SD = .57), than Pakistani children (M = 4.30, SD = .67), F(1.96) = 4.609, p < .05,  $\eta^2_p = .046$ .

Finally, we report on the pictorial card sorting test where children could rank four alternatives (playing alone, playing with friend, playing with a stranger and playing with the iCat) in terms of their preference. To analyze the results we used the Friedman Test, and to test for pairwise significance we used the Wilcoxon Signed-Rank Test. Both tests were run separately for both cultures. For this analysis, we mapped the card ranking on a 1 to 4 scale where 1 means the card was ranked as the first choice and 4 means the card was ranked as the last choice (see Figure 6). The result of Friedman Test for Pakistani children showed that playing with friend was ranked on top (M=1.46), followed by playing with iCat (M=1.66), then playing with stranger (M=3.19) and playing alone was ranked the lowest (M = 3.69),  $\chi^2(3) = 110.82$ , p < .001. The Wilcoxon Signed-Rank Test revealed a significant difference between all comparisons. For Dutch children, the Friedman Test revealed that children's first choice was 'playing with iCat' (M = 1.53), followed by 'playing with friend' (M = 1.69), 'playing alone' (M = 3.27) and 'playing with stranger' (M = 3.61),  $\chi^2(3) = 111.00$ , p < .001. The Wilcoxon Signed-Rank Test revealed a significant difference among all comparisons except for one (playing with iCat vs. playing with friend).

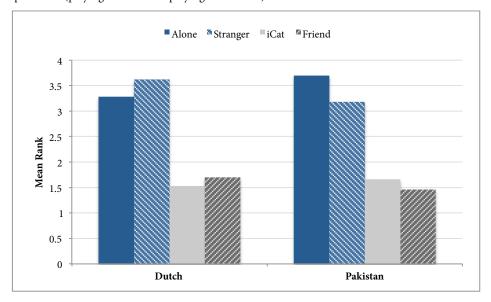


Figure 6: Rank preferences for four game playing conditions on the scale of 1 (high) to 4 (low)

# 3. Experiment 2: Perception Test

In the previous section we discussed children's self-reports on how they appreciated the interaction with the iCat, and compared the results with those obtained from children playing alone or with a friend. During the game playing session, children were also videotaped. These recordings comprise a rich collection of social behaviour and of verbal and non-verbal emotional responses to winning or losing a game. In this section we present judgements of selected recordings from children (taken from all three conditions). The task of judges is to guess whether the child has just won or lost the game. In this way we can objectively estimate the expressiveness of the children in the different conditions; more correct guesses indicate that children are more expressive.

# 3.1 Stimuli

We randomly selected 80 children who played games with the iCat, 40 Dutch and 40 Pakistani. For comparisons, we also used clips of 144 children from the 'alone' and 'pairs' condition, which were collected earlier, mentioned in the introduction of this chapter. We balanced for age and gender and for each child we chose the first winning game (in which children made a correct prediction for the last card) and the first losing game (in which the final guess turned out to be incorrect) from the video recordings. Furthermore, from the clips of child-pairs, we randomly selected one child from each pair by zooming in on his/her face. In the case of the iCat condition, we also zoomed in on the child's face so that the iCat was never visible in the stimuli. Half of the children sitting on the right chair and half of the children sitting on the left chair were selected. The video snippets were cut from the moment the final card was turned until the main response of the child was finished.

This resulted in 4 sets of stimuli:  $stimuli\ 1$  – 80 Dutch iCat clips [= (10 (iCat - 8 year old boys) + 10 (iCat - 8 year old girls) + 10 (iCat - 12 year old boys) + 10 (iCat - 12 year old girls)) x 2 (1 win + 1 lost)],  $stimuli\ 2$  – 80 Pakistani iCat clips [similar to stimuli 1],  $stimuli\ 3$  – 96 Dutch alone and pairs clips [= (12 (Alone - 8 years old alone) + 12 (Alone - 12 years old) + 12 (Pairs - 8 years old) + 12 (Pairs - 12 years old)) x 2 (1 win + 1 lost)],  $stimuli\ 4$  – 96 Pakistani alone and pairs clips (similar to stimuli 3).



Figure 7: Representative stills of children for winning and loosing games during three conditions

These four sets of stimuli were presented in a perception test to observers for a rating experiment. In this experiment, stimuli were presented in a random order, in a vision-only format so that participants could not rely on auditory cues (e.g. "Yahooo!" sound used as feedback to children that they had won the game).

# 3.2 Participants

144 Dutch adults participated in the perception experiments (M = 21 years, SD = 2.7 years), with a roughly equal number of men and women.

# 3.3 Procedure

Four group experiments were conducted and in each experiment 36 Dutch viewers viewed one set of stimuli. For every experiment, groups of participants were invited into a quite classroom where the computer screen was projected on the classroom wall using a projector. Participants were told that they would see a set of stimuli in which children were showing their emotions after winning or losing a game and that their task was to determine whether the children had just won or lost their game (Figure 7 shows representative stills from each clip). Each stimulus was preceded by a number displayed on the screen indicating the upcoming stimulus, and followed by a six second interval during which participants could fill in their score on the answer form. The actual experiment was preceded by a short training session in which 3 clips were shown (different from the ones shown in the actual set of stimuli) to make participants familiar with the stimuli and the experimental task. If everything was clear, the actual experiment started. During the experiment there was no interaction between participants and experimenter.

# 3.4 Data analysis

In the following, we first statistically compare the results across the three main conditions of interest (Alone, Friend, iCat) using independent samples t-tests (an analysis of variance could not be used since we compare data collected in various experiments, as described above). We compared through two t-tests how the results for the iCat condition relate to that for the alone and together conditions, respectively. For a detailed analysis, we then zoomed in on the iCat data only, where tests of significance were performed using a repeated measure analysis of variance (ANOVA) with three within-subjects factors, namely Culture (levels: Pakistani, Dutch), Age (levels: 8 years, 12 years), Game status (levels: winning, losing) and with the percentage of correct classification as the dependent variable. Our working hypothesis is that when children are more expressive, participants will be better able to determine whether a child had won or lost. A higher percentage of correct classifications is therefore related to more nonverbal expressiveness.

# 3.5 Results

Figure 8 summarizes the results, and presents a very clear picture. Children playing a game with the iCat are more expressive than children playing alone (more correct classifications), but less expressive than children playing with a friend. This pattern is identical across the two cultures, but Pakistani children appear to be more expressive than Dutch ones.

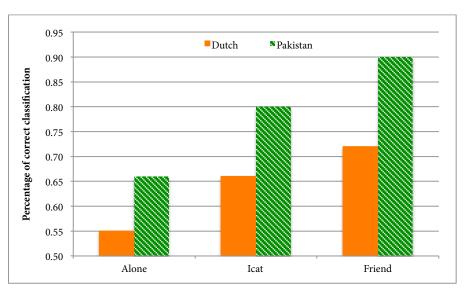


Figure 8: Percentage of correct classification of children winning and losing for three game playing conditions and two cultures (higher percentage correct = more expressive, lower percentage correct = less expressive, 50% = chance level)

We first statistically compare the data across the three conditions for determining the main effect. An independent samples t-test confirmed that the percentage of correct classification for the iCat condition (M = .73) is higher than for the alone condition (M = .61), t(142) = 8.73, p<.001. In turn, the percentage of correct classification of children playing with a friend (M = .81) is higher than that for children playing with the iCat (M = .73), t(142) = 5.09, p<.001.

Next, we turn to the iCat data. First of all, we found that the percentage of correct classifications is higher for Pakistani children (M = .80) than for Dutch children (M = .66), F(1,29) = 230.68, p < .66.001,  $\eta^2_p$  = .89. In addition, overall the percentage of correct classifications is higher for 8 year olds (M = .75) than it is for 12 year olds (M = .71), F(1,29) = 22.31, P < .01,  $\eta^2_P = .45$ , indicating that younger children are more expressive when playing with the iCat than older children. Furthermore, the percentage of correct classifications is higher for losing games (M = .82) than it is for winning games (M = .64), F(1,127) = 148.33, p < .001,  $\eta^2_p = .84$ .

Given the differences in how Dutch and Pakistani children of different age groups respond to winning and losing while playing with iCat, the interaction effects are especially interesting. We found a significant interaction between Culture and Age (F (1,29) = 4.25, p < .05,  $\eta_p^2 = .13$ ). This interaction can be explained by the fact that while playing game with iCat, the Dutch children aged 8 (M = .67) and aged 12 (M = .64) are comparable in their expressiveness, while the Pakistani 8-year olds (M = .83) are overall more expressive (higher percentage correct) than the Pakistani 12 year olds (M = .77); see Figure 9.

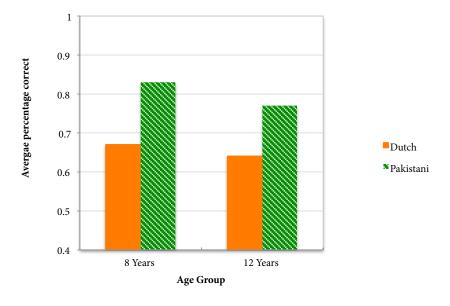


Figure 9: Average percentage correct as a function of Culture and Age

We also found a significant interaction between culture and game status (F(1,29) = 119.28, p <.001,  $\eta_p^2$  = .80). This interaction can be explained as follows: for Dutch children, losing games are more often classified correctly than winning ones, while for Pakistani children the difference between these two is much smaller, as can be seen in figure 10.

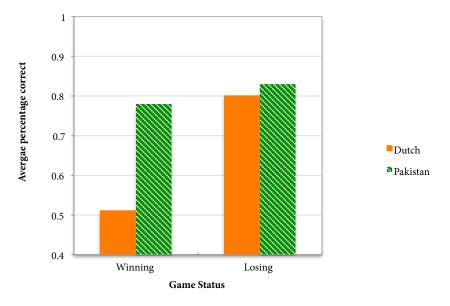


Figure 10: Average percentage correct as a function of Culture and Game Status

# 4. Observational Analysis

#### 4.1 Introduction

The perception tests clearly showed that children are more expressive when playing a game with an iCat than when playing the game alone, although they are more expressive till when playing with a friend. In this section, we describe a detailed observational analysis of the children's non-verbal behavior in the various conditions. This was partly done to find out what kinds of features participants of our perception tests may have used as cues in their judgments, but also to gain insight into specific behaviors that children display in various game contexts.

# 4.2 Development of observation scheme

We developed a new observation scheme for observing and quantifying children's responses at two different moments during the game, i.e., (1) children's affective response after winning or losing a game and (2) children's social interaction during the game play. The first part of the scheme focused on the game outcome, to quantify the positive (joy after winning) and negative (disappointment after losing) emotions. This part of the coding scheme was developed on the basis of informal observations of video clips taken from three game playing conditions (playing alone, playing with iCat, and playing with friend) and on an existing coding scheme (Shahid et al., 2011). This resulted in two main categories (see appendix 1 for details): 1) positive affect (subcategories: smile, laughter, clapping, jumping, moving, winning arm gesture, verbal response (e.g. "yes", "yahoo", etc.)), and 2) negative affect (subcategories: sadness (frown), anger, closing eyes, covering face, head down, verbal response (e.g. "ohoo", "no", etc.).

For this part of the coding scheme, some chosen facial features are roughly comparable to the well-known Action Units (AU's) of the Facial Action Coding Systems (FACS) proposed by (Ekman and Friesen, 1978) (AU 12 for smiling, AU 51-58 for turn taking or head movement, AU 4 for frowning, etc.). All visual features and relevant stills of social behaviors can be found in the appendix 1. In addition to these visual features, we also added one auditory feature, i.e. the verbal response to a winning or losing situation, even though these could not have served as cues in the perception study, given that sound was removed from the clips.

The second part of the coding scheme focused on observing social interactions during gameplay. We used video recording from two conditions (playing with iCat, playing with friend) as a point of reference for identifying different social behaviors. We based this part of the scheme on the social interaction framework in gameplay (Kort et al., 2007) (e.g. measuring (non) verbal communication, immediacy, monitoring, etc.) and guidelines for measuring social play (Isbister, 2010). All together this resulted in six categories: 1) physical connectedness (subcategories: hugging, touching, hitting), 2) responsiveness (subcategories: stool moved (for sitting closer to the partner), turn to partner, turn taking (guessing numbers one by one), gaze partner (looking into the partner's eyes) and 3) compliance (subcategories: accepting certain proposal, accepting uncertain proposal.

We coded the presence and absence of these features and counted the number of times a particular behavior (e.g. smiling, touching, hugging, etc.) occurred in the selected fragments. In the case of eye-gaze behavior, we coded the duration of the behavior for a fixed interval. Even when we coded for the social interaction in the 'playing with friend' condition, we coded the behavior of only one player, thus ignoring the behavior of his/her team partner.

#### 4.3 Stimuli

The stimuli for coding affective response were the same as we used previously in the perception test. For coding social interaction, the stimuli were extended, so as to include the decision process leading up to the final choice. Sound was not muted, as it contained conversation about the last guess and verbal responses to winning and losing. Furthermore, coding was done on the original and not on the zoomed clips, to ensure that the participant was clearly visible during the entire interaction.

#### 4.4 Procedure

The first author and two independent assistants, one Dutch and one Pakistani, performed the labeling of the observational data. The procedure was as follows: The first author trained the two other coders until an acceptable level of inter-rater agreement was established before starting the coding. After the coding, the inter-rater reliability for each feature was measured. All three coders coded all fragments and scored the number of times different behaviors mentioned in the coding list happened in the given stimulus. The coding was done in two rounds where each coder watched all fragments twice, while being blind for condition. In the first round, all coders coded the 'affective' dimension (positive, negative or neutral response to winning and losing) of the gameplay. In the second round, each coder primarily coded the social dimension of gameplay. Coders also coded the verbal response to winning and losing during this round.

# 4.5 Inter-rater reliability

At the end of the coding sessions, the coding of the principle researcher and the other two coders were checked for the percentage agreement for all categories separately. The inter-rater agreement ranged from 63% to 88% for positive affect, 69% to 92% for negative affect, 74% for neutral, 81% for tense, 87% to 94% for physical connectedness, and 79% to 91% for responsiveness. For further analyses, disagreement between three labelers was resolved through majority voting.

# 4.6 Data Analysis

For finding the similarities and differences in the exhibition of positive and negative cues during the three game playing conditions, two separate univariate ANOVAs were performed for positive and negative affect categories with culture (levels: Pakistani, Dutch) and presence (levels: alone, iCat, friend) as between subject factors, and mean number of visual features as the dependent variable.

# 4.7 Results

Table 3 summarizes the results and shows the number and type of visual features displayed by Pakistani and Dutch children in three game-playing conditions. In general, it can be seen that more features are shown in the together conditions than in the alone condition, with the iCat condition sandwiched in between. Furthermore, the number of features shown by Pakistani children is overall higher than the number by the Dutch children.

Affect Category	Features	Dutch			Pakistani		
		Alone	iCat	Friend	Alone	iCat	Friend
Positive	Smile (happy)	16	18	21	16	21	25
	Laugh	0	2	5	1	7	14
	Clapping	0	3	5	2	9	16
	Jumping/Moving	0	4	10	3	8	22
	Winning gesture	1	7	11	2	13	21
	Verbal response	2	4	13	3	8	10
То	tal positive Features	19	38	65	27	66	108
Negative	Sad (frown)	17	19	22	16	18	21
	Anger	0	1	4	0	2	5
	Closing eyes	8	10	14	7	15	21
	Covering face	1	6	10	4	8	14
	Head down	2	4	6	2	6	9
	Verbal response	3	10	16	5	7	8
Tot	al Negative Features	31	50	72	34	56	78
T	otal Visual Features	50	88	137	61	122	186

Table 3: Number of positive and negative features shown by children during three game playing conditions

The two-way ANOVA for the positive affect category (Figure 11) revealed that children were more expressive and showed the most number of positive cues while playing with friends, the least number of cues while playing alone and the number of positive cues shown while playing with the iCat was in between, F(2,287) = 42.708, p < .001. The two-way ANOVA also showed a main effect of culture, in the sense that Pakistani children showed a higher number of visual cues than their Dutch counterparts, F(1,287) = 21.670, p < .001.

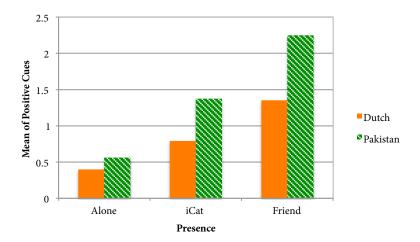


Figure 11: Mean of positive features shown by Pakistani and Dutch children

The two-way ANOVA for the negative affect category (Figure 12) revealed that children showed the most number of negative cues while playing with friends, the lowest number of cues while playing alone and the number of negative cues shown while playing with the iCat was in between, F (2,287) = 18.823, p < .001. For the negative cues category, no significant effect of culture was found which suggest that both Dutch and Pakistani children were equally expressive after losing a game.

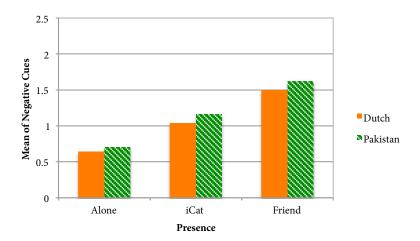


Figure 12: Mean of negative features shown by Pakistani and Dutch children

We also explored to what extent there is a relation between the percentage of correct classification of games (perceived status of a game by judges) and the number of visual features shown in three game playing conditions. For this analysis, we linked the observational data with the perception test results. In this way, for each clip used in the perception test and observational analysis, we had a one to one mapping between percentage of correct classification and number visual cues.

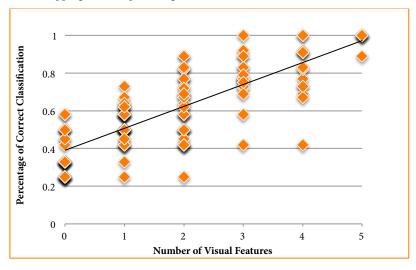


Figure 13: Correlation between percentage of correct classification and number of visual features shown by Dutch children

We performed separate correlational analyses for the two cultures. These correlation analyses revealed that the percentage of correct classification increases as a function of the number of visual features shown by children. Figure 13 and 14 shows that, both for Dutch (r = .672, p < .001) and Pakistani (r = .799, p < .001) children, the increase in the amount of visual features present in a stimulus gives rise to better classification. This correlation is significantly higher for Pakistani children than for Dutch children.

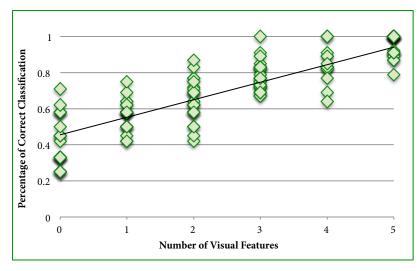


Figure 14: Correlation between percentage of correct classification and number of visual features shown by Pakistani children

To find out which social cues children exhibited while interacting with the iCat and with friends, a two-way ANOVA was run for 'physical connectedness', 'responsiveness, and 'compliance' categories with culture (levels: Pakistani, Dutch) and partner (levels: iCat, together) as between subject factors and mean of visual features as the dependent variable.

Category	Features	Dutch		Pakistan	
	_	iCat	Together	iCat	Together
Physical connected-	Hugging	1	0	3	13
ness	Touching	4	8	19	26
	Hitting	0	2	0	7
	Total features	5	10	22	46
Responsiveness	Stool moved	6	15	25	30
	Turn to partner	23	30	45	40
	Gaze Partner	13	19	38	31
	Turn taking	43	39	35	30
	Total features	85	103	143	131
Compliance	Accepting certain proposal	42	45	44	37
	Accepting uncertain proposal	12	6	21	10
	Total features	54	51	65	47

Table 4: Number of social cues shown by children during three game playing conditions

Table 4 summarizes the results. For physical connectedness, the two-way ANOVA showed that Pakistani children were more socially connected with partners than Dutch children, F (1,191)= 46.585, p < .001,  $\eta^2_p = .20$ . We also found that children were more connected while playing with their friends than with the iCat, F (1,191)= 14.378, p < .001. We also found a significant interaction between culture and partner, F (1,191)= 2.038, p < .05,  $\eta^2_p = .03$ , which indicates that although in both cultures children were more connected while playing with their friends as compared to playing with iCat, this difference in the two game playing conditions is higher for Pakistani children.

For responsiveness, the two-way ANOVA showed that Pakistani children were more responsive to their partners than Dutch children, F(1,191)=40.333, p<.001,  $\eta^2_p=.29$ . We also found a significant interaction between culture and partner, F(1,191)=4.083, p<.01,  $\eta^2_p=.04$ , which

indicates that Dutch children were more responsive to their friends but Pakistani children were more responsive to iCat. No main effect of partner was found.

For compliance, the two-way ANOVA showed that children more often agreed with the iCat than with their friends, F(1,191) = 7.219, p < 0.01,  $\eta^2 p = 0.04$ . We also found a significant interaction between culture and presence, F(1,191) = 4.061, p < .05,  $\eta^2_p = .02$ , which can be explained by the fact that although in both cultures children agreed more with iCat than with their friends, this difference in agreement between the two game playing conditions was much larger for the Pakistani children.

#### 5. Discussion

In this chapter we studied how children of different age groups (8 vs. 12 year old) and belonging to different cultural backgrounds (Dutch vs. Pakistani) experience playing a game with a robot in comparison to playing the same game alone or with a friend. To the best of our knowledge, our study is the first to investigate cross-cultural child-robot interaction in playful settings using such an experimental paradigm. In this study, unlike many of the previous cross-cultural HRI studies, we systematically used a combination of self-reports, objective measures and behavioral analyses and in the next section we summarize our major findings and reflect on them in a detailed manner.

# 5.1 Child-robot interaction: playing alone or together

In general the results of the subjective measures, the perception test and behavioral analyses are very consistent. Subjective fun scores suggest that children have more fun playing with the robot than playing alone, but have more fun still when playing with a friend. Overall the experience of playing with the iCat is much closer to playing with a friend than playing alone: the mean difference of the fun scores for the iCat and Alone conditions is about twice that of the iCat and Friend condition. A perception test of selected fragments shows a similar trend and indicates that children are more expressive when playing with the robot than they are when playing alone, but less expressive than when playing with a friend. In a similar vein, finally, the behavioral analyses indicate that children showed the highest number of visual and social cues while playing with a friend, the lowest number of cues while playing alone, while the number of cues shown while playing with a robot was in between these two extremes.

So, to what extent did children respond to and showed attitudes towards the iCat as if it were a human partner? The answer to this question is not straightforward. On the one hand, disregarding the cultural and age differences, most of the children indeed strongly felt the presence of the iCat as a game partner and also showed a tendency of treating the iCat as a human partner. Many of the elements and social behaviors of "normal" child-child interaction were present during child-robot interactions such as responding to the iCat after every guess, or gazing at the iCat, showing sympathy and affection (as shown in figure 15). Furthermore, most of the children were positive about their interaction with the iCat and almost all types of visual (negative and positive) cues present when interacting with a friend were also present during their interaction with the iCat.





Figure 15: Left: A 8 year old Dutch girl padding (showing empathy) iCat after losing a game, Right: a 12 year old Pakistani girl saying goodbye by kissing the iCat.

On the other hand, however, there are key differences in the way children responded to the iCat and to a human partner. Compared to playing with the iCat, children were not only more expressive while playing with their friends, but they also showed more nonverbal cues. Furthermore, the coding of the social cues revealed differences between playing with friends and playing with iCat. For example, children jumped/moved more while interacting with their friends and they touched their friends more often than they did with the iCat. Children also sometimes hit their friends after losing a game, but they (fortunately) never hit the iCat. Children, by moving their stools, sat in closer proximity to their friends than was the case with the iCat.

Although, in general, playing with a friend was appreciated more than playing with a robot, there were important individual differences, in particular because of cultural background and age of the child participants.

# 5.2 Child-robot interaction: Effect of culture and age

When looking at the effects of culture and age, the results of the subjective measures, the perception test, and the behavioral analysis are again consistent. Subjective fun scores suggest that in all three game playing conditions, Pakistani children reported more fun than Dutch children. In both cultures, 8 year olds generally reported more fun than 12 years old. The results of the GEQ dimensions suggest that Pakistani children were more engaged and showed a stronger interest in playing with the iCat again, compared to Dutch children. Furthermore, 8 year old children found playing with the iCat more engaging than 12 year olds and 8 years olds were also unhappier when the game was over. Both Pakistan and Dutch children equally appreciated and liked the iCat but overall 8 year olds liked the iCat more than 12 year olds. These findings are also in line with the findings of the post-game interviews. After the experiment, a number of Dutch children showed an interest in buying the iCat. One child said, "I would like to buy this iCat and I am willing to pay from my pocket money". Furthermore 8 year old children liked the iCat as a game partner more than 12 year old children and thought it was as helpful as their friends would be. A number of Pakistani children, mostly young, asked the researcher about the shop where they could buy this robot. This result is consistent with the survey results of (Dautenhahn et al., 2005), where they

showed that as compared to older participants, young participants showed strong interest in having a robot companion as a friend at home.

Both Pakistani and Dutch Children appreciated the very basic movements and behaviors of the iCat such as moving the head, eyes, and lips and the display of feedback with lights. From retrospective interviews with these children, we learned that the feeling that children had a rich social interaction was caused by the emotional response of the iCat with appropriate facial expressions, the fact that the conversation proceeded in a natural language, and the fact that the iCat quickly reacted to different actions.

In terms of the perceived smartness of the iCat, 12 year old children (compared to the 8 year old children) thought that the iCat was more intelligent and was able to give them a decent number of right answers. Furthermore, Dutch children rated the smartness of the iCat higher than Pakistani children. One possible explanation for this effect is the 'level of expectations' associated with the iCat by children. In the post-game interviews, we learned that younger children associated higher expectations with the iCat than older children and when these expectations were not fully met, i.e., sometimes iCat's guess was wrong, children got disappointed.

In the perception experiment, we found that the percentage of correct classification was higher for Pakistani children than Dutch children while playing with the iCat. Moreover, both Pakistani and Dutch 8 year old children were more expressive, as shown by more correct classifications, than 12 year old children interacting with iCat. This result are in line with the theory that younger children are more expressive than older ones (Thompson, 1994). The perception experiment also revealed that both Pakistani and Dutch children were equally expressive while showing negative emotions (after losing), while Pakistani children were more expressive than Dutch children while showing positive emotions (after winning). The perception test scores are also very much in line with the fun scores and GEQ self-reports in the sense that the children's appreciation of the robotic partner during collaborative game play is reflected in the degree to which they become expressive. In sum: culture and age not only had an effect on children's subjective feelings towards collaborative gameplay with the iCat, but these factors also shaped their expressive behavior.

The results of observational analysis also support this. Observational coding for the affective dimension revealed that Pakistani children showed more positive cues than their Dutch counterparts, while the number of negative cues shown by children during all three game playing conditions was comparable across both cultures. This is in line with the perception test results, which showed that Pakistani and Dutch children were equally expressive after losing while playing with iCat. The correlational analysis for Pakistani and Dutch cultures showed that for both cultures there was a strong positive correlation between the percentage of correct classifications and the number of visual features shown by children after winning or losing a game. The results of the correlational analysis not only strengthen the results of the perception test but also endorse it as a practical method to get an objective measure of the affective states of participants.

The results of the observational analysis of the social dimension showed remarkable differences between the two cultures while interacting with a robotic partner and a human partner. Pakistani children hugged and particularly touched iCat more often than their Dutch counterparts who almost never touched the iCat and also rarely touched their friends. Pakistani children were also found to be more responsive to the iCat than their Dutch counterparts. Dutch children on the other hand were more responsive to their friends than to iCat. Dutch children almost never moved their stool while interacting with iCat and rarely gazed at iCat, especially not during the guessing process.



Figure 16: Top row: 12 year old Dutch girl (left) and Pakistani girl (right) after losing a game, bottom row: 8 year old Dutch boy(left) and Pakistani boy (right) communication with the iCat after losing a game

Both Pakistani and Dutch children agreed with iCat when it expressed itself in a certain way, yet Pakistani children (more often than Dutch ones) also agreed relatively frequently when the iCat showed uncertainty about the next guess. Although Dutch children accepted more uncertain proposals from the iCat than from their friends, they were overall very strict about their first guess and they hardly ever changed their opinion on the basis of the iCat's recommendations, especially not when the iCat was uncertain.



Figure 17: Pakistani 12 years old girl praying before the final card, a square highlights the praying gesture

In terms of specific cross-cultural differences, Pakistani children also showed some culture specific gestures which were never shown by Dutch children such as dancing after winning a game, praying gesture while waiting for the final answer (Figure 17), 'grabbing both ears' after losing (as an apology to partner). Dutch children never prayed or showed different culture-specific gestures.

Another cross-cultural difference was the way Dutch and Pakistani children communicated with iCat. Compared to Dutch children who almost always used 'informal you' ("jij" rather than "u") for addressing the iCat, Pakistani children, depending on age, used both formal and informal variants of 'you'. A majority of 8 year old children used informal version of 'you' ("Tum" in Urdu) and this variant of 'you' is usually used between friends or persons of same age groups. On the other hand, many 12 year old children used the formal variant of 'you' ("Aap" in Urdu). Note that the iCat always used informal you, for both cultures. Dutch children did not raise any concerns about the iCat addressing them in an informal manner but a few Pakistani 12 year old children did notice this behavior. One 12 year old girl said, "I am always using "Aap" for her but she kept on saying "Tum" to me. I think she should start talking in a good manner."

# Considerations for CRI methodology and theory

Our study provided clear results on how children of different age groups and cultural backgrounds treat an entertaining social robot in a natural social setting. Our method of using two control conditions (playing 'alone' and playing 'together' with a friend) for evaluating child-robot interaction seems promising and reveals an interesting dimension of this interaction where the overall experience of playing a game with a modern social robot falls in between these two extremes. Although playing with a robot is more fun than playing alone, and approaches playing with a friend, for achieving the latter level more attention should be paid on designing a better interaction between children and robots.

Our evaluation method can be useful, for example, for interactions with autistic children or children with special needs, given that robotic technology is also designed for improving the socioemotional skills of children. Our method provides a new way to 'benchmark' CRI by comparing it to other conditions. We feel this is more informative than the evaluation in many existing HRI studies where participants merely rate their experience in an absolute sense. If one wants to improve CRI and make it more human- and life-like, it would seem important to compare it with human-like benchmarks.

In terms of theory, in line with the observations of (Melson et al., 2005), we also believe that the media equation theory, stating that people treat computers as human actors, holds for CRI on a more general level. As mentioned earlier, despite the iCat's limitations, many children (especially young ones in both cultures of interest) were able to associate social attributes to the iCat, imagined that the iCat is socially competent and interacted with iCat in socially appropriate manner. We realized that this effect is probably as strong as it is because the iCat was able to show basic social competence and responded to the game situations in an appropriate manner (showing sadness after loosing, showing surprise on an unexpected question or reaction of the child, appreciating the child's input by showing a smile in the case of correct guess). We believe that the well-controlled wizard of Oz set-up helped to achieve a very realistic CRI scenario. This is also what children seemed to appreciate most during the post-experiment interviews.

Nevertheless our results also clearly showed that cultural background and age differences played a big role in the way children interacted with iCat. Unlike traditional computers, robots (particularly social ones) are not only more responsive but also display more life-like and social properties, which make children of different age groups and cultures sensitive to such robotic devices, as we found, for instance, concerning the use of formal and informal 'you'. In general, our results show that age and culture are important factors to take into account when developing CRI applications. Cross-cultural robot design should go beyond mere translations and high-level localizations as also insisted by (Wang et al., 2010). The underlying social norms of a particular culture should be part of the robot design process.

# 6. Conclusion

In this chapter we took a new approach to study child-robot interaction, by asking whether playing a game with a modern social robot is more similar to playing this game alone or with a friend. We focused on children of different age groups (younger vs. older children) and belonging to different cultural backgrounds (Dutch vs. Pakistani). This study consisted of two main experiments. In the first experiment, we collected data using a deterministic card guessing game, which children played in one of three conditions (alone, with a friend or with the iCat). After the game playing session, we collected children's subjective fun reports and overall game experience. In the second experiment, we used a perception test as an objective evaluation method, for measuring the emotional response of children after winning or losing a game. In the last part of this study, we did an observational analysis and coded the video data along affective and social interaction dimensions.

Our study highlighted some important, but previously understudied aspects of child-robot interaction in an authentic social setting. The results showed that the overall experience of playing a game together with a modern social robot falls in between the playing 'alone' and playing 'together with a friend' conditions. Although playing with a robot is more fun than playing alone, and approaches playing with a friend, more work is needed before playing with a social robot is truly like playing with a friend. Our results also show clear differences in the way children of different age groups and cultural backgrounds interacted, not only with a human partner but also with a robot. In general, we found that Pakistani children had richer social interactions than Dutch children, and that younger children were more positive about the robot than older children (12 years old). Based on these results, we propose that one way of designing a better interaction between children and robots is to take individual differences such as cultural background and age explicitly into account during the design process.

# Appendix 1

Smile



Laugh



Clapping



Wining gesture



Moving



Jumping



Sad (frown)



Head down



Closing eyes



Covering face



# (Social Interaction cues)

Hugging



Touching



Hitting



Chair moved (Example of small distance)



Turn to partner /Eye-gaze (example of 'more distance' as well)



#### References

Aubrey, S., 2009. A Cross-Cultural Discussion of Japan and South Korea and How Differences Are Manifested in the ESL/EFL Classroom. Asian Social Science, 5(5), 34-39.

Bartneck, C., Nomura, T., Kanda, T., Suzuki, T., Kennsuke, K., 2005. A cross-cultural study on attitudes towards robots. HCI International, 1981-1983.

Bartneck, C., Suzuki, T., Kanda, T., A Nomura, T., 2007. The Influence of People's Culture and Prior Experiences with AIBO on their Attitude Towards Robots. AI Society, 21, 217-230.

Baxter, P., Belpaeme, T., Cañamero, L., Enescu, V., Brussel, V.U., 2011. Long-Term Human-Robot Interaction with Young Users. In Proceedings of the IEEE/ACM HRI workshop on Robots interacting with children, Lausanne, Switzerland.

Breazeal, C., 2002. Designing Sociable Robots. MIT Press, Cambridge.

Bursie, C.I., 2004. Dog-like Behavior Selection for an Aibo Robot Dog, Technology, 94.

Castellano, G., Leite, I., Pereira, A., Martinho, C., Paiva, A., McOwan, P., 2010. Affect Recognition for Interactive Companions: Challenges and Design in Real World Scenarios. Journal on Multimodal User Interfaces, 3(1-2), 89-98.

Cho, H.K., Park, K.P., Han, J.H., Min, D.K., Ko, K.H., 2008. Education+Robots: the Vision and the Action Plans. Communications of the Korea Information Science Society, 26(4), 55-64.

Conn, K., Sarkar, N., Stone, W., 2008. Online Affect Detection and Robot Behavior Adaptation for Intervention of Children With Autism. IEEE Transactions on Robotics, 24, 883-896.

Dautenhahn, K., Woods, S., Kaouri, C., Walters, M.L., Koay, K.L., Werry, I., 2005. What is a robot companion-Friend, assistant or butler? International Conference on Intelligent Robots and Systems, IEEE, 1192-1197.

Ekman, P., Friesen, W., 1978. Facial Action Coding System: A Technique for the Measurement of Facial Movement. Consulting Psychologists Press, Palo Alto.

Elfenbein, H., Ambady, N., Mandal, M., Harizuka, S., 2002. Cross-cultural Patterns in Emotion Recognition: Highlighting Design and Analytical Techniques. Emotion(2), 75-84.

Evers, V., Maldonado, H.C., Hinds, P., 2008. Relational vs . Group Self-Construal : Untangling the Role of National Culture in HRI. Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction. ACM, 255-262.

Fong, T., Nourbakhsh, I., Dautenhahn, K., 2003. A survey of socially interactive robots. Robotics and Autonomous Systems, 42, 143-166.

Francois, D., 2009. Facilitating Play Between Children with Autism and an Autonomous Robot, Adaptive Systems Research Group, School of Computer Science. University of Hertfordshire.

Fussell, S.R., Kiesler, S., Setlock, L.D., Yew, V., 2008. How people anthropomorphize robots. Proceedings of the 3rd ACMIEEE international conference on Human robot interaction, 145-152.

Han, J., Hyun, E., Kim, M., Cho, H., Kanda, T., Nomura, T., 2009. The Cross-cultural Acceptance of Tutoring Robots with Augmented Reality Service. International Journal of Digital Content Technology and its Applications, 3(2).

Han, J.-H., Jo, M.-H., Jones, V., Jo, J.-H., 2008. Comparative Study on the Educational Use of Home Robots for Children. Journal of Information Processing Systems, 4, 159-168.

IJsselsteijn, W.A., Kort, Y.A.W.d., Poels, K., 2007. The Game Experience Questionnaire: Development of a self-report measure to assess the psychological impact of digital games. TU Eindhoven.

Isbister, K., 2010. Enabling social play: A framework for design and evaluation, in: Bernhaupt, R. (Ed.), Evaluating User Experiences in Games: Concepts and Methods. Springer.

Kahn, P.H., Freier, N.G., Kanda, T., Ishiguro, H., Ruckert, J.H., Severson, R.L., Kane, S.K., 2008. Design patterns for sociality in human-robot interaction. Proceedings of the 3rd international conference on Human robot interaction HRI 08, ACM Press.

Kahn, P.H., Friedman, B., Perez-Granados, D.R., Freier, N.G., 2006. Robotic pets in the lives of preschool children. Interaction Studies, 7(3), 405-436.

Kaiser, S., Wehrle, T., 1996. Situated emotional problem solving in interactive computer games. In N. Frijda (Ed.), Proceedings of the International Society for Research on Emotions, 276-280.

Kanda, T., Hirano, T., Eaton, D., Ishiguro, H., 2004. Interactive Robots as Social Partners and Peer Tutors for Children: A Field Trial. Human-Computer Interaction, 19, 61-84.

Kerepesi, A., Kubinyi, E., Jonsson, G.K., Magnusson, M.S., Miklósi, A., 2006. Behavioural comparison of human-animal (dog) and human-robot (AIBO) interactions. Behavioural Processes, 73, 92-99.

Kort, Y.A.W.d., Ijsselsteijn, W.A., Gajadhar, J.B., 2007. People, Places and Play: A Research Framework for Digital game Experience in a Socio-spatial Context. DiGRA Conference.

Kozima, H., Nakagawa, C., Yasuda, Y., 2007. Children-robot interaction: a pilot study in autism therapy. Progress in Brain Research, 164, 385-400.

Lee, K.M., Park, N.K., Song, H., 2004. Can a Robot be Perceived as a Developing Creature?: Effects of Artificial Developments on Social Presence and Social Responses toward Robots in Human-robot Interaction. International Communication Association Conference, New Orleans, LA.

Looije, R., Neerincx, M.A., Lange, V.D., 2008. Children's responses and opinion on three bots that motivate, educate and play. Journal of Physical Agents, 2(2), 13-20.

Markopoulos, P., Read, J., MacFarlane, S., Hoysniemi, J., 2008. Evaluating Children's Interactive Products: Principles and Practices for Interaction Designers (Interactive Technologies). Morgan Kaufmann.

Melson, G.F., Kahn, P.H., Beck, A.M., Friedman, B., Roberts, T., Garrett, E., 2005. Robots as Dogs? -Children's Interactions with the Robotic Dog AIBO and a Live Australien Shepherd. Proceedings of CHI 2005, ACM Press, Portland, Oregon, 1649-1652.

Mitchell, R.W., Hamm, M., 1997. The interpretation of animal psychology: Anthropomorphism or behavior reading? Behaviour, 134, 173-204.

Nomura, T., Suzuki, T., Kanda, T., Han, J., Shin, N., Burke, J., Kato, K., 2007. What People Assume about Robots: Cross-Cultural Analysis between Japan, Korea, and the USA, Human Robot Interaction. I-Tech Education and Publishing, pp. 275-288.

Oh, K., Kim, M., 2010. Social Attributes of Robotic Products: Observations of Child-Robot Interactions in a School Environment. International Journal of Design, 4, 45-55.

Oshlyansk, L., 2007. Cultural Models in HCI: Hofstede, Affordance and Technology Acceptance. Swansea University.

Pepe, A.A., Ellis, L.U., Sims, V.K., Chin, M.G., 2008. Go, Dog, Go: Maze Training AIBO vs. a Live Dog, An Exploratory Study. Anthrozoös, 21, 71-83.

Read, J.C., 2008. Validating the Fun Toolkit: an instrument for measuring children's opinions of technology. In the Journal of Cognition, Technology & Work, 10(2), 119-128.

Reeves, B., Nass, C., 1996. The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places. Cambridge University Press, New York.

Ribi, F.N., Yokoyama, A., Turner, D.C., 2008. Comparison of Children's Behavior toward Sony's Robotic Dog AIBO and a Real Dog: A Pilot Study. Anthrozoös: A Multidisciplinary Journal of The Interactions of People Animals, 21, 245-256.

Robins, B., Dickerson, P., Stribling, P., Dautenhahn, K., 2004. Robot-mediated joint attention in children with autism: A case study in robot-human interaction. Interaction Studies, 5, 161-198.

Salter, T., Michaud, F., Larouche, H., 2010. How wild is wild? A taxonomy to characterize the 'wildness' of child-robot interaction. International Journal of Social Robotics.

Salter, T., Werry, I., Michaud, F., 2008. Going into the wild in child-robot interaction studies: issues in social robotic development. Intelligent Service Robotics, 1, 93-108.

Shahid, S., Krahmer, E., Swerts, M., 2008a. Alone or Together: Exploring the Effect of Physical Copresence on the Emotional Expressions of Game Playing Children Across Cultures. In the International Conference of Fun and Games, Eindhoven, the Netherlands, Springer.

Shahid, S., Krahmer, E., Swerts, M., 2008b. GamE: Game as a method for Eliciting Emotions. Proceedings of the Measuring Behaviour Conference (MB 2008), Maastricht, the Netherlands.

Shahid, S., Krahmer, E., Swerts, M., 2011. The effect of co-presence and culture on the facial expressions of game playing children: An experimental exploration. Submitted.

Shibata, T., Wada, K., Tanie, K., 2003. Subjective evaluation of a seal robot at the national museum of science and technology in Stockholm 12th IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN 2003), San Francisco, USA, 397-407.

Shibata, T., Wada, K., Tanie, K., 2004. Subjective evaluation of a seal robot in Burunei. 13th IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN 2004), Kurashiki, Japan, 135-140.

Shinozawa, K., Reeves, B., Wise, K., Lim, S., Maldonado, H., 2003. Robots as New Media: A Cross-Cultural Examination of Social and Cognitive Responses to Robotic and On-Screen Agents. Conference of the International Communication Association, San Diego.

Short, E., Hart, J., Vu, M., Scassellati, B., 2010. No Fair!! An Interaction with a Cheating Robot. 5th ACM/IEEE International Conference on Human-Robot Interaction, 219-226.

Spencer-Oatey, H., 2000. Rapport management: basic concepts and issues, in: Spencer-Oatey, H. (Ed.), Culturally Speaking: Managing Relations in Talk Across Cultures, London.

Thompson, R., 1994. Emotion Regulation: A Theme in Search of Definition. The Development of Emotion Regulation.

Verhaegh, J., Hoonhout, J., Bingley, P., Meerbeek, B., 2007. iCat as a game buddy for children. 2nd ACM/IEEE International Conference on Human-Robot Interaction, 177-183.

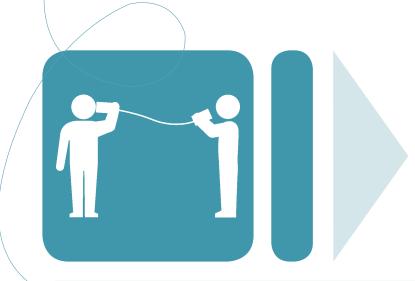
Vink, M., 2011. The iCat as a natural interaction partner:Play Go Fish with a robot., Media and Knowledge Engineering. Delft University of Technology, Delft.

Wang, L., Rau, P.-L.P., Evers, V., Robinson, B.K., Hinds, P., 2010. When in Rome: the role of culture & context in adherence to robot recommendations. Proceeding of the 5th ACMIEEE international conference on Humanrobot interaction, ACM, 359-366.

Weil, M.M., Rosen, L.D., 1995. The psychological impact of technology from a global perspective: A study of technological sophistication and technophobia in university students from twenty-three countries. Computers in Human Behavior, 11(1), 95-133

Woods, S., 2006. Exploring the design space of robots: Children's perspectives. Interacting with Computers, 18, 1390-1418.





# CHAPTER 4

Video-mediated and Co-present Gameplay: Effects of Mutual Gaze on Game Experience, Expressiveness and Perceived Social Presence



# Video-mediated and Co-present Gameplay: Effects of Mutual Gaze on Game Experience, Expressiveness and Perceived Social Presence

# ABSTRACT.

We study how pairs of children interact socially and express their emotions while playing games in different communicative settings. In particular, we study how such interactions can vary for environments that differ regarding the level of mediation and feelings of co-presence. To this end, we investigate how variation in natural eye-gaze influences the players' game experience and perceived sense of presence in a technologically mediated environment, and to what extent this experience resembles that of a natural co-present environment. The results show that eye-gaze is a crucial ingredient for the perceived sense of presence, enjoyment and positive game experience; conversely, the absence of mutual eye-gaze dramatically effects the quality of interaction in the video-mediated environment. The results also show that overall the mediated gameplay with mutual gaze present was appreciated the most, followed by the co-present gameplay, while playing in the mediated environment without gaze turned out to be the least ideal one. The results of this study stress the importance of mutual gaze, and we therefore argue that it should become an integral component of future video mediated communication systems, particularly in those designed for playful settings and children.

Shahid, S., Krahmer, E., Swerts, M. (In Press). Video-mediated and Co-present Gameplay: Effects of Mutual Gaze on Game Experience, Expressiveness and Perceived Social Presence. Interacting with Computers (Elsevier). Earlier versions of this chapter appeared as: Shahid, S., Krahmer, E., Swerts, M., 2011. Using Facial Expressions and Social Presence Self-reports for Investigating the Role of Eye-gaze in Computer-mediated Gameplay. In the proceedings of the Conference of the International Society for Research on Emotions, ISRE'11, Kyoto, Japan.

# 1. Introduction

Video-mediated communication (VMC) systems allow remote users to communicate synchronously using verbal, nonverbal and interpersonal cues (Finn et al., 1997; Giannakos et al., 2011). One of the most important features of the VMC medium is its ability to replicate a communication experience which is closer to face-to-face communication, by supporting the feeling of social presence amongst its users (Kirk et al., 2010). Social presence during mediated communication positively effects the feeling of psychological accessibility and enhances affective connectedness (Biocca et al., 2001; Tu and McIsaac, 2002). In VMC systems, many factors contribute to the creation of social presence (Biocca et al., 2003; Kim et al., 2011). It includes system characteristics such as audio-video quality, audio-video delay, and image size (McCulloch and Oborne, 1999; Tang, 1992), but also the fact that VMC allows of the transmission of visual cues such as gestures, facial expressions and gaze. (Annetta et al., 2009; Bruce, 1996; O'Malley et al., 1996).

In this study, we emphasize one particular visual cue namely 'mutual gaze' (i.e. when two individuals make eye-contact), considered as the most sophisticated form of gaze-awareness (Cook, 1977; Fagel and Blanche, 2010; Kleinke, 1986), and investigate its role in creating social presence and enriching social interaction in VMC systems. It has been argued that, as in face-to-face communication, VMC systems should support mutual gaze between two parties (Grayson and Monk, 2003) because lack of gaze not only effects the system's overall usability and the naturalness of the medium (Ohno, 2005) but also diminishes the sense of intimacy between remote parties (Karahalios, 2004; Mukawa et al., 2005).

However, even though prior research explored the effect of different levels of gaze-awareness on the quality of VMC systems, the findings of these studies have not always been straightforward and a number of important aspects have not received sufficient attention to date. In many studies, the absence of mutual gaze in VMC systems was taken as a limitation of the medium (Mukawa et al., 2005; Vertegaal and Ding, 2002), and most of the time its effect was only measured on the system's usability and users' task performance (Grayson and Monk, 2003). It has also been reported that the success of VMC systems largely depends on the situation in which it is used in and the tasks it is designed for (O'Malley et al., 1996), but the contexts explored in VMC research have often been limited in that they mostly focused on settings where remote participants, generally adults, were involved in professional tasks related to business, work or education (Knoblauch, 1999; Masoodiana et al., 2000).

In this chapter, we focus on mutual gaze in a non-professional, game playing context and study how the presence and absence of mutual gaze during video-mediated gameplay effects the perceived social presence and user experience of physically-apart children. Based on the existing line of research (Bailly et al., 2010; Fullwood and Doherty-Sneddon, 2006; Grayson and Monk, 2003; Monk and Gale, 2002), we predict that the ability or inability to look into the partner's eye (eye-contact), even during playful interactions, can strongly influence the perceived social presence, game experience and player's behaviours. Even though the focus of this study is on the role of mutual gaze during collaborative gameplay, we also compare the mediated gameplay conditions with a copresent condition to see to what extent the mediated experience resembles that of a natural copresent environment in this more playful context.

# 1.1 Video-mediated Communication in Playful Settings

Although many VMC systems have been designed for a variety of purposes and in a number of fields, research in the area of VMC has been primarily directed to work related and business settings (Finn et al., 1997). Examples include systems for distant collaborative meetings and computersupported cooperative work (Panteli and Dawson, 2010), healthcare related activities (Demiris et al., 2009), and for various distance education activities (Fels and Weiss, 2001). Furthermore, a number of high production and resolution systems were designed and deployed by big corporations, such as HP's Halo and Cisco's Tele-Presence systems (Szigeti et al., 2009). These systems were all primarily developed for adults, and associated studies for evaluating these systems also tended to focus on adult user groups (Karahalios and Donath, 2004; Karahalios, 2004; Mueller et al., 2003).

As video technology is becoming more affordable and accessible, VMC systems are moving from typical work settings to home environments for casual, non-work and social interactions (Wheatley and Basapur, 2009). The use of VMC technology for non-work settings, particularly for fun and playful situations such as game playing is still understudied. Furthermore, when it comes to playful situations, children are the natural group to target, but not much research has been conducted with regard to VMC and children (Bobick et al., 1999; Giannakos et al., 2011).

There are only a handful of studies where researchers observed children's gameplay in VMC and face-to-face settings. In one study, (Yarosh et al., 2009) observed parent-child pairs playing a board game in a video-enabled communication environment using a shared tabletop. They found that children were highly engaged with their parents during gameplay and interacted socially. In another study, researchers observed college students playing the social game "mafia" using VMC (Batcheller et al., 2007). The results show that people in the VMC conditions had similar levels of fun, satisfaction, and frustration to those that play while co-present. However, the gameplay via VMC introduced new challenges in terms of managing attention and signaling remote players. In one study, researchers investigated the free game play of young children over a VMC channel (Yarosh et al., 2010). They showed that children were largely successful in playing games using VMC, but that the problems of managing visibility and attention made VMC gameplay more difficult than face-toface gameplay; in the end children preferred collocated play over VMC play.

Although, these previous studies revealed some important aspects of playful interactions during VMC, especially in comparison to face-to-face communication, the findings are arguably still preliminary and indecisive. For example, it remains unclear which factors play a fundamental role in making a VMC experience as playful and fun as a face-to-face experience. Therefore, in this chapter, we focus on young children in a game-playing context and investigate how children play collaborative games and establish a social bond in both mediated and co-present conditions.

#### 1.2 Mutual Gaze

Do we feel our game partner is socially present with us when he/she sends us a message in an online game? Does the level of social presence increase when we are engaged with him/her in an audio only or audio-video call? The concepts of being "socially present" and "socially absent" have become much more relevant after the arrival of different communication systems (Biocca et al., 2003). Many modern communication technologies (including, for example, speech interfaces, agent-based ecommerce systems, virtual environments and so on), despite their diversity, share a common goal, which is to improve social presence during interactions (Biocca and Harms, 2002). Video-mediated communication systems are no exception in this regard.

Video-mediated communication systems capture and transmit a wide range of visual cues during dyadic interactions, and it has been argued that these visual cues contribute considerably to the feeling of social presence among remote participants during mediated interactions (Biocca et al., 2001). It has also been suggested that different visual cues in video-mediated communication are particularly helpful for children because these cues offer better means for conversation and support playfulness more strongly during remote communication (Ballagas et al., 2009). One important visual cue, which received considerable attention in the body of VMC research, is 'mutual gaze'. Mutual gaze and its impact on perceived social presence is also the main focus of our study.

Previous research has shown the crucial role of mutual gaze in establishing a social bond between partners, communicating emotional and interpersonal information, providing feedback, avoiding distraction and regulating conversations (Argyle and Dean, 1965). A review of the HCI literature also shows a strong connection between presence and absence of gaze and perceived social presence particularly in virtual environments (Bailenson et al., 2001), during human-robot interaction (Fong et al., 2003), and during avatar mediated communication (Garau et al., 2001; Steptoe et al., 2010). Mutual gaze seems to have an effect on task performance and perceived social presence during VMC (Bondareva et al., 2006).

Different studies of the development of eye contact behavior in children have also found that eye contact is an important conversation resource for children (Argyle, 1972; Mirenda et al., 1983) and from their early ages they start using it in a systematic manner for regulating conversations (Krantz et al., 1983). The amount of eye contact is lowest during ages 4 to 5 (pre-school) but rapidly increases and reaches the highest level around age 7- 8. It decreases during early adolescence and after the ages 15 to 16 it again gradually increases to a more standard adult level (Ashear and Snortum, 1971; Levine and Sutton-Smith, 1973). Since the use and importance of eye contact seems to vary with different age groups, it is interesting to study, as we do in this chapter, the role of eye contact during VMC with younger children (7-8 year olds).

A number of approaches have been used for evaluating the effect of mutual gaze on different aspects of VMC systems. Most of the time, the effect of mutual gaze has been investigated either on the system's usability (where traditional usability measures are used for evaluating the system) or on the user's performance (where different efficiency related measures are used). In this chapter, contrary to much previous work, we focus on investigating the effect of presence or absence of mutual gaze on perceived social presence of children using direct measures, in particular their non-verbal expressiveness, in playful situations.

# 1.3 Media Capacity Theories

Research on VMC is often couched in terms of media capacity theories, such as social presence theory and media richness theory, which are based on the principle that different media vary in their capacity to carry personal and social communicative cues during remote communication (De Greef and IJsselsteijn, 2001). The social presence theory, introduced by Short, Williams and Christie

(Short et al., 1976), was the first to discuss social presence in the context of mediated environments, and defined it as 'the degree of awareness of another person in the interaction and the subsequent appreciation of an interpersonal relationship'. The research related to social presence is quite diverse and many different definitions and concepts of social presence have emerged since the inception of social presence theory (Ijsselsteijn and Riva, 2003; Waterworth and Waterworth, 2003), but these different definitions usually relate to two perspectives (Weinel et al., 2011), with each perspective having its own pros and cons (Riva et al., 2011).

The first perspective, in line with social presence theory, focuses on the importance of technology and the role of the medium in creating social presence. It looks at the effect of variations in medium settings on social presence (Wise et al., 2004). Many researchers call this type of sense of presence as 'Media Presence' (Lombard and Ditton, 1997; Schloerb, 1995). In our study, we follow this perspective and investigate social presence as influenced by a medium and its underlying characteristics. For example, how does the mere presence of mutual gaze contribute to social presence and social behavior? The other perspective emphasizes the significance of individual differences and perceptions and their impact on social presence (Lee, 2004; Riva and Davide, 2001; Waterworth and Waterworth, 2001). Under this perspective, researchers investigate social presence as the ability of communicators and look at individuals performing different tasks in mediated settings (Rourke et al., 1999). It is also called 'Inner Presence', a more general psychological phenomenon, not essentially linked to experience of medium (Riva et al., 2011).

Another theory, related to the first perspective of social presence, is media richness theory, proposed by (Daft and Lengel, 1984). According to this theory, a "medium's capacity for immediate feedback, the number of cues transmitted, channels utilized, personalization, and language variety" all influence its degree of information richness (Daft and Lengel, 1986). The theory proposes that the task performance will increase if the richness of the medium matches with task needs. According to media richness theory, face-to-face communication has the highest richness and numeric communication (e.g. numeric spread sheets) the lowest. The richness of computer-mediated communication systems falls somewhat on this continuum, where systems showing more visual cues such as VMC systems are richer and closer to face-to-face communication experience and systems without a visual channel, such as audio conferencing and chat systems, are lower in richness. Systems with high media richness are considered much more suitable for complex social tasks such as negotiations or decision-making tasks, where many interpretations of the existing information are possible (Bordia, 1997; Dennis and Valacich, 1999; Steptoe et al., 2010).

In this chapter, we position our research questions in terms of media capacity theories. Media richness theory explains the potential capacity of a medium to communicate different cues (e.g. gaze, wink, lip movement, etc.) via a communication channel and in our study it is evaluated by manipulating different levels of eye gaze in three game playing conditions (VMC with mutual gaze, VMC with no gaze, and co-present). The social presence theory describes the sense of collocation and being there with a remote user during a mediated interaction, and it is evaluated by a set of social presence questionnaires, user's emotional expressions and social behavior exhibited by children during gameplay.

#### 1.4 The current studies

The goal of this chapter is to study the role of mutual gaze in VMC from a fresh perspective: we focus on young (7-8 year old) children that play a card game in one of three conditions that differ in the possibilities for mutual gaze: (1) a condition in which children play via VMC and have the possibilities of mutual gaze, (2) another VMC variant which is exactly like the previous one, except that mutual gaze is not possible, and (3) a condition in which children are physically co-present and interact in a face-to-face manner. We study how the various conditions, and the possibilities for mutual gaze that they offer, influence the extent to which children enjoy the game, feel co-present with their game partner, and how these various conditions influence the non-verbal expressiveness of the children (thereby offering a direct way to measure the impact the conditions have on how children enjoy the gameplay). Expressiveness is studied both by asking judges to interpret non-verbal behaviour in a perception experiment and by annotating different visual cues in the recordings of the children.

The rest of this chapter is organized as follows. In section 2, we describe our data collection procedure, which was based on a card-guessing game. 132 children (66 pairs), equally distributed over the three experimental conditions, played this game. We then present the results of game experience and social presence questionnaires filled out by children after the gameplay sessions. In section 3, we discuss the results of a perception experiment in which we asked adult observers to guess -based solely on the facial expressions of individual children- whether they had just lost or won their most recent game. Observational analysis, which is the topic of section 4, looks at the different visual (positive and negative) features and social cues shown during game playing sessions. Finally, in section 5 we synthesize all results, discuss them in the light of existing findings and draw conclusions for designing better VMC systems for playful settings and children.

# 2. Experiment 1: Mediated and non-mediated game playing

# 2.1 Material: The Card game

In this study, we once again used the card game presented in Chapter 2 and developed in (Shahid et al., 2008a), as a tool for investigating our research questions. In this game, players have to guess whether the next number of a sequence will be bigger or smaller than the previous (reference) number. When the game starts, players see a row of six cards on the screen where the number of the first card is visible ('1' in the case of the example in Figure 1) and the other five cards are placed upside down so the numbers are hidden. All the numbers on the cards are between 1 to 10 and a number displayed once is not repeated in a particular game. Players have to guess whether the number on the next card will be higher or lower than the previous number. Once players have guessed the number, the relevant card is turned around and the correct answer is visible on the screen. Players are also informed about the correctness or incorrectness of their guess via a characteristic non-speech sound. If players make a wrong guess, the game is finished and they move to the next game. If players guess the number correctly then they are asked to guess the next number. Players only win a game if they guess all numbers in the row correctly.

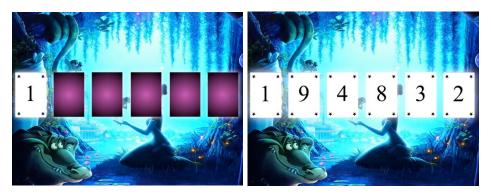


Figure 1: Start of a game (left) and end of the game (right)

The card game was developed using Adobe Flash\*. Appropriate colorful images were chosen for the game background and different animations were used to turn cards around in an attempt to make the game more attractive for children. During the experiment, children played six games, and could in theory win six coins (one for each game that the won). However, unknown to the children, each game was completely deterministic, and two different game alternatives were used. In the first alternative, a rational decision procedure would result in winning the game, and in the second alternative, being rational would result in losing the game. Figure 1 is an example of the losing variant: the most probable outcome for the final card would be a "higher" number than 3, but guessing "higher" would make this a losing game. One winning scenario is shown in figure 2 where 10 is shown on the second last card. No number can be bigger than 10 and that is why children were expected to win this game. Winning and losing games were mixed in the sequence, starting and ending with a variant in which children were likely to win. Before the actual experiment started, children played a practice game. During the gameplay, children were encouraged to discuss every next guess with their partner.

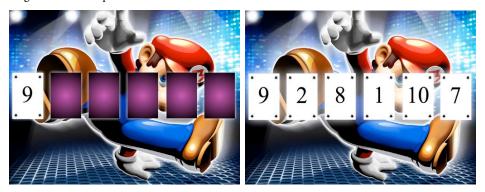


Figure 2: Winning Scenario - Start of a game (left) and end of the game (right)

### 2.2 Participants

88 Dutch children (44 pairs) participated in this study of which 44 children (22 pairs) played the game in the 'mutual gaze' condition and the remaining 44 children played the game in the 'no gaze' condition. To compare the two mediated conditions ('no gaze' and 'mutual gaze') with a non-mediated, co-present condition, we used the data from Chapter 2. In that study, 44 children (22 pairs) participated in the experiment and played the card game while sitting next to each other. All together this resulted in 132 participants spread over three conditions (mediated-mutual gaze, mediated-no gaze, non-mediated/co-present). All children were native speakers of Dutch and were around 8 years old (group 4 in the Dutch elementary school system). In all conditions we balanced for gender (equal number of boys and girls). Parents and teachers gave prior written permission for their child to participate, and signed a consent form stating that the recordings could be used for research purposes.

# 2.3 Setup

#### 2.3.1 Mediated Conditions

The experimental setup was the same for both mediated conditions, except for the gaze manipulation. Children played the game in pairs, and each child was directed to a different room. Communication between the rooms was made possible via a wireless Internet connection. The experimental setup for children was the same in both rooms. Two adults supervised the experiment. During the actual experiment, one was in a separate space, and could control the Card Game and observe the whole experiment, without being noticed by the children. In both rooms, two chairs were placed in front of a desk on which a Windows XP laptop computer was placed, on which the card game ran. Behind the desk, a camcorder was placed on a tripod in such a way that it could record the children's face and upper part of their body. Before the start of each experiment, the camera was adjusted to the height of the participating child.



Figure 3: (Part of the) experimental setup: Position of the card game, camcorder and participant.

The two chairs placed in front of the desk were turned to each other with an angle of about 45 degrees. One chair was reserved for the child participant, on the other chair a 24-inch computer monitor was placed vertically (comparable to children's height). A live video stream of the second participant, sitting in the other room, was shown on this monitor. This monitor was placed in such a way that the participant sitting next to the monitor could see the live video stream properly (figure

4). This was all done to make the sitting environment and experimental setup as similar as possible to the co-present condition.



Figure 4: Mutual gaze Condition: Webcam on the top of the side screen where partner's live video is shown (a rectangle shows the camera position)

The 24-inch monitor, placed on the chair, was attached to a MacBook, which was placed behind the monitor (visible in figure 5). We used iChat (Mac's embedded audio-video conferencing utility) for enabling the audiovisual communication between the two children. A different Mac computer, connected wirelessly with the MacBooks in the room, allowed us to control the connection between the two rooms, without actually entering.



Figure 5: No gaze Condition - Webcam on the top of the front screen where card game is projected (a square shows the camera position)

A high definition (1080p), separate webcam was attached to the MacBooks in the respective rooms. In the case of the mutual gaze condition, it was placed on top of the 24 inch monitor as shown in figure 4. By looking at the monitor where the partner's video steam was broadcasted, the participants were able to see each other's faces simultaneously. In the case of the no gaze condition, the webcam was placed on the top of the windows XP computer showing the card game, as shown in the figure 5. This condition did not allow the participants to see each other's faces simultaneously and therefore no eye-contact was possible. All they could see was partner's left or right side of the face.

# 2.3.2 Co-present Condition

Data for the co-present condition was collected in essentially the same way, except that both children were in the same room. For this a quiet room in the school was used. The children were positioned on two chairs, placed in front of desk on which a computer monitor was placed. The computer monitor was used to show the card game. Behind the computer monitor a camcorder was placed on a tripod in such a way that it could record the children's faces and the upper part of their body. The experimenter was positioned behind a screen, outside the visual field of the participants. A laptop was used by the experimenter to control the game.

#### 2.4 Procedure

The experiment was conducted in three elementary schools located in Tilburg and Helmond, the Netherlands. Children were asked to form pairs themselves. In all conditions, children received instructions in pairs. The experimenter always welcomed the children and started a small discussion by asking a few questions to break the ice ("How old are you? Do you like to play games?" Etc.). After this introductory phase, the experimenter gave spoken instructions, telling the children about the game, the game rules and the coins they could win. In the case of mediated conditions, an audiovideo session was already established and participants were informed about the audiovisual setup and how they could see and talk to each other using this communication system. When the children seemed to understand the rules and the setup, they were asked to take their position. In the mediated conditions, each child was brought to a different room, and asked to sit down on the chair next to the chair containing the monitor showing their friend. In the non-mediated condition, children stayed in the same room and sat down in the chairs next to each other.

When both children were in their chairs, the experimenter started a practice game ("So you only have to say whether the next card is higher or lower. This is just an exercise and it doesn't really count"). After this exercise, the experimenter asked the children whether they had any questions. If not, the experimenter left the children's field of vision and the first experimental game was started. The experiment did not have a fixed duration because different pairs of children took different time. On average, each session lasted for approximately 10 to 15 minutes.

After the sixth and final game, the experimenter asked the participants to fill in a post-experiment questionnaire. After this, the experimenter asked the children some open questions. To finish the whole session, the experimenter congratulated the children and informed them that they could trade in their coins for an individual gift (a key ring or chocolates). Overall, the game worked well. All children made the logical choices they were expected to make in most of the games, so that each pair of children lost at least two games and won at least two games. Furthermore, not a single child reported any suspicion that the game was not real, but a deterministic simulation. Figure 6 shows some representative stills of children winning a game in each of the conditions.





Figure 6: Reaction of children after winning the game in the mutual gaze (top left), no gaze (top right) and copresent (bottom) conditions.

#### 2.5 Measurements

To measure the experience of fun after playing the games in the three conditions (mediated-mutual gaze, mediated- no gaze, co-present), we used the 'funometer', which is a part of the fun toolkit [8]. For analysis, scores of the fun measurements were mapped on a 1-7 point Likert scale, where 1 represented "no fun at all" and 7 represented "a lot of fun". For evaluating the overall game experience, we adapted a questionnaire from (IJsselsteijn et al., 2007; Read, 2008), resulting in eight questions distributed over two dimensions, 1) endurability (e.g. "I would like to play this game again"), and 2) engagement (e.g. "the game was exciting, I was fully involved in the game"), all on 5 point Likert scales.

For evaluating the children's social experience, perceived social presence, and connectedness, we adapted (i.e., selected relevant questions and translated them in Dutch) the network minds social presence measure (NMSPM) for children (Biocca and Harms, 2002), to which we added some questions about partner evaluation (Garau et al., 2001). The adapted NMSPM contained sixteen questions distributed over four categories, 1) perceived co-presence (e.g. "I was often aware of my partner"), 2) perceived message understanding (e.g. "my partner understood what I meant"), 3) perceived affect understanding (e.g. "I understood how my partner felt") and 4) partner evaluation (e.g. "I trusted my partner") all on 5 point Likert scales. For analyses, negatively formulated questions were recoded, so that high numbers always represented a positive score. The NMSPM questionnaire was only administered during the mediated (no gaze and mutual gaze) conditions. Before using this questionnaire during the experiment, we translated it into the Dutch. We then showed the text to a schoolteacher and asked her whether or not children would be able to understand this. After some modifications, we pre-tested the questionnaire with 3 children ( $M_{age} = 7.86$ ). After the final modifications, we used the questionnaire in our study.

# 2.6 Statistical Analysis

To test for significance on the fun scores we conducted three univariate analysis of variance with one between subject factors, namely presence (levels: co-present, mutual gaze, no gaze) and with the mean of the fun, engagement and endurability scores as the dependent variables. All pairwise comparisons were performed using Tukey's HSD.

In the case of NMSPM, we zoomed in on the two mediated conditions (no gaze and mutual gaze). In this case, four separate ANOVAs were performed for each different NMSPM category (perceived co-presence/ connectedness, perceived message understanding, perceived affect understanding, and partner evaluation), and with one between subjects factor, namely presence (levels: no-gaze, mutual gaze). The internal consistency of the NMSPM questionnaires was measured using Cronbach's alpha and was very good (.76 <  $\alpha$  < .87) for four categories. Due to the high Cronbach's alpha values, all items belonging to a particular category were merged for each condition and the average mean of each category is used in the analysis.

#### 2.7 Results

Table 1 summarizes the results on the fun, engagement and endurability questions and shows a rather consistent picture. The children reported having most fun while playing in the mutual gaze condition (M = 6.47, SD = .64), least fun while playing in the no gaze condition (M = 5.57, SD = 0.92) with the co-present condition in between (M = 6.15, SD = .73), F(2,119) = 17.09, p < .001,  $\eta_F^2 = .22$ . All pairwise comparisons are statistically significant.

Fun Toolkit	Conditions				
Categories	No gaze	Mutual gaze	Co-present		
Fun	5.57 (.92)	6.47 (.64)	6.15 (.73)		
Engagement	4.07 (.94)	4.60 (.89)	4.47 (.60)		
Endurability	3.97 (.03)	4.72 (.63)	4.45 (.81)		

Table 1: Average Fun scores with standard deviation (SD) on a 7-point Likert scale (1= very negative, 7= very positive) and average Engagement and Endurability scores (with SD) on a 5-point Likert (1= very negative, 5= very positive).

The results of the engagement dimension showed that children were roughly equally engaged with each other in the mutual gaze (M = 4.60, SD = .89) and co-present (M = 4.47, SD = .60) conditions, while their level of engagement was relatively low in the no-gaze condition (M = 4.07, SD= .94), F(1,119) = 7.93, p < .01,  $\eta_p^2 = .12$ . With the exception of the co-present vs. mutual gaze comparison, all pairwise comparisons are statistically significant.

The results for the endurability (interest in playing the game again) dimension revealed that children who played in the mutual gaze condition showed the strongest interest in playing again (M= 4.72, SD = 0.63), followed by children who played in the co-present condition (M = 4.45, SD = .81), while children in the no gaze condition were least interested in playing again (M= 3.97, SD= 1.03), F (1,119) = 15.72, p < .001,  $\eta_p^2 = .21$ . Pairwise comparisons revealed that all comparisons are statistically significant.

Next we describe the results of the NMSPM scale for mediated conditions (no gaze and mutual gaze) of which table 2 summarizes the results. The results for the perceived co-presence category reveal that children were more involved and felt the presence of their partner stronger when playing in the mutual gaze condition (M = 4.57, SD = .51), than in the no gaze condition (M = 4.02, SD = .51) .73), F(1,79) = 15.34, p < .001,  $\eta_p^2 = .16$ .

We did not find a significant difference between the mutual gaze and no gaze conditions on the message understanding dimension (F < 1). However, children who played the game in the mutual gaze condition indicated that they were better able to understand the emotional state of their partner (M = 4.47, SD = .76), than children who played in the no gaze condition (M = 3.80, SD = .89), F  $(1,79) = 20.66, p < .001, \eta_p^2 = .21.$ 

Finally the game playing condition had a significant effect on the partner's evaluation. Children trusted their game partners more and appreciated their involvement more in the mutual gaze condition (M = 4.63, SD = .69) than in the no gaze condition (M = 4.15, SD = .50), F(1,79) = 13.91, p $<.001, \eta_p^2 = .17.$ 

Social Presence Dimensions	Conditions	
	No gaze	Mutual gaze
Co-presence	4.02 (.73)	4.57 (.51)
Message Understanding	4.30 (.72)	4.41 (.63)
Affect Understanding	3.80 (.89)	4.48 (.76)
Partner Evaluation	4.15 (.50)	4.63 (.69)

Table 2: Mean (and standard deviation) of the four categories of Network Minds Social Presence Measure on a 5point Likert (1= very negative, 5= very positive)

#### 3. Experiment 2: Perception Test

In the previous section we presented children's self-reports on how they appreciated the game playing experience in two mediated conditions, and we also compared the results with those obtained from children playing in the non-mediated (co-present) condition. We found that children in the mutual gaze condition scored significantly higher on fun and endurability than children in the co-present condition. In the case of NMPSP, children in the mutual gaze condition scored significantly higher than children in the no gaze condition on all dimensions of the NMPSP, with the exception of message understanding. But how do the children actually behave in the various conditions? Can we tell from their non-verbal behavior that they have most fun in the mutual gaze condition?

To test this, we ran a perception experiment in which we showed short clips (taken from all three conditions) and presented them to judges, who were asked for each clip to guess whether the child appearing in the clip had just won or lost the game. In this way, we can estimate the expressiveness of the children in the various conditions; more correct guesses indicate that children are more expressive. Running such perception tests is a standard procedure in many research disciplines (e.g., speech and emotion research) and allows us to collect objective judgments of non-verbal behavior.

#### 3.1 Stimuli

We randomly selected 20 pairs of children from each of the mediated conditions (mutual gaze and no gaze) of Experiment I. From these 40 pairs, we selected the response of their first winning game (in which they made a correct prediction for the last card) and the response of their first losing game (in which the final guess turned out to be incorrect). From each pair, we quasi-randomly selected one child, making sure that half of the children came from one room, and half from the other, and that half of the children were boys and half were girls. In all clips, we zoomed in on the face of the child, making sure that the computer screen, placed next to the participants, was never visible in the final clip.

For comparisons, we also used clips of 20 pairs of children from the 'co-present' condition, where winning and losing responses were selected in the same way as above. Furthermore, from the clips of the co-present condition, we randomly selected one child from each pair, and zoomed in on his or her face. Half of the children sitting on the right chair and half of the children sitting on the left chair were selected. In all cases, the video snippets were cut from the moment the final card of the game was turned until the main response of the child was finished. The average length of the stimuli was similar across the three conditions. This resulted in 3 sets (one for each condition), each consisting of 40 clips (2 (win/lose) x 20 children).

#### 3.2 Participants

90 Dutch adults participated in the perception experiments, with a roughly equal number of men and women.



Figure 7: Representative stills of children for winning (right) and losing games (left) during no gaze (top), mutual gaze (middle) and co-present (bottom) conditions

#### 3.3 Procedure

Three group experiments were conducted and in each experiment 30 Dutch viewers viewed one set of stimuli. For every experiment, groups of participants were invited into a quiet classroom where a computer screen was projected on the classroom wall using a beamer. Stimuli were projected without sound, to avoid that participants could pick up auditory cues for winning or losing. Participants were told that they would see a set of 40 stimuli in which children were showing their emotions after winning or losing a game and that their task was to determine whether the children had just won or lost their game. Each stimulus was preceded by a number displayed on the screen indicating the upcoming stimulus, and followed by a six second pause during which participants could fill in their score on the answer form. The actual experiment was preceded by a short training

session in which 3 clips were shown (different from the ones shown in the actual set of stimuli) to make participants familiar with the stimuli and the experimental task. If everything was clear, the actual experiment started. During the experiment there was no interaction between participants and experimenter.

#### 3.4 Statistical Analysis

To test for significance we performed a repeated measurement analysis of variance (ANOVA) with two within-subjects factors, namely game status (levels: win, lost), and gender (levels: boy, girl) and with one between-subjects factor: presence (levels: no gaze, mutual gaze, co-presence) and with percentage of correct classification as the dependent variable. All pairwise comparisons were performed using Tukey's HSD.

#### 3.5 Results

First of all, we found that the average number of correct classifications is highest for children playing in the mutual gaze condition (M = .84), lowest for children playing in the no gaze condition (M = .68), with the scores for children playing in the co-present condition in between (M = .74), F(2,87) = .68, P < .001,  $P_p^2 = .60$  (all pairwise comparisons significant).

In addition, judges found it easier to see whether children lost their game (M = .80) than whether they won (M = .70), F (1,87) = 28.21, P < .001,  $\eta_P^2$  = .245, and judges made more correct classifications for boys (M = .77) than for girls (M = .72), F (1,87) = 12.40, P < .01,  $\eta_P^2$  = .125. Besides these main effects, there was one significant two-way interaction, between game status and presence (F (1,70) = 24.177, P < .001,  $\eta_P^2$  = .257). This interaction can be explained by inspection of Figure 8: there it can be seen that for both no gaze and co-present conditions, losing clips are more often classified correctly than winning clips, but in the case of the mutual gaze condition both winning and losing are classified correctly about equally often.

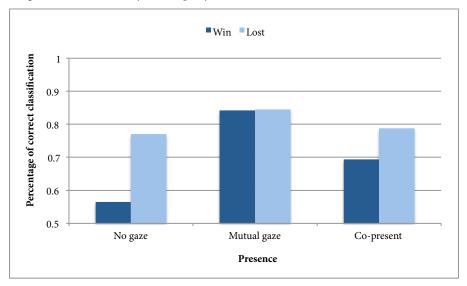


Figure 8: Average percentage of correct classifications as a function of presence and game status.

#### 4. Observational Analysis

The perception tests clearly showed that judges perceive children to be most expressive when playing in the mutual gaze condition, least expressive in the no gaze condition and that their expressiveness is judged to be in between these two extremes while playing in the co-present condition. In this section, we describe a detailed observational analysis of the children's actual non-verbal behavior, and correlate these with the scores from the perception test, to see which cues our judges relied on. We also code cues of their social interactions in the three conditions, to see how these conditions influenced children's social behavior.

#### 4.1 Development of observation scheme

To quantify behavior in terms of (1) their affective responses to winning or losing a game and (2) the children's social interactions during gameplay, we developed two new observation schemes. The first focused on observing and quantifying positive (joy after winning) and negative (disappointment after losing) emotions. This part of the coding scheme was developed on the basis of an existing coding scheme (as discussed in Chapter 3), and extended based on informal observations of video clips taken from the various conditions.

This resulted in two main categories (see appendix 1 for details): 1) positive affect (subcategories: smile, laughter, clapping, jumping/moving, winning arm gesture), and 2) negative affect (subcategories: sadness (frown), anger, closing eyes, covering face, head down). In addition to these visual features, we also added one auditory feature, namely the verbal response to winning or losing. For this part of the coding scheme, a number of facial features are roughly comparable with Action Units (AU's) of the Facial Action Coding Systems (FACS) proposed by (Ekman and Friesen, 1978), e.g. AU 6 & 12 for smile/happiness.

The second part of the coding scheme focused on observing social interaction during gameplay. This scheme was designed on the basis of the informal consultation of video recordings from all conditions. All together this resulted in two categories: 1) connectedness (subcategories: thumbs up, waving, virtual (touching)), and 2) responsiveness (subcategories: chair moved (for sitting closer to the partner), eye fixation (looking at each other for 1.5 or more seconds), turn to partner, and turn taking between players (guessing the upcoming cards one by one)).

We coded the presence and frequency of a feature and the number of times a particular behavior (e.g. smiling, touching, hugging, etc.) occurred in the selected fragments. In a few cases, we coded the duration of the behavior for a certain interval (e.g. eye-fixation). During the coding, the unit of analysis was always one (player).

#### 4.2 Stimuli

The stimuli used for the non-verbal analysis of positive and negative affect were the same as those presented previously in the perception test. For coding social interaction, essentially the same set of stimuli was used, but in a extended duration (including the discussion leading up to the final higher or lower decision) and with sounds not muted. Furthermore, for coding social interaction we did not zoom in on one participant, but looked at both participants and their interaction in the way it was recorded.

#### 4.3 Procedure

The first author and one independent researcher (trained by him) performed the labeling of the observational data. Both coders worked independently, after which the inter-rater reliability for each feature was measured and was found to be substantially high for all categories (ranging from 61% to 92%). The coding was performed in two rounds: first focusing on non-verbal cues related to affect, and then focusing on social cues.

#### 4.4 Data Analysis

To test for significant differences we conducted separate ANOVAs with presence (levels: no gaze, mutual gaze, co-present) as the between subjects factor and mean number of visual features as the dependent variable.

#### 4.5 Results

Table 3 summarizes the number and type of visual features related to affect displayed by children in three game-playing conditions. In general, it can be seen that both for positive and negative effect the number of features is the highest for the mutual gaze category, slightly lower for the co-present category and the lowest for the no gaze category.

Affect	Features	Game		
Category	_	No gaze	Mutual gaze	Co-present
Positive	Smile (happy)	17	22	18
	Laugh	3	14	5
	Clapping	4	13	3
	Jumping/Moving	4	14	9
	Winning gesture	5	13	9
	Verbal response	3	12	11
Т	otal positive Features	36	88	55
Negative	Sad (frown)	19	19	20
	Anger	5	4	4
	Closing eyes	5	14	13
	Covering face	4	11	8
	Head down	4	7	5
	Verbal response	5	15	14
To	otal Negative Features	42	70	64
	Total Visual Features	78	158	119

Table 3: Number of positive and negative features shown by children during three game playing conditions

The one-way ANOVA for the affective dimension revealed that children showed most (non)verbal cues while playing in the mutual gaze condition, the least while playing in the no gaze condition, with the co-present condition was in between, F(2,119) = 14.15, p < .001,  $\eta_p^2 = .19$ . All pairwise comparisons are statistically significant.

We also explored to what extent these visual (positive and negative) cues correlate with the percentage of correct classifications. To do so, we correlated the number of different cues observed in a stimulus with the percentage of correct classifications in the perception test and we indeed found a strong and statistically significant positive correlation between the two (Pearson r = 0.81, N = 120, p < .001). Figure 9 clearly shows that the increase in the amount of visual features present in stimuli gives rise to better classification.

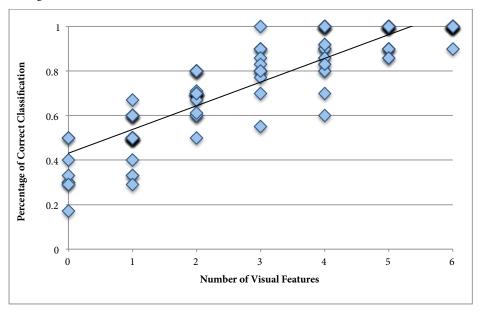


Figure 9: Correlation between percentage of correct classification and number of visual features shown by children

The one-way ANOVA for the social dimension revealed that the number of social cues and behaviors shown by the children in the mutual gaze condition are much higher compared to the no gaze and co-present conditions, F(2,119) = 35.04, p < .001,  $\eta_p^2 = .37$ . With the exception of the copresent vs. no gaze comparison, all pairwise comparisons are statistically significant.

Category	Features	Presence			
	_	No Gaze	Mutual gaze	Co-present	
Connectedness	(Virtual) Touching	1	8	6	
	Thumbs up	3	9	0	
	Waving	2	14	0	
	Total features	5	31	6	
Responsiveness	Stool moved	16	28	9	
	Turn to partner	28	39	24	
	Eye fixation	13	39	15	
	Turn taking¹	23	38	34	
	Total features	80	144	82	
	Total Social features	86	175	88	

Table 4: Number of social cues shown by children during three game playing conditions

#### 5. Discussion

In this chapter, we studied how children experienced playing a card game with a friend, comparing three conditions: two video-mediated conditions, one which allows for natural, mutual gaze, and one which does not, and one condition in which game-playing children were sitting next to each other without a mediating interface. The general results are remarkably consistent: children seem to enjoy playing the card game least in the no gaze condition, while their appreciations for the mutual gaze and co-present condition are comparable (often with a preference for the former over the latter). Below we give a more elaborate discussion of several of our findings.

<sup>&</sup>lt;sup>1</sup> Children were instructed to guess the next number turn by turn (first guess by player A, second guess by player B, so on). We observed whether or not turn taking process was respected during a particular game.

<sup>&</sup>lt;sup>2</sup> Abbreviations: SIKS – Dutch Research School for Information and Knowledge Systems; CWI – Centrum voor Wiskunde en Informatica, Amsterdam; EUR – Erasmus Universiteit, Rotterdam; KUB – Katholieke Universiteit Brabant, Tilburg; KUN - Katholieke Universiteit Nijmegen; OU – Open Universiteit; RUL – Rijksuniversiteit Leiden; RUN – Radbourd Universiteit Nijmegen; TUD – Technische Universiteit Delft; TU/e - Technische Universiteit Eindhoven; UL – Universiteit Leiden;

#### 5.1 Effect on Game Experience and Social Presence

The post game questionnaires revealed that children had more fun, and were more engaged in the mutual gaze than in the no gaze condition, and children in the mutual gaze condition were also more interested in playing again. These results are in line with the media richness theory, and strengthen the results of previous studies where researchers reported similar effects with different user groups and in different domains, i.e., the presence of mutual gaze increases users' satisfaction and pleasantness of tasks. In addition, we found that children in the co-present condition always scored higher for these scales than those in the no gaze condition but not quite as high as those in the mutual gaze condition. However, keeping into account the differences of subjective ratings between co-present and mutual gaze conditions, it is clear that both conditions are closely related to each other with overall high positive scores.

The results of the networked minds social presence measure (NMSPM) scale showed that the absence or presence of mutual gaze during video-mediated game play has a number of effects on the way the games were experienced by the child players. The mutual gaze had a direct effect on the perceived intimacy and closeness of children. Children reported that they felt more strongly connected with each other and felt the presence of their partner in the same physical space in the mutual gaze condition, as compared to the no gaze condition. This is an interesting new finding in the context of children's remote game play, and is in line with the findings of (Bondareva et al., 2006), who showed the direct influence of presence or absence of mutual gaze on perceived social absence of adults.

Results also show that mutual gaze enabled remote players to understand and synchronize each other's emotional responses in an adequate manner. These findings strengthen the already established but less emphasized argument that gaze not only has a communicative function but it also assists in regulating joint arousal, a function which is also closely connected to intimacy and tele-presence (Vertegaal, 1999; Vertegaal and Ding, 2002). Our results show that the findings of previous studies are not restricted to formal tasks performed by adults but can also be seen in more playful situations with children.

However, regarding the 'affective understanding' dimension, we observe that children in the two mediated conditions were able to understand each other's messages without problem. Mutual gaze or lower level of social presence did not have an effect on the perceived message understanding. Although mutual gaze was not present in one condition, other visual cues, system characteristics, and children's own capacity could have helped them in understanding each other (Wise et al., 2004). Furthermore, the task at hand did not require extensive explanation or argumentation, which might have positively effected the 'perceived understanding' scores.

Overall the results of three categories of the NMSPM and their relationship can be interpreted in terms of what Kim and others (Kim et al., 2011) have previously reported, as they showed that social presence is developed on the basis of shared attention, mutual support, open communication and affective connectedness. According to their model, mutual attention and understanding are relatively easy and affective connectedness is quite difficult to achieve in a mediated environment. This is also in line with the argumentation of (Biocca et al., 2003), creators of the NMSPM construct

(Biocca and Harms, 2002), who claim that social presence follows a pattern, that starts with spatial presence and develops into psychological involvement and affective connectedness.

The different game conditions revealed some other social consequences as well. Children trusted their partners more in the mutual gaze condition than in the no gaze condition, and found their partners friendlier and more attentive, compared to the no gaze condition. In the mutual gaze condition, children also respected the logical guesses of partners more often and did not stick to their own guesses even when it was their turn to decide. Apparently, higher level of trust in each other's opinion and friendly attentiveness also led to more games won during the mutual gaze condition. In view of VMC and children's gameplay, it would be interesting to explore the relationship between trust and game status more in future work.

#### 5.2 Effect on Emotional Expressions

In the perception experiment, we found that overall the number of correct classifications of the child data in terms of whether they had won or lost was rather high, showing that facial expressions and other nonverbal features have clear cue value in that respect. Interestingly, the percentage of correct classification was the highest for the mutual gaze condition and the lowest for the no gaze condition, which implies that children were very expressive while playing games in the presence of mutual gaze and absence of mutual gaze had a negative effect on expressiveness. In the co-present condition, children were also rated more expressive than in the no gaze condition but not quite as high as in the mutual gaze condition. The perception test scores are very much in line with the fun scale and fun toolkit self-reports in the sense that the children's appreciation of the game is reflected in the degree to which they become expressive. Moreover, the results of the perception test support the social presence self-reports, particularly in as far as the 'affective understanding' dimension of the NMSPM scale is concerned. In sum: mutual gaze not only had an effect on children's subjective feelings but it also influenced their expressive behaviour.

The perception test also showed that in the case of no gaze and co-present conditions, children were more expressive after losing a game than winning which is in accordance with the results of a previous study (described in chapter 2) using the game play setting (Shahid et al., 2011b) and with data of adult participants (Shahid et al., 2008b). However children were equally expressive after winning or losing a game in the mutual gaze condition, which is not in accordance with our previous findings. One possible explanation for this result could be that this is due to a ceiling effect, namely that in this condition children were so excited about the game that it made them very expressive overall, irrespective of the game status.

The results of the observational analysis also agree with the results of self-reports and perception test scores. Observational codings for the affective dimension revealed that children showed the highest number of visual cues, both negative and positive, in the mutual gaze condition, the lowest number of visual cues in the no gaze condition and the number of visual cues shown in the copresent condition was in between two mediated conditions. There was also a strong positive correlation between percentage of correct classifications and the number of visual features shown by children after winning or losing a game. The results of a correlational analysis not only strengthen the results of perception test but also endorse it as a practical method to get an objective response about the affective states of participants.

The results of the social dimension revealed that children showed more social cues and felt more connected with each other in the mutual gaze condition than in the no gaze and the co-present conditions, whereas the no gaze and the co-present conditions did not differ. Interestingly, only for the mediated gaze condition, children produced typical gestures, like thumbs up and waving, to give feedback or get attention from the other player, which could have strengthened the feeling of social connectedness. Similarly, it turned out that children turned to their partner more often in the mutual gaze condition. Furthermore, in the mutual gaze condition, players always waited for the response of their partner for having a true eye contact (eye-fixation). In the co-present condition, players may have felt less of a need to constantly turn to their partner and establish eye contact as players were sitting side by side in the same physical space and were sure about each other's response. In the case of no gaze condition, children did not turn to their partners very often because they might have realized that turning to partner does not give any additional feedback about the partner's state. In the mutual gaze condition, players may have felt a stronger desire to give each other feedback and synchronize their actions because turning to partner does give additional feedback about the partner's current state. This appears to be in line with existing studies in the context of VMC (Bondareva et al., 2006; Bordia, 1997), where it has been shown that people in mediated conditions (with possible mutual gaze) gesture more (hyper-gesturing) than the co-present condition for conveying the message.

#### 5.3 Mutual Gaze in Co-present and Video-mediated Communication

And finally, it is noteworthy that our analyses revealed that children tend to appreciate the mutual gaze condition more than the co-present one. Both subjective self-reports and objective assessments were higher for the mutual gaze than for the co-present condition. *Prima facie*, these findings seem to contradict a general prediction of the media-richness theory, that face-to-face communication is the ideal form of communication and that people generally prefer it over a mediated communication experience. It should be stressed, however, that with a few notable exceptions (Yarosh et al., 2009; Yarosh et al., 2010), most studies that found evidence for this assumption, were conducted in professional settings with adult participants. We think it is precisely the fact that our experiments involved young participants engaging in playful interaction that accounts for our findings.

There is also a general consensus among researchers in the computer-mediated communication (CMC) research area, that the success of a CMC system largely depends on the type of task it is designed for (Bordia, 1997; Hollingshead et al., 1993). Under right settings and over time the CMC experience can not only match the face-to-face experience (Thurlow et al., 2004), it can even be hyperpersonal (Walther, 1992, 1996). A couple of studies conducted in the gameplay context, one with adults (Bos et al., 2002) and another one with children (Batcheller et al., 2007), also showed that mediated experience can nicely match the face-to-face experience, however with some limitations.

We explain our results along similar lines and argue that the simplicity of the game-playing task, the naturalness of the mutual gaze condition's setup, and its close resemblance with the co-present condition played a key role in shaping the results the way they are. We also conjecture that in the mediated mutual gaze condition, the children have to "work somewhat harder" to communicate than in the face-to-face condition. Obviously, the same applies to the no gaze mediated condition, but since communication is not facilitated there, the children fail in their attempts to properly

communicate. That children work a bit harder can also be seen from the frequency of social cues displayed by children in this condition. Plausibly, the extra effort pays off in fun, engagement and endurability as well and children appreciated the video-mediated mutual gaze condition over the copresence one.

Another possible explanation of these results relates to the novelty of the mediated gaze conditions. The 'novelty effect' is potentially relevant in this case because the child praticipants presumably never played such a game using VMC before and it has been shown that children are usually more positive about new products during their first few uses (Markopoulos et al., 2008). Plausibly, this 'novelty effect' caused children to appreciate the video-mediated mutual gaze condition over the co-presence one. Crucially, however, is the observation that mutual gaze is a necessary precondition for this novelty effect, since children clearly didn't appreciate the no gaze condition.

Interestingly, we found these effects even though the mutual gaze was not perfect, in the sense that players could see each other's gaze without true eye-contact. In other words, even when mutual gaze may not be ideal, it can still improve the mediated communication. See also (Grayson and Monk, 2003), who argue that depending on the task and the context, mutual gaze awareness in typical video conferring systems may support similar communicative functions as true mutual gaze. This is particularly relevant with regard to young children, where previous research shows that children do not always go for direct eye contact. Rather they sometimes seem to prefer "eye-to-face" contact (or focus on a boarder area around eyes) because it is unusual for them to fixate eyes on a small area (Argyle and Cook, 1976).

#### 6. Conclusion

This study aimed to explore the effect of variations in level of gaze on the perceived social presence, emotional expressions, social interaction and game playing behavior of children during video-mediated and co-present communication situations. This study consisted of two experiments where the first experiment was primarily used to collect data. We used a deterministic card guessing game and asked children to play games in three different conditions. After the game playing session, we collected children's subjective response to fun, engagement, and perceived social presence. In the second experiment, we used a perception test, as an objective evaluation method, for measuring the emotional response of children after winning and losing a game. In the last part of this study, we did an observational analysis and coded the video data along affective and social interaction dimensions.

The results show that compared to the no-gaze condition where pairs did not feel a social bond, the feeling of social presence and connectedness was high in the mutual gaze condition and children experienced the mutual gaze condition at least as intensely as co-presence. Furthermore, children had more fun and showed rich social and emotional behavior in the presence of mutual gaze than in the absence of it. Overall the co-present condition was also much more appreciated over the no gaze condition with high self-reports and expressiveness scores but not quite as high as those in the mutual gaze condition. The study has yielded new and interesting insights into the video mediated gameplay of young children. The children we focused on are particularly relevant because during this age (7-8 years) they exploit mutual gaze for supporting and regulating communication. This

study suggests that mutual gaze is a crucial ingredient for establishing a joyful video mediated game play and without it the quality of interaction is dramatically effected.				

## Appendix: Observational Scheme

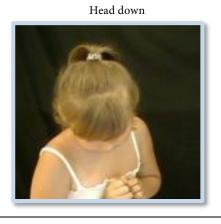
## **Emotional Cues**

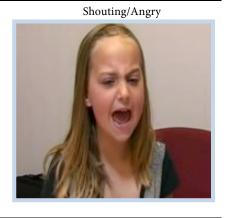












Sad (frown)



Moving/Jumping



Closing eyes



Covering face



## (Social Interaction cues)

(Virtual) Touching



Waving



Thumbs up



Turn to partner /Eye fixation



#### References

Annetta, L.A., Minogue, J., Holmes, S.Y., Cheng, M.-T., 2009. Investigating the impact of video games on high school students' engagement and learning about genetics. Computers & Education, 53(1), 74-85.

Argyle, M., 1972. Non-verbal communication in human social interaction, in: Hinde, R. (Ed.), Nonverbal communication. Cambridge University Press, New York.

Argyle, M., Cook, M., 1976. Gaze and Mutual Gaze. Cambridge University Press, UK.

Argyle, M., Dean, J., 1965. Eye contact, distance and affiliation. Sociometry, 28, 289-304.

Ashear, V., Snortum, J.R., 1971. Eye Contact in Children as a Function of Age, Sex, Social, and Intellective Variables. Developmental Psychology, 43(3), 289-304.

Bailenson, J.N., Blascovich, J., Beall, A.C., Loomis, J.M., 2001. Equilibrium Theory Revisited: Mutual Gaze and Personal Space in Virtual Environments. Presence Teleoperators Virtual Environments, 10, 583-598.

Bailly, G., Raidt, S., Elisei, F., 2010. Gaze, conversational agents and face-to-face communication. Speech Communication, 52, 598-612.

Ballagas, R., Kaye, J.J., Ames, M., Go, J., Raffle, H., 2009. Family communication: phone conversations with children. Interaction Design for Children (IDC 09), ACM, 321-324.

Batcheller, A.L., Hilligoss, B., Nam, K., Rader, E., Rey-Babarro, M., Zhou, X., 2007. Testing the technology: playing games with video conferencing. Proceedings of the SIGCHI conference on Human factors in computing systems (CHI 07), San Jose, CA., ACM Press, 849-852.

Biocca, F., Harms, C., 2002. Defining and measuring social presence: Contribution to the networked minds theory and measure. Proceedings of the 5th International Workshop on Presence, 7-36.

Biocca, F., Harms, C., Burgoon, J.K., 2003. Toward a More Robust Theory and Measure of Social Presence: Review and Suggested Criteria. Presence, 12(5), 456-480.

Biocca, F., Harms, C., Gregg, J., 2001. The Networked Minds Measure of Social Presence: Pilot Test of the Factor Structure and Concurrent Validity Co-Presence. Paper presented at the International Workshop on Presence, Philadelphia, PA, 9-11.

Bobick, A.F., Intille, S.S., Davis, J.W., Baird, F., Pinhanez, C.S., Campbell, L.W., Ivanov, Y.A., Schutte, A., Wilson, A., 1999. The KidsRoom: A Perceptually-Based Interactive and Immersive Story Environment. Presence (Teleoperators Virtual Environments), 8(4), 369-393.

Bondareva, Y., Meesters, L.M.J., Bouwhuis, D.G., 2006. Eye contact as a determinant of social presence in video communication. Human Factors in Telecommunication (HFT).

Bordia, P., 1997. Face-to-Face Versus Computer-Mediated Communication: A Synthesis of the Experimental Literature. Journal of Business Communication, 34, 99-118.

Bos, N., Olson, J., Gergle, D., Olson, G., Wright, Z., 2002. Effects of Four Computer-Mediated Communications Channels on Trust Development. Human Factors in Computing Systems (CHI), ACM Press.

Bruce, V., 1996. The Role of the Face in Communication: Implications for Videophone Design. Interacting with Computers, 8(2), 166-176.

Cook, M., 1977. Gaze and mutual gaze in social encounters. American Scientist, 65, 328-333.

Daft, R., Lengel, R., 1984. Information richness: A new appraoch to managerial behavior and organizational design, in: Cummings, L.L., Staw, B.M. (Eds.), Research in Organizational Behaviour. JAI Press, Homewood, IL.

Daft, R., Lengel, R., 1986. Organizational Information Requirements, Media richness and Structural Design. Management Science, 32(5), 554-571.

De Greef, P., IJsselsteijn, W., 2001. Social Presence in a Home Tele-Application. Cyber Psychology and behaviour, 4(2), 307-315.

Demiris, G., Oliver, D.P., Wittenberg-Lyles, E., Washington, K., 2009. Video-mediated Communication in Hospice Interdisciplinary Team Meetings: Examining Technical Quality and Content. AMIA Annual Symposium proceedings AMIA Symposium AMIA Symposium, 2009, 135-139.

Dennis, A.R., Valacich, J.S., 1999. Rethinking media richness: towards a theory of media synchronicity. Proceedings of the 32nd Annual Hawaii International Conference on Systems Sciences, IEEE.

Ekman, P., Friesen, W., 1978. Facial Action Coding System: A Technique for the Measurement of Facial Movement. Consulting Psychologists Press, Palo Alto.

Fagel, S., Blanche, H., 2010. On the Importance of Eye Gaze in a Face-to-Face Collaborative Task. Proceedings of the 3rd international workshop on Affective interaction in natural environments (AFFINE '10), 81-85.

Fels, D.I., Weiss, P.L.T., 2001. Video-mediated communication in the classroom to support sick children: a case study. International Journal of Industrial Ergonomics, 28, 251-263.

Finn, K.E., Sellen, A.J., Wilbur, S.B., 1997. Video-Mediated Communication. Routledge Publishers.

Fong, T., Nourbakhsh, I., Dautenhahn, K., 2003. A survey of socially interactive robots. Robotics and Autonomous Systems, 42, 143-166.

Fullwood, C., Doherty-Sneddon, G., 2006. Effect of gazing at the camera during a video link on recall. Applied Ergonomics, 37, 167-175.

Garau, M., Slater, M., Bee, S., Sasse, M.A., 2001. The impact of eye gaze on communication using humanoid avatars. Proceedings of the SIGCHI conference on Human factors in computing systems CHI 01, 309-316.

Giannakos, M., Chorianopoulos, K., Johns, P., Inkpen, K., Du, H., 2011. Children's Interactions in an Asynchronous Video Mediated Communication Environment. Interact 2011, Lisbon, Portugal, 199-206.

Grayson, D.M., Monk, A.F., 2003. Are you looking at me? Eye contact and desktop video conferencing. ACM Transactions on Computer-Human Interaction, 10, 221-243.

Hollingshead, A.B., McGrath, J.E., O'Corner, K.M., 1993. Group Task Performace and Communication Technology: A longitudinal Study of Computer-mediated and face-to-face work groups. Small Group Research, 24(3), 307-333.

Ijsselsteijn, W., Riva, G., 2003. Being There: The experience of presence in mediated environments, in: Riva, G., Davide, F., Ijsselsteijn, W. (Eds.), Being There: Concepts, effects and measurements of user presence in synthetic environments. Ios Press, Amsterdam.

IJsselsteijn, W.A., Kort, Y.A.W.d., Poels, K., 2007. The Game Experience Questionnaire: Development of a self-report measure to assess the psychological impact of digital games. TU Eindhoven.

Karahalios, K., Donath, J., 2004. Telemurals: Linking remote spaces with social catalysts. SIGCHI Conference on Human Factors in Computing Systems (CHI), ACM press, 615-622.

Karahalios, K.G., 2004. Social Catalysts: enhancing communication in mediated spaces. Doctoral Dissertation- Massachusetts Institute of Technology, 1-108.

Kim, J., Kwon, Y., Cho, D., 2011. Investigating factors that influence social presence and learning outcomes in distance higher education. Computers & Education, 57, 1512-1520.

Kirk, D.S., Sellen, A., Cao, X., 2010. Home video communication: Mediating "closeness". Proceedings of the 2010 ACM conference on Computer supported cooperative work (CSCW 10), 135.

Kleinke, C.L., 1986. Gaze and eye contact: a research review. Psychological Bulletin, 100, 78-100.

Knoblauch, H., 1999. Video-Mediated Communication. Kathleen E. Finn, Abigail J. Sellen, Sylvia B. Wilbur (eds.). Computer Supported Cooperative Work, 8, 299-301.

Krantz, M., George, S.W., Hursh, K., 1983. Gaze and Mutual Gaze of Preschool Children in Conversation. The Journal of Psychology, 113(1), 9-15.

Lee, K.M., 2004. Presence, explicated. Journal of Communication, 14, 27–50.

Levine, M.H., Sutton-Smith, B., 1973. Effects of Age, Sex, and Task on Visual Behavior During dyadic Interaction Developmental Psychology, 9(3), 400-405.

Lombard, A., Ditton, T., 1997. At the heart of it all: the concept of presence. Journal of Computer Mediated-Communication, 3(2).

Markopoulos, P., Read, J., MacFarlane, S., Hoysniemi, J., 2008. Evaluating Children's Interactive Products: Principles and Practices for Interaction Designers (Interactive Technologies). Morgan Kaufmann.

Masoodiana, M., Apperleya, M., Frederickson, L., 2000. Video support for shared work-space interaction: an empirical study. Interacting with Computers, 7(3), 237-253.

McCulloch, M.S., Oborne, D.J., 1999. Video conferencing Systems: Telepresence and Selection Interviews. Ergonomics Society Annual Conference, Leicester, UK.

Mirenda, P.L., Donnellan, A.M., Yoder, D.E., 1983. Gaze behavior: A new look at an old problem. Journal of Autism and Developmental Disorders, 13(4), 397-409.

Monk, A.F., Gale, C., 2002. A look is worth a thousand words: Full gaze awareness in video-mediated conversation. Discourse Processes, 33, 257-278.

Mueller, F., Agamanolis, S., Picard, R., 2003. Exertion interfaces: sports over a distance for social bonding and fun. SIGCHI conference on Human factors in computing systems (CHI 03), Ft. Lauderdale, Florida, USA 561-568.

Mukawa, N., Oka, T., Arai, K., Yuasa, M., 2005. What is Connected by Mutual Gaze? — User's Behavior in Video-mediated Communication. Extended abstracts on Human factors in computing systems (CHI 05), 1677.

O'Malley, C., Langton, S., Anderson, A., Doherty-Sneddon, G., Bruce, V., 1996. Comparison of face-to-face and video-mediated interaction. Interacting with Computers, 8, 177-192.

Ohno, T., 2005. Weak gaze awareness in video-mediated communication. CHI 05 extended abstracts on Human factors in computing systems CHI 05, Portland, Oregon, 1709-1712.

Panteli, N., Dawson, P., 2010. Video conferencing meetings: Changing patterns of business communication. Information Systems Journal, 16, 88-99.

Read, J.C., 2008. Validating the Fun Toolkit: an instrument for measuring children's opinions of technology. In the Journal of Cognition, Technology & Work, 10(2), 119-128.

Riva, G., Davide, F., 2001. Communications through virtual technologies: Identity, community and technology in the communication age. IOS Press, Amsterdam.

Riva, G., Waterworth, J.A., Waterworth, E.L., Mantovani, F., 2011. From intention to action: The role of presence. New Ideas in Psychology, 29(1), 24-37.

Rourke, L., Anderson, T., Garrison, D.R., Archer, W., 1999. Assessing social presence in asynchronous text-based computer conferencing. Journal of Distance Education, 14, 50-71.

Schloerb, D., 1995. A quantitative measure of telepresence. Presence: Teleoperators, and Virtual Environments, 4(1), 64–80.

Shahid, S., Krahmer, E., Swerts, M., 2008a. Alone or Together: Exploring the Effect of Physical Copresence on the Emotional Expressions of Game Playing Children Across Cultures. In the International Conference of Fun and Games, Eindhoven, the Netherlands, Springer.

Shahid, S., Krahmer, E., Swerts, M., 2008b. Real vs. acted emotional speech: Comparing Caucasian and South-Asian speakers and observers. Proceedings of the Speech Prosody, Brazil, 669-672.

Shahid, S., Krahmer, E., Swerts, M., 2011a. Child-Robot Interaction: How does playing a game with a social robot compare to playing a game alone of with a friend? Submitted.

Shahid, S., Krahmer, E., Swerts, M., 2011b. Child-robot interaction: playing alone or together? . CHI (Extended Abstracts), 1399-1404.

Short, J., Williams, E., Christie, B., 1976. The social psychology of telecommunications. Contemporary Sociology.

Steptoe, W., Steed, A., Rovira, A., Rae, J., 2010. Lie tracking: social presence, truth and deception in avatar-mediated telecommunication. Proceedings of the 28th international conference on Human factors in computing systems CHI 10, ACM, 1039-1048.

Szigeti, T., McMenamy, K., Saville, R., Glowacki, A., 2009. Cisco TelePresence Fundamentals. East, 592.

Tang, J.C., 1992. Why Do Users Like Video? Studies of Multimedia-Supported Collaboration. Computer Supported Cooperative Work, 1, 163-196.

Thurlow, C., Lengel, L., Tomic, A., 2004. Computer Mediated Comunication: Social Interaction and the Internet. Sage, Thousand Oaks, CA.

Tu, C.-H., McIsaac, M., 2002. The Relationship of Social Presence and Interaction in Online Classes. American Journal of Distance Education, 16(3), 131-150.

Vertegaal, R., 1999. The GAZE Groupware System: Mediating Joint Attention in Multiparty Communication and Collaboration. Human Factors in Computing Systems (CHI), ACM press.

Vertegaal, R., Ding, Y., 2002. Explaining Effects of Eye Gaze on Mediated Group Conversations: Amount or Synchronization? ACM conference on Computer supported cooperative work, ACM, 278-285.

Walther, J.B., 1992. Interpersonal effects in Computer-mediated Interaction: A Relational Perspective. Communication Research, 19(1), 52-90

Walther, J.B., 1996. Computer-mediated Communication: Impersonal, interpersonal, and Hyperpersonal Interaction. Communication Research, 23(1), 3-43.

Waterworth, J.A., Waterworth, E.L., 2001. Focus, locus, and sensus: the three dimensions of virtual experience. Cyberpsychology and Behavior, 4(2), 203-213.

Waterworth, J.A., Waterworth, E.L., 2003. The meaning of presence. Presence-Connect, 3(3).

Weinel, M., Bannert, M., Zumbach, J., Hoppe, H.U., Malzahn, N., 2011. A closer look on social presence as a causing factor in computer-mediated collaboration. Computers in Human Behavior, 27, 513-521.

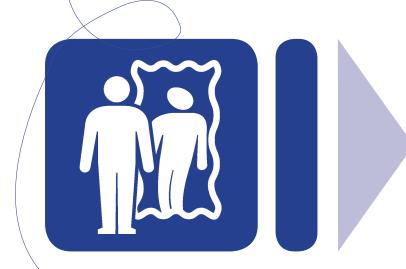
Wheatley, D.J., Basapur, S., 2009. A comparative evaluation of TV video telephony with webcam and face to face communication. Proceedings of the seventh European conference on interactive television (EuroITV 09), Leuven, Belgium ACM Press.

Wise, A., Chang, J., Duffy, T., Valle, R.D., 2004. The Effects of Teacher Social Presence on Student Satisfaction, Engagement, and Learning. Journal of Educational Computing Research, 31, 247-271.

Yarosh, S., Cuzzort, S., Müller, H., Abowd, G.D., 2009. Developing a Media Space for Remote Synchronous Parent-Child Interaction. Interaction Design for Children (IDC), 97-105.

Yarosh, S., Inkpen, K.M., Brush, A.J.B., 2010. Video playdate: toward free play across distance. Interaction Design for Children (IDC), 1251-1260.





# CHAPTER 5

Positive Affective Interactions: The Role of Repeated Exposure and Social Presence.



# Positive Affective Interactions: The Role of Repeated Exposure and Social Presence

#### ABSTRACT.

We describe and evaluate a new interface to induce positive emotions in users: a digital, interactive adaptive mirror. We study whether the induced affect is repeatable after a fixed interval (Study 1) and how social co-presence influences the emotion induction (Study 2). Results show that participants systematically feel more positive after an affective mirror session that this effect is repeatable, and stronger when a friend is co-present.

A journal paper based on this chapter is submitted for publication. Earlier versions of this chapter appeared as: 1) Shahid, S., Krahmer, E.J., Swerts, M.J.G., Melder, W., and Neerincx, M., 2009. You Make Me Happy: Using an Adaptive Affective Interface to Investigate the Effect of Social Presence on Positive Emotion Induction. In the Proceedings of the IEEE 3<sup>rd</sup> International Conference on Affective Computing and Intelligent Interaction, ACII'09. Amsterdam, the Netherlands, and as 2) Shahid, S., Krahmer, E.J., Swerts, M.J.G., Melder, W., and Neerincx, M., 2009. Exploring the Social and Temporal dimension of Emotion Induction using an adaptive Affective Mirror. In the proceedings of the ACM CHI Conference on Human Factors in Computing Systems (CHI EA '09), Boston, USA.

#### 1. Introduction

The days of most people contain one or more episodes of slightly boring inactivity; standing in a queue for a ticket, waiting on a train, sitting in front of a computer before starting another day of work. During episodes such as these, one may get a slight feeling of ennui, and one's mood may start to linger. Wouldn't it be great if there were a way to induce positive emotions in people during situations such as these?

In this chapter we describe and extensively evaluate our attempt of developing such an application: an interactive digital laughing mirror that tries to make people happier. The basic idea is that this mirror deforms a person's face in an adaptive way, during a game-like interaction that lasts only a few minutes. The application could be positioned, say, in a location where queues are likely to form, or activated in the morning on a worker's personal computer, just before starting the work of the day.

Of course, such application contexts raise many difficult questions that need to be addressed before they could become a reality. But they also pose specific evaluation questions: for example, it is not only relevant whether such an application indeed leads to more positive feelings in users the first time they use it, but also whether the effect is repeatable, or gone after the first time. In addition, since there may be other people around in the example scenarios, it is important to see what the effect is of different social settings on the effectiveness of the application. These are the questions that we address in this chapter.

#### 1.1 Positive Emotions

Traditionally, research on emotion in psychology has tended to focus more on negative feelings and emotions (Sheldon and King, 2001) and how to understand and lessen the distress of people (Csikszentmihalyi and Csikszentmihalyi, 2006; Seligman and Csikszentmihalyi, 2000), while positive emotions have generally received less scholarly attention (Fredrickson and Cohn, 2008; Gross, 1999). A similar pattern can be observed in the field of Human-Computer Interaction (HCI), where the main focus is often on the negative feelings and emotions of users when interacting with a computer (e.g., recognizing human frustration) (Fernandez and Picard, 1997; Picard, 1999). There are relatively few examples where researchers primarily focused on positive emotions, either for personalizing the interaction or for improving the overall feeling of human users and developing their lives on a short or long term basis (Calvo and D'Mello, 2010).

This relative lack of interest in positive emotions is somewhat surprising, since it is well-known that positive emotions serve a bigger purpose than merely showing that a person is feeling fine (Fredrickson, 1998). Positive emotions may directly influence a person's behaviour, and help in achieving positive goals (Fredrickson, 2001). It has been shown that positive emotions enhance creativity (Vosburg, 1998; Zhou and George, 2003), trigger optimistic feelings about the self (emotional well being) (Fredrickson and Joiner, 2002), increase self esteem (Ikegami, 2002), and help in achieving both short-term and long-term goals (Loewwnstein and Lerner, 2003; Lyubomirsky et al., 2005).

In view of these positive effects of positive emotions, it would clearly be beneficial if there were ways of inducing positive feelings in persons, yet changing someone's emotional state from negative or even from neutral to positive is not a straightforward task (Seligman et al., 2005). People may use different cognitive and behavioural strategies to increase their everyday feelings of happiness (Boehm and Lyubomirsky, 2009; Sin and Lyubomirsky, 2009). Although this may vary from one person to another, it has been found that factors like 'social affiliation' (spending time and doing activities with friends and family members, supporting and helping others, etc.), 'mental control' (not thinking too much about unhappy things and events), 'active and passive leisure' and particularly 'direct attempts to remain happy' (performing activities which directly involves happiness, laughing, etc.) have a positive effect on emotional state (Boehm and Lyubomirsky, 2009; Tkach and Lyubomirsky, 2006). There are also a number of scientifically proved happiness increasing, positive emotion inducing and sustaining strategies (Lyubomirsky, 2007), such as fostering social relationships, expressing gratitude, savoring, meditative practices, and increasing both short-term and longterm flow experiences (Bryant, 1989; Chesney et al., 2005; Csikszentmihalyi, 1990; Emmons and McCullough, 2003; Gross et al., 2006; Shapiro et al., 2005). To some extent, such strategies have also been used in HCI for building a better user experience. For example, the 'encouraging social relations' strategy has been applied in gaming environments for a better game play (Ochs et al., 2008). Flow experience has also received a lot of attention in HCI and particularly in the gaming industry (Chen, 2007; Sweetser and Wyeth, 2005). Informally, 'flow' can be defined as an optimal state of enjoyment, where participants (players) are completely absorbed in an activity (game playing), and a balance is achieved between the task at hand (challenge) and the user's experience (game playing skills) (Sweetser and Wyeth, 2005).

#### 1.2 Is laughter the best medicine?

Charles Darwin (Darwin, 1972) said that "Joy, when intense, leads to various purposeless movements – dancing about, clapping the hands, stamping, etc., and to loud laughter." Modern research has indeed shown that laughter may work as an effective medicine (stress-reliever) and may also contribute to positive affect (Bennett and Lengacher, 2008; Nichols and Zax, 1977).

Various researchers have studied laughter directly, for instance by eliciting it (Kori, 1989; La Pointe et al., 1990; Mowrer et al., 1987) or by looking at recordings (Trouvain, 2001). Much of this research suggests that spontaneous laughter is the only form of the genuine laughter (Ruch and Ekman, 2001) and that it positively influences the emotional state of humans (Provine, 2000). In addition, the relation between laughter and smiling has been the subject of study (Haakana, 2010) where it was found that smiles may occur as a 'pre-laughing device', and that they can be a used as a substitute for laughter ('silent laughter'), for instance if the social conditions call for this.

In general, there is evidence that laughter is a social phenomenon and that human smiles and laughter may fluctuate with a change in social context (Glenn, 2003; Jefferson, 1979). A humorous film, watched together with friends, elicits more smiles and more laughter, and works better in inducing a positive emotion, than a film watched alone. In different social settings, people tend to regulate their smiles and laughter in an appropriate manner (Preuschoft, 1997), but compared to smiling, facial expressions that supplement laughter are under less deliberate control (Devereux and Heffner, 2007). While there are some studies that addressed the effect of humorous stimuli on eliciting laughter in different social situations (Chapman, 1973; Devereux and Ginsburg, 2001),

(Fuller, 1974), there are to the best of our knowledge no studies that looked into the repeatable nature of these elicitation procedures.

One important thing to note is that laughter in itself is not an emotion; although it is usually taken as an indication of a positive emotional state, it does not necessarily signal such a state. On the one hand, people can find something truly amusing without laughing out loud, while on the other hand, people can produce a fake laugh without finding something amusing or without having any positive feelings. Furthermore, other emotional states such as, for example, nervousness and embarrassment can also give rise to laugher (Vaid and Ramachandran, 2001).

In a somewhat similar vein, it is known that, compared to smiles, it is more difficult to fake laughter (Glenn, 2003). In the case of smiles, usually a distinction is made between natural, spontaneous smiles (the so-called Duchenne smiles) and fake, unfelt smiles (the non-Duchenne smiles). A number of behavioural markers can help differentiate between the Duchenne and non-Duchenne smiles, including the timing and intensity of the smile, and laterality of the mouth, and the appearance of "crows' feet wrinkles" around the eye (Fredrickson, 1998; Soussignan, 2002). For observers, differentiating fake from natural laughter requires more effort because it involves multiple modalities and more complex behavioural markers.

Given the many positive effects that genuine laughter, and positive emotions in general, may have on a person's well-being and productivity, it would be interesting to develop an affective computing application that makes people laugh and feel better. Here we describe and evaluate such an application, with a focus on repeatability and social presence.

#### 1.3 Current Study

In this chapter we describe a new, multimodal affective interface concept: the Affective Mirror (AM), which tries to make users happy, by showing a distorted ("funny") representation of their face. It integrates automatic emotion detection from both face and voice, and uses the fused, perceived emotional state of the user as a "trigger" for selecting different audiovisual effects and for creating a 'flow' experience. The AM may be likened to a traditional carnival mirror, but obviously extends beyond this by its interactive and adaptive nature.

Here, we not only describe the design of this new interface concept but also report on two extensive evaluation studies. We aim to test the effectiveness of the AM in inducing positive emotions and try to gain insight into two specific issues. First, much past experimental work into the expression and detection of emotion has been based on studies that consist of a single experiment where participants are typically tested once through a specific emotion induction procedure (Martin, 1990). Consequently, it is unknown to what extent the reported changes in emotional state in these studies are repeatable. In study 1, we look at the effect of repeated, individual sessions in front of the AM. Second, not much is known about possible social factors that may have an influence on the expression of emotion (Manstead, 2005). Usually, past experiments in this field are conducted with single participants, even though there are indications that the mere presence of another person may have an effect on the extent to which people show their emotions (Coan and Allen, 2007). In Study 2, we look at the effectiveness of the AM as a function of physical co-presence of a friend.

Evaluating affective interfaces is complex and requires more than the "standard" usability evaluations (Isbister and Hook, 2005). In both studies, we collect different kinds of data of participants interacting with the AM, including overall user experience measurements and self-reported emotion scores. In addition, we argue that a lot can be learned from the facial expressions of participants, which is why in both studies we also conduct perception experiments in which independent judges rate the facial expressions of participants.

#### 1.4 Chapter overview

The remainder of this chapter is organized as follows. First, we describe the design and different software and hardware components of the AM. After this we present our first study where we look into the effect of repeated, individual sessions in front of the AM. We report on two experiments where the first experiment measures the induced emotion and the overall user experience from users' self reports and the second, perception experiment concentrates on the facial expressions of participants. We then present our second study in which we look at the effects of physical copresence of a friend on the emotions of users. We finally reflect on both studies in the general discussion section and conclude the chapter by presenting some qualitative analyses and observations.

#### 2. Affective Mirror

#### 2.1 The General Idea of the AM

The Affective Mirror (AM) is an affective multimodal interface that adapts itself to the user's perceived affective state. The purpose of the AM is to try and elicit laughter in people and thereby induce positive emotions in them. While doing this, it also intends to help in collecting spontaneous expressions of positive emotions.

The basic idea of AM is simple: it detects the state of the user and then provides audiovisual feedback in the form of a distortion of the user's face, just like a traditional 'Carnival Mirror'. The AM is different from a traditional 'Carnival Mirror' though, because it is not static but creates an interactive 'production-perception-adaptation' loop thereby creating a rich user experience. The amount and type of facial distortions depend on the detected levels of laughter and smiling. The more participants laugh, the more they progress in different levels of distortions.

The concept of our affective mirror is different from the one introduced by Picard (Picard, 1997) where she used the term for a machine which literally mirrors users' emotions by first understanding a user's emotions and then reflecting them back to a user (for example via an avatar). The persuasive mirror is another, somewhat related interface concept that unobtrusively monitors the behaviour of users and assists them in improving and maintaining a healthy lifestyle (Andrés del Valle and Opalach, 2006). A few other 'distorting mirror' concepts have been developed in the past, such as a 'multimodal caricature mirror' in which an avatar's visual and auditory expressions reflect the user's emotional expressions in an enlarged manner (Olivier et al., 2005) and a 'conventional virtual distortion mirror' that tries to make users laugh by transforming their faces but without involving any real time emotion recognition.

Even though the idea of distorting a user's face has been explored before, the AM is the first system that actively tries to change a user's emotional state in an interactive and adaptive way, and moreover, is the only system whose effect on the emotional state of users has been extensively studied.

The first version of the AM was developed in the MultimediaN project (Melder et al., 2007). For the current version, various new features have been added, in particular for improving the sensing system and the general appearance. In this chapter, we mainly focus on the use and experimental evaluation of the AM, but we do give a general overview of the system's technical components and setup.

#### 2.2 AM's Working Scenario

This section presents a brief interaction scenario between a user and the AM. The user sits in front of the mirror when the session starts. The mirror captures the user's face, and displays a distorted version of the face to the user together with a corresponding funny sound. Typical manipulations are effects where eyes are blown up, or the mouth corners of the user are raised in an extreme fashion, or the face appears in a swirl mode. Such visual distortions tend to generate surprise effects in the user's face, or make the users laugh. These facial expressions are in turn detected by the system, after which the user is confronted with a more extremely distorted version of his/her face.

#### 2.3 Game-like Interaction

The Affective Mirror has a game-like interaction. A reward and penalty system has been implemented in the AM that relies on continuous scoring. An interaction with the AM consists of five levels, and participants require 100 points to move to the next level. One episode of recognized laughter will result in an increase of 10 points. The possibility exists to change these setting, which gives an opportunity to make the mirror more susceptible to laughter.

If participants laughed a lot, then they could move to the next level quickly (reward) but if participants did not laugh then they had to stay in a particular level for a longer duration (punishment). The minimum time in each level was set to 25 seconds and the maximum to 45 seconds. These settings were experimentally decided based on user feedback on development trials.

#### 2.4 System Components and Functionality

The AM senses the user's state by interpreting observational user data. The affect sensing system is based on a visual subsystem and a vocal subsystem that detect smiles and laughter. The affect recognition system captures laughter and affective verbal expressions in the voice, and facial expressions from the frontal video stream. Fusion of the recognition subsystems results in monitoring the overall user experience and adapts the AM to the current user state.

The AM consists of a number of software modules and hardware components. The most important software components are Face Reader (FaceReader, 2008), Laughter Recognition in Speech (LARS) (Truong and Raaijmakers, 2007) and a multimodal fusion module. In this section, we will briefly describe the key subcomponents that were used to evoke and recognize the emotional state of users. The overall structure of all subcomponents and the user-system interaction loop is shown in figure 1.

Even though both FaceReader and LARS sometimes make incorrect predictions about the user's emotional state, this does not damage the interaction. The recognized emotional state merely acts as a trigger for new visual distortions, and is never communicated directly to the user.

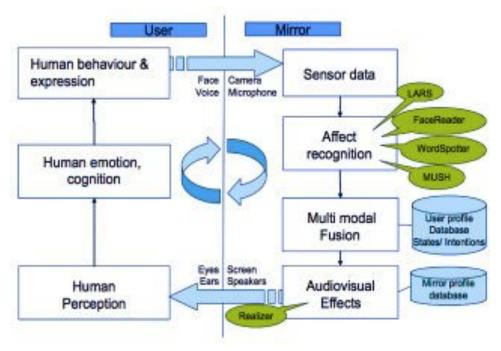


Figure 1: Systematic view of Affective mirror components and underlying interaction loop

#### 2.4.1 Facial Expression Recognizer

We used FaceReader for analysing facial expressions of participants. This software is maintained by Noldus (NoldusTechnologies, 2011) and automatically analyses a person's facial expressions and makes an estimate of a person's emotional state. The software is pre-trained on a database of visual expressions of emotion from a large number of participants.

A live video stream from the user is transmitted to FaceReader, which then analyses it frame by frame and stores the assessment of a person's emotions in a text file. FaceReader also outputs a value on the valence scale (positive–negative). The recognition scores are used in AM to determine whether evidence is found in the face of smiling or laughing. A screenshot of FaceReader while analysing one participant's video stream is shown in Figure .



Figure 2: A screenshot of FaceReader during the AM session. The horizontal green bar (right) shows that the picture quality (input to the FaceReader) is excellent. The vertical green bar (left) shows the perceived valence and the pink horizontal bars (extreme left) show the emotion categories. In this case, FaceReader categorizes the face as 'happy' with a high degree of certainty (purple bar crossing the dotted vertical line).

#### 2.4.2 Laughter Recognition in Speech

The vocal laughter recognizer (LARS) was originally developed at TNO (Truong and Raaijmakers, 2007). It detects auditory laughter by first differentiating silence from speech and laughter and then performing voice activity detection to filter out non-speech and non-laughter sounds. LARS works on the basis of trained acoustic models for laughter, speech and silence, where the ICSI Meeting Recorder Corpus (Janin et al., 2003) was used for this training.

#### 2.4.3 Multi Modal Fusion

In the multimodal expression analysis module, several classification systems are trained to work with multiple modalities for getting a reliable hypothesis on the user state (Truong et al., 2007). In the AM, we developed a dedicated "multi modal fusion (MMF)" module, and we used the decision level fusion technique of (Pantic and Rothkrantz, 2003) for building the MMF. In the decision level fusion, the user state is determined after the subsystems have taken their decision for a single modality.

#### Audiovisual Affective Feedback

To make a person laugh, audiovisual effects are added to the user's mirror image as shown in figure 3. The real-time graphical filters, the geometric image deformations to the user's face, are created with the "Realizer", an audio/video-rendering engine that is part of the KeyWorx platform. KeyWorx is an open source platform to develop, integrate and deploy applications with multi-user/ multimedia features, under Open Source license (KeyWorx, 2011). The "Realizer" also uses Apple Core Image video effects, which requires Apple Mac OSX to run. To achieve a stimulating and adaptive behaviour for the mirror, we created a video effect adaptation that depends on a) time (time elapsed in current user state, time elapsed in current level and time elapsed without laughter are effect modifiers), b) valence (recognized value from the facial subsystem), c) recognized laughter score and d) a random binomial quantity for creating slightly unpredictable distortions.



Figure. 3: Examples of visual distortions created by the AM

#### 2.5 System Architecture

#### 2.5.1 Eye Catcher

To capture participants' live video stream and for displaying their video stream on the screen after appropriate facial distortions, we used the Eye Catcher (QConferencing, 2011). The Eye Catcher is a essentially a videophone with a high-resolution camera installed behind the screen that is invisible to the participants sitting in front of it because of its half-silvered mirror. The camera captures the live video of the participants while they are watching their distorted representation on the mirror. The Eye Catcher provides the AM with a frontal video stream, without the bias of a camera standing next or on top of the screen. This not only improves the performance of FaceReader, but also allows users to have a real 'mirror'-experience of looking straight into their own eyes, without the notion of being recorded by a camera.

#### 2.5.2 Hardware and Software Setup

For running the realizer, we used an iMac with Mac OSX. For running the FaceReader, we used a Dell laptop with Windows XP, service pack 2, and the .NET framework. The modules LARS and MMF were implemented on a Win32 architecture, using the multimedia audio drivers and were also running on a Windows machine. The video signal came from an analogue surveillance camera that is part of the Eye Catcher and positioned behind the half silvered mirror. The camera outputs a y/c (luminance and chrominance) signal or s-video signal, which we split with an s-video splitter. One s-video output was available for the facial expression recognition (SVA) and one to use as the mirror image that the user actually sees (SVB). The SVA output was connected to a USB Video grabber that provided digital video streaming to the laptop. The second output SVB was connected to an Imaging Source video-to-firewire converter that provided the iMac with a digital video stream. The audio signal came from a standard desktop microphone that was connected to the microphone input of a SoundCard. The video stream sent to the laptop (FaceReader) for online analysis was recorded and stored on the windows machines for further offline analysis. Screen capturing software was used for capturing the video of the realizer (distorted faces of participants). Figure gives a behind the scenes impression of the hardware/software setup.



Figure 4: Behind the scenes look at the setting of the Affective Mirror

## 3. Study 1: Repeated Exposure

The aim of study 1 is to find out whether users indeed feel more positive after a session with the AM, and whether this effect is repeatable over time. We also study how users express their emotions during this interaction and whether the repeated exposure influences this expressiveness. We compare a "natural" condition with a control condition in which participants are asked to suppress their laughter. We added this suppress condition to check whether participants would not "fake" their subjective response and laughter (to please the experimenter). In addition, if participants fail to suppress and start laughing, this would be strong evidence for the effectiveness of the AM.

We conducted two experiments in this study. In the first experiment, we exposed participants to the AM in two sessions and we collected their subjective responses about their emotional state and the overall user experience. This experiment also resulted in a rich set of spontaneous expressions of emotion, which was used in the second experiment. In the second experiment, we ran a perception test where judges objectively rated the expressivity of participants on the basis of their facial expressions.

Our hypothesis is that participants will report a positive change in emotions after interacting with the affective mirror in both conditions and after both sessions. The positive change in emotions may be smaller but still substantial after the second session. The adaptive and dynamic nature of AM will amuse participants and it will be difficult for participants to suppress their laughter, particularly in the first session.

#### 3.1 Experiment 1: Data Collection and Subjective AM Experiences

#### 3.1.1 Participants

Participants were 40 (26 females) Dutch undergraduate students (M = 22 years, SD = 3.1 years), who participated for course credits. Participants were randomly assigned to an experimental condition. Participants were recruited using the Tilburg University's participant-pool.

#### 3.1.2 Procedure

The experiment lasted approximately 70 minutes, and was individually performed in an experimental lab as shown in figure 5. To create a pleasant and informal atmosphere, the lab was adorned with a SpongeBob poster and this poster was present in all experimental conditions. Each experimental session consisted of three parts, where parts 1 and 3 provided the repeated interactions with the AM and part 2 consisted of a controlled "waiting period". Upon arrival in the experimental lab, the participant was seated in a comfortable chair facing the Affective Mirror, which was placed on a small table. Two microphones, the experiment task booklet and physiological sensing equipment were also placed on the same table. At this point the mirror was not active.

After seating adjustments, the experimenter introduced himself and briefly described the purpose of the experiment, after which the participant was asked to fill an informed consent form. All participants gave written consent to record and use audiovisual data for research purposes. After this, the participant filled in a self-report emotion questionnaire ("At this moment, I feel...") derived from (Mackie and Worth, 1989), consisting of six 7-point bipolar semantic differential scales with positive and negative adjectives (happy/sad, pleasant/unpleasant, satisfied/unsatisfied, content/discontent, cheerful/sullen, high spirits/low-spirited). The order of the adjectives was randomized; for processing negative adjectives were mapped to 1 and positive ones to 7. Following this, electrodes for measuring galvanic skin response (GSR) and heart rate were attached to the participant and this was followed by a rest period of 5-7 minutes for recording the baseline physiological measurements of the participant (the results of biophysical data analysis are not

presented in this chapter). The participant was also fitted with a tie clip microphone for recording the audio required for the AM.



Figure 5: Individual participant in study 1

After these pre-session measurements, the experimenter started the actual AM session and left the room. In the "suppress laughter" condition, participants were instructed not to laugh during their interaction with the AM. In the natural, "show laughter" condition, participants received no further instructions except that they simply had to look into the Affective Mirror. Each session lasted 3-5 minutes depending on the amount of detected laughter. As soon as the final level was over, the experimenter entered the room and gave the participant the same emotion questionnaire as before.

Next, the participant was asked to fill in additional questionnaires about the overall user experience (UX) and fun of the AM experience, marking the end of the first part. The UX questionnaire was derived from (IJsselsteijn et al., 2007), consisting of eighteen 5-point Likert scale items. These eighteen items were divided into five distinct but related categories: Naturalness of the AM ("The AM was rather artificial..."), enjoyment ("It was fun to sit in-front of the mirror..."), feeling of time ("Time passed quickly/I found it a waste of time..."), play again ("I would like to play again/ playing again would be boring..."), and induce laughter ("The mirror stimulated amusement..."). Positive and negative questions were randomly combined in the questionnaire. For data-processing purposes, negative questions were recoded such that 1 always represents very negative and 5 represents very positive. The Funometer from Read's fun toolkit (Read, 2008) was used for measuring fun.

Next, the experimenter asked the participant to follow him to another room where the controlled waiting period took place, for which the participant was asked to sit in a quiet room for 40 minutes. There they were asked to perform various cognitive tasks, which were neither cognitively overloaded nor emotionally sensitive. After this period, the experimenter asked the participant to follow him back to the room where part 1 had been conducted. Upon arrival, the participant filled in the self-report emotion questionnaire again, after which the second session with the AM started, which proceeded in the way described above. At the end of the second session, participant filled the emotion questionnaire and answered few questions about the overall experience during an informal interview. After the final interview, the participant was debriefed and thanked for the participation.

Finally, participants received a token of appreciation, shown in the figure 6, which consisted of a printed Score Card showing the perceived laughter statistics (amount and duration of recognized laugher from face and voice) together with a funny picture of participant.

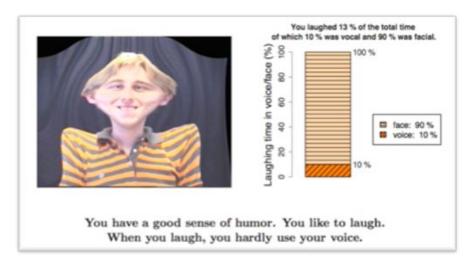


Figure 6: Scorecard presented as a token of appreciation

# 3.1.3 Statistical Analysis and Design

The experiment had a mixed design, with Time of emotion measurement (4 levels: 1, 2, 3, 4, with 1 and 3 pre-measurements and 2 and 4 post-measurements of the two respective sessions) as within-variable and Condition (2 levels: suppress laughter, show laughter) as between-variable and the self-reported emotion scores as well as various user experience scores as the dependent variables. The internal consistency of the self-reported emotion questionnaire was measured using Cronbach's alpha and was very good (.83 <  $\alpha$  < .92). Checks for statistical significance were performed with a repeated measures analyses of variance where the Bonferonni correction was used for making pairwise comparisons. The internal consistency of the user experience (UX) questionnaires was also measured using Cronbach's alpha and was very good (.77 <  $\alpha$  < .85) for all five categories. Due to the high Cronbach's alpha values, all items belonging to a particular category were merged for each condition and the average mean of each category is discussed in the results section. For the fun question, the data from the 'funometer' free scale was mapped on the 1-10 scale where 1 represented no fun at all and 10 represented a lot of fun. Checks for statistical significance on the UX categories were performed using independent sample t-tests.

#### 3.1.4 Results

Self-report emotion questionnaire

Figure 7 summarizes the results and reveals a remarkably consistent picture. First of all, it can be seen that participants' self reported emotion scores were not influenced by Condition, F<1. Thus, participants report the same emotional state, irrespective of whether they were in the "suppress laughter" or in the "natural" condition. However, the time of the emotion measurement did have a strong, significant effect on the self-reported emotion scores, F(3,114) = 131.166, p < .001,  $\eta^2 p =$ .775.

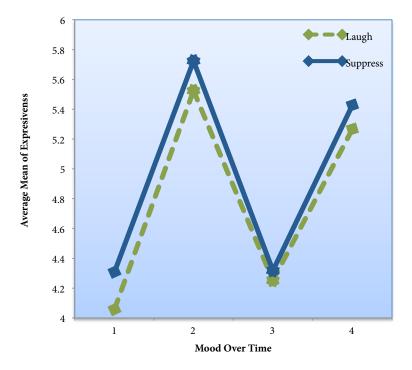


Figure 7: Average self-reported emotion scores as a function of Condition and Time of the emotion measurement. The blue straight line represents the 'suppress laughter' condition and the green dotted line represents the 'show laughter' condition. The x-axis represents the four emotion measurement points and y-axis represents the average self-reported emotion scores on a 7-point scale ranging from 1 (very negative) to 7 (very positive)

Inspection of figure 7 reveals that the first session with the AM worked exactly as intended; participants report more positive feelings after sitting in front of the AM than before (pre-emotion (1): M = 4.18, post-emotion (2): M = 5.63). Right before the second session with the AM starts, the self-reported emotion scores are back to the initial level (pre-emotion (3): M = 4.28), but the second session again works well and afterwards participants report more positive feelings (post-emotion (4): M = 5.36). Pairwise comparisons show that measurements 1 - 2 and 3 - 4 differ significantly (p < 1.06) .001), none of the other comparisons is statistically significant. The interaction between Condition and Time of measurement was not significant, F < 1.



Figure 8: Distorted Face of users (Right), Users' reaction to their distorted face (Left)

## *User experience and fun questionnaire*

Table 1 summarizes the results of the user experience questionnaire. It can be seen that participants scored well above neutral (M > 3) for all constructs, irrespective of condition. This shows that participants were positive about the interaction with the AM and the overall user experience. No effect of condition was found for any of the categories, with the exception of the Enjoyment category where participants enjoyed the experience more in the natural, laugh condition (t (38) = 2.316, p < .05). It was also found that participants in the Laugh condition had more fun (M = 8.45, SD = 1.09) than participants in the Suppress condition (M = 7.50, SD = 1.19) (t (38) = 2.52, p < .01). Apart from this, it was found that the user experience was very similar in the Laugh and the Suppress condition.

Category	Condition	Mean (Std. Dev.)
Laughter Induction	Laugh	3.85 (.489)
	Suppress	3.70 (.544)
Play Again	Laugh	4.05 (.510)
	Suppress	3.80 (.616)
Feeling of Time	Laugh	4.10 (.541)
	Suppress	4.15 (.489)
Naturalness of AM	Laugh	3.90 (.641)
	Suppress	3.65 (.933)
Enjoyment	Laugh	4.35 (.489)
	Suppress	3.90 (.618)

Table 1: Mean (and standard deviation) for five ux questionnaire categories on a 5-point scale (1=very negative, 5=very positive)

# 3.2 Experiment 2: Perception Test

#### 3.2.1 Stimuli

From all the participants of Experiment 1, we selected their response during the last 10 seconds of level 3 of the AM session. This resulted in 2 clips for each participant, one clip from the first session and a second clip from the second session. All together, this resulted in 80 stimuli [20 clips (round 1suppress condition) + 20 clips (round 1 – laugh condition) + 20 clips (round 2 – suppress condition) + 20 clips (round 2 - laugh condition)]. The duration of each clip was approximately 10 seconds. Stimuli were presented to participants in a random order during a group experiment.

#### 3.2.2 **Participants**

35 Dutch adults, with a roughly equal number of men and women, participated in the perception experiments. None had participated in Experiment 1.

#### 3.2.3 Procedure

A group of participants was invited into a classroom where the stimuli were projected on a white screen using an overhead projector. Participants were told that they would see 80 video clips of adults sitting in front of a video camera in different experimental conditions and their job, as viewers, was to rate the expressiveness of these adults based on their facial expressions. No other information about the experiment and different conditions was given to viewers. They were asked to mark the expressiveness using a 7 point Likert scale ranging from 1 (not expressive at all) to 7 (very expressive). Each stimulus was preceded by a number displayed on the screen indicating the upcoming stimulus, and followed by a six second interval during which participants could fill in their score on the answer form. Before the actual experiment, there was a short training session in which 4 clips were shown (different from the ones shown in the actual experiment) to make participants familiar with the stimuli and the experimental task. If there were no further questions, the actual experiment started which lasted for approximately 15 minutes. During the experiment there was no interaction between participants and experimenter. After the experiment, participants were debriefed about the stimuli.

## 3.2.4 Statistical Analysis

To test for significance we performed a repeated measures analysis of variance with two withinsubjects factors, namely repeated-exposure (levels: first session, second session) and condition (levels: show laughter, suppress laughter) and with average mean of expressiveness as the dependent variable.

#### 3.2.5 Results

The results are summarized in figure 9. We found that the average rated expressiveness of participants is higher during session 1 (M = 4.3, SD = 1.02) than during session 2 (M = 3.6, SD = .83), F(1,19) = 11.834, p < .01,  $\eta^2 p = .384$ ). In addition, as one would expect, participants were perceived as more expressive in the 'show laughter' condition (M = 4.2, SD = 1.01) than in the 'suppress laughter' condition (M = 3.7, SD = .82), F (1,19) = 6.153, p < .05,  $\eta^2 p$  = .245. There was no significant interaction effect.

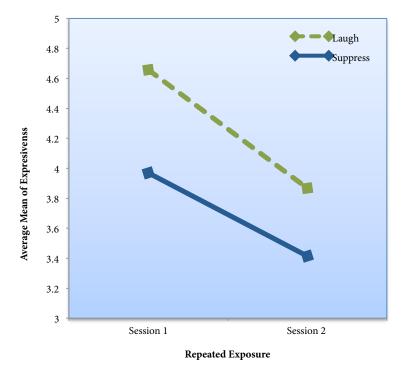


Figure 9: Average perceived expressiveness as a function of Condition and Time of the emotion measurement. The blue straight line represents the 'suppress laughter' condition and green dotted line the 'show laughter' condition. The x-axis shows the two sessions of interaction with the AM and y-axis represents the expressiveness on a 7-point scale ranging from 1 (not expressive at all) to 7 (very expressive).

#### 3.3 Discussion

Study 1 offers a first evaluation of the Affective Mirror, focussing on the effects of repeated exposure. We found that, in general, the interface worked quite well in inducing positive emotions and that participants reported that they indeed felt more positive after their AM session. As one would expect, this is a temporary effect, which was essentially over after the 40-minute controlled waiting interval. However, we also discovered that the positive affect induced by the affective mirror is repeatable, in that participants again reported positive feelings after their second interaction with the Affective Mirror. Interestingly, the results revealed that for the reported emotions it did not matter whether participants were instructed to suppress their laughter or not. In fact, various participants failed to suppress their laughter, which we take as important evidence that the AM indeed elicits spontaneous laughter in participants.

Furthermore, the results from the user experience (UX) questionnaire are consistent with the self-report emotion questionnaire. Participants not only felt better after the AM session but they also indicated that they had fun while interacting with the AM and overall reported a good user experience. Participants appreciated the naturalness of the Affective Mirror, reported that the time passed very quickly during the affective mirror session, showed an interest in sitting in-front of the

mirror again, and reported that the AM was indeed funny and elicited laughter, irrespective of whether they were in the natural or in the suppress condition. As one would expect, participants in the natural condition reported enjoying the interaction more than those in the suppress condition. This result is also in line with results of the fun questionnaire, where participants reported having more fun in the natural condition compared to the suppress condition. These findings are corroborated by the qualitative data collected in the interview afterwards, where one participant explained the reason of having less fun in the suppress condition in this manner "It is an interesting game with one bad rule ... 'Suppress your laughter'. Actually it means: do not play the game. I enjoyed it but could have enjoyed it much more if I was allowed to smile, react, move, whatever!" (Here and elsewhere, we give English translations of Dutch original statements.) Another participants who interacted with the AM in the natural condition appreciated the AM in this manner: "I have an exam in a few minutes. Usually experiments are really boring but this one made me happy. I am feeling good. You should put it outside the examination hall (with smile)."

The second experiment revealed that participants were rated as more expressive during session 1 (particularly in the 'show laughter' condition) than during session 2. If we compare these results with self-report emotion and user experience questionnaires then we can conclude that AM not only 'induced' positive emotions (participants felt positive after the session) but also 'elicited' positive expressions and participants enjoyed the sessions. However, the perception test also suggests that the expressiveness of participants dropped during the second session, even though the self-reports of the participants did not reveal a significant difference between sessions 1 and 2. It is important to realize that the self-reports were collected at the end of a session with the AM, while the stimuli for the perception test were taken from the middle of the interaction, and only show a small part of the interaction, which makes it somewhat difficult to compare the self-reports and judged expressiveness scores.

# 4. Study 2: Physical co-presence

The aim of this study was to see whether there is an effect of physical co-presence on the effectiveness of the AM. For this we compared single participants (who sat in front of the AM alone) with participants who sat in front of the AM in the presence of a friend. Again, we compared a "natural" condition with a control condition in which participants are asked to suppress their laughter. Similar to study 1, we conducted two experiments in this study. In the first experiment, we collected subjective responses of participants about their emotional state and the overall user experience. This experiment once again resulted in an extensive dataset of spontaneous expressions of emotion, which was sampled for the second experiment, in which judges rated the expressivity of participants on the basis of their facial expressions.

We hypothesize that co-presence will have a positive effect on the induced emotions (participants are expected to report feeling more positive) and will result to more visual expressiveness. We also hypothesize that it will be more difficult for participants to suppress their laughter when sitting together with a friend than when they are alone, which is also likely to influence the expressiveness scores.

# 4.1 Experiment 3: Data collection

## 4.1.1 Participants

Participants were 94 (55 female) Dutch undergraduate students (M = 21 years, SD = 2.4). Of these, 54 participated in self-selected pairs, consisting of friends. Pairs were randomly assigned to either a suppress (N = 14 pairs) or a natural condition (N = 13 pairs). The 40 individuals were those of Study 1, where we only used data from their first session. All participants received a course credit for doing this experiment.

#### 4.1.2 Procedure

The procedure of experiment 3 was essentially the same to that of experiment 1 (session 1), which allows us to compare data from individuals with data from pairs. In the case of pairs of participants, it was first decided which participant would sit in front of the AM first. The other participant sat on the left side of the participant in front of the AM, such that the other participant could look at both the friend's face and at the distortions of the affective mirror. The setup is shown in figure 10. After this session, which proceeded exactly as described for study 1 (including measurements and questionnaires), the two participants switched positions, and another session with the AM was initiated in exactly the same way as before. When both participants had been in front of the AM, they were debriefed and thanked for their participation with an individual score card.



Figure 10: Pair of participants in study 2.

## 4.1.3 Statistical Analysis and Design

The internal consistency of the self-reported emotion questionnaire was measured using Cronbach's alpha and was again very good (.82 <  $\alpha$  < .93). Two separate analyses were run. First we analysed the data from the pairs in a mixed between-within design, with Time of emotion measurement (2 levels: pre and post) as within-variable and Condition (2 levels: suppress laughter, show laughter) and Turn (2 levels: first, second) as between variables. Next we compared the data from individuals and pairs in a comparable mixed between-within design, with Time of emotion measurement as within-

variable, and Co-presence (2 levels: alone or together) as between-variables and the self-reported emotion scores and the various usability experience questions as the dependent variables. Checks for statistical significance were performed with repeated measures analyses of variance in the case of the self-report emotion scores. In the case of user experience (UX) questionnaire, only the data from the pairs was analyzed (data from individual participants have already been discussed above). The internal consistency of the UX questionnaires was measured using Cronbach's alpha and was very good (.80 <  $\alpha$  < .88) for four categories with an exception of  $\alpha$  = .69 for the 'Naturalness of the AM' category. Due to the high Cronbach's alpha values, all items belonging to a particular category were merged for each condition and the average mean of each category is discussed in the results section. Data from the 'funometer' scale were analysed as before.

#### 4.1.4 Results

Self-report emotion questionnaire

Analysis of the pairs' data revealed that there was no significant effect of Condition nor of Turn (both F<1). These two factors did not interact with any of the other factors. In other words, for the self-reported emotion scores it does not matter whether participants from pairs sat in front of the mirror first or second. In addition, as in study 1, it does not matter whether participants were asked to suppress their laughter or not. Based on these results, we aggregated the data of all participants across Condition and Turn for further analysis.

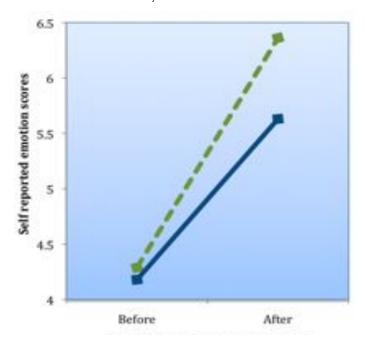


Figure 11: Comparison of the changes in self-reported emotion scores for members of pairs and for individual participants. The blue straight line represents the emotion scores for individuals and green dashed line represents the pairs. The x-axis represents the two emotion measurement points (before and after a sessions with the AM) and the y-axis represents the average self-reported emotion scores on a 7-point scale ranging from 1 (very negative) to 7 (very positive)

Figure 11 summarizes the results for the comparison between pairs and individuals. It can clearly be seen that participants report overall more positive scores after their session with the AM (M = 6.36) than before (M = 4.29), F(1,92) = 957.170, p < .001,  $\eta^2 p = .912$ ). Interestingly, this effect is stronger for participants who participated with a friend than for single participants, as shown by the significant interaction between Time of the emotion measurement and co-presence, F(1,92) = 29.249, p < .001,  $\eta^2 p = .241$ .

User experience and fun questionnaire

Table 2 summarizes the results of the user experience questionnaire for pairs of participants. Again, no significant effect of the Turn was found for any category of the UX questionnaire. Based on this result, we once more aggregated the data of all participants across Turn.

Category	Condition	Mean (Std. Dev.)
Laughter Induction	Laugh	4.29 (.460)
	Suppress	4.04 (.720)
Play Again	Laugh	4.39 (.629)
	Suppress	4.00 (.693)
Feeling of Time	Laugh	4.11 (.497)
	Suppress	4.23 (.765)
Naturalness of AM	Laugh	3.75 (.645)
	Suppress	3.69 (.618)
Enjoyment	Laugh	4.50 (.509)
	Suppress	4.12 (.516)

Table 2: Mean of five UX questionnaire categories on a 5-point scale (1=very negative, 5=very positive) for Pairs

It can be seen that participants scored above neutral (3) for all constructs, irrespective of whether they participated in Laugh or Suppress condition. This shows that participants were positive about the interaction with the AM and the overall user experience. A t-test revealed a significant effect of the condition for the 'enjoyment'  $(t\ (52) = 2.756,\ p < .01)$  and 'feeling of time' category  $(t\ (38) = 2.184,\ p < .05)$ . No effect of condition was found for the other three categories. Finally, it was found that participants in the Laugh condition had somewhat more fun  $(M = 8.96,\ SD = .90)$  than participants in the Suppress condition  $(M = 8.15,\ SD = 1.02)$   $(t\ (52) = 2.966,\ p < .01)$ .

# 4.2 Experiment 4: Perception Test

### 4.2.1 Stimuli

From all the pairs that participated in Experiment 3, we randomly selected 20 for the perception test (10 from each condition). For both participants in each pair, we selected their response during the last 10 seconds of the level 3 of the AM game. This resulted in 2 clips for each pair: one clip from the first turn (participant 1) and the other clip from the second turn (participant 2). For individuals, the stimuli from the Experiment I were used. This resulted in 80 stimuli: 40 from Pairs [10 clips (turn 1-suppress laughter) + 10 clips (turn 2 - suppress laughter) + 10 clips (turn 3 - show laughter) + 10 clips (turn 4 - show laughter) + 10 clips (turn 5 - show laughter) + 10 clips (turn 6 - show laughter) + 10 clips (turn 7 - show laughter) + 10 clips (turn 8 - show laughter) + 10 clips (turn 9 - show laughter) + 10 clips (turn 1 - show laughte

clips (turn 2 - show laughter)] + 40 from individuals [20 clips (round 1- suppress condition) + 20 clips (round 1 – laugh condition)].

#### 4.2.2 **Participants**

35 Dutch adults, with a roughly equal number of men and women, participated in the perception experiment. None participated in any of the other experiments.

#### 4.2.3 Procedure

The procedure of was the same as that of experiment 2 (study 1).

#### 4.2.4 **Statistical Analysis**

Tests for significance were performed using a repeated measurement analysis of variance with two within-subjects factors, namely co-presence (levels: alone or together) and condition (levels: show laughter, suppress laughter) and with average mean of expressiveness as the dependent variable.

#### 4.2.5 Results

The results, summarized in figure 12, show that participants who sat in front of the AM together with their friends were perceived as more expressive (M = 6.1, SD = .73) than those who sat alone in front of the AM (M = 4.2, SD = .93), F(1,19) = 52.810, p < .001,  $\eta^2 p = .735$ ).

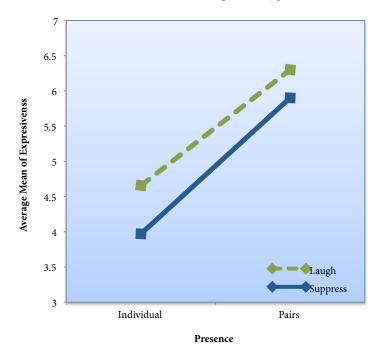


Figure 12: Mean of expressiveness as a function of Condition and co-presence. The blue (bottom) line represents the 'suppress laughter' condition and green (top) dotted line represents the 'show laughter' condition. The x-axis represents the two variations in co-presence and y-axis represents the mean of expressiveness on a 7-point scale ranging from 1 (not expressive at all) to 7 (very expressive)

Even though participants in the suppress condition were perceived as slightly less expressive, the difference was not significant (F < 1), and neither was the interaction between co-presence and condition. We separately tested (for the members of pairs of participants) whether it mattered whether they interacted with the AM during the first or second turn, but no difference in expressiveness was found.

## 4.3 Discussion

In general, the analysis of the self-reported emotion scores in study 2 confirm the findings of study 1: the affective mirror succeeds in inducing positive emotions in participants, irrespective of whether they are in the "natural" condition or in the "suppress laughter" condition. In addition, this study showed that the effect is stronger for people who interacted with the AM with a friend present, suggesting that social factors strengthen the effectiveness the AM. These results are consistent with the results found by (Wagner and Lee, 1999; Wagner and Smith, 1991) that doing an activity together with a friend results in feeling better than doing an activity alone. These results also strengthen the results of other studies (Gajadhar et al., 2008; Shahid et al., 2008) where the effect of social presence on game players was investigated and it was revealed that the game experience and emotional response increase when a friend is physically present.

The analysis of the UX and fun questionnaire score in study 2 is also in line with that of study 1: participants appreciate the interaction with the AM. They reported that they found the interaction with the AM hilarious, engaging, natural and, most importantly, fun and enjoyable. Pairs of participants in the natural condition did enjoy it more than those in the suppress condition, for reasons similar to those expressed in study 1. Comparing Table 1 (mean score for individuals) with Table 2 (mean score for pairs) reveals that participants who interacted with the AM in pairs were more positive about the AM than those who participated alone. They enjoyed their session more, showed a stronger interest in playing again, and were highly convinced that the AM induces laughter and is fun. Interestingly, participants in the suppress condition reported that it that it was difficult to suppress their laughter, which we take as evidence that the AM indeed elicits laughter to some extent.

These findings are also in line with qualitative findings reported during post-session interviews. One participant from a pair said "It is simply not possible to continuously suppress your laughter and keep a straight face. Even if you are very good in controlling, you will still smile and break once." People were less critical about the naturalness of the suppression task while sitting in pairs. Another participant said, "Gosh! This mirror inspects and maybe understands me. After a long suppression it does something magical and I break." Both participants of a pair equally enjoyed the AM, irrespective of their turn. In some sense this can be interpreted as a repeated exposure effect (each pair participates in two sessions with the AM), and the results suggest that the effect of repeated exposure might be stronger for pairs, even when the gap between exposures during study 2 is much shorter than the gap during study 1. One participant said, "My turn was second. (...) The first session was hilarious. I thought when I would sit, it would be less funny but that was not the case. There were many new visual effects and it looked like a new experience."

The second experiment revealed interesting results about the emotional expressions of participants: clearly participants in pairs were much more expressive than individual participants, which is perfectly in line with the self-reports, and confirms that social co-presence strengthens the effectiveness of the AM. Participants in pairs not only felt happier after their session with the AM, it was also easily detectible from their facial expressions.

### 5. General Discussion and conclusion

The Affective Mirror is a novel interface concept for inducing happiness in participants. The AM creates a game-like situation and adapts itself intelligently based on the user's perceived current affective state. Two experimental evaluation studies revealed that participants indeed report feeling more positive after a session in front of the AM, that they enjoyed interacting with the AM and were quite expressive during their interaction with the AM. Study 1 revealed that this effect is repeatable in the sense that the AM induces positive emotions after both sessions, although the expressiveness of participants decreases over time.

Study 2 showed that the positive effect of the AM is stronger for pairs of participants than for individual participants; when people participate in pairs, they report feeling happier and enjoying the interaction more, and independent judges rate them as more expressive. In both studies we asked half of the participants to suppress their laughter. This was done as a control, on the one hand to avoid "demand effects" (the AM is trying to make people laugh, and we want to avoid participants from laughing because they suspect that that is what the experimenter demands from them), but also to see whether the AM is funny enough to make people laugh even when they are instructed not to. The latter was clearly the case, which lends further support to the general conclusion that the AM succeeds in making people laugh and feel happier.

A closer inspection of the participants' behaviour in front of the AM also offers additional evidence of this. One striking thing that can be observed is that participants frequently respond to the AM's transformations by making "weird" faces themselves (figure 13). In addition, they moved their body in and out of the screen, and moved their arms and torso. These playful interactions were more frequent in study 2. One participant said, "One of the reasons I like this is mirror is that I can make my own distortions and transformations while playing with the mirror."

Interestingly, suppressing laughter appeared to be more difficult in the presence of friend (study 2). It was frequently clear from participants' facial expressions that they were indeed trying to suppress their laughter, and were not always successful (Fig. 14). Participants in pairs also consulted and occasionally communicated with their partners during the interaction and tried to involve them in the whole experience.



Figure 13: Top: Participants playing with the AM and transforming their faces in response to AM's distortions. Bottom: Participants physically moving their position and bodies while interacting with AM during show laughter condition



Figure 14: Participant trying to suppress laughter in study 1 (top-left), participant laughing after failing to suppress laughter in study 2 (top-right), participants using their hands to hide their laughter after failing to suppress it during study 2 (bottom row),

In conclusion, both the quantitative and qualitative analyses show that the AM is an effective tool inducing feelings of happiness and eliciting laughter. This can be illustrated nicely with a quote from one participant who claimed that the AM is a real mood changer and said, "This mirror is like an Aspirin... a tablet for happiness"



Figure 15: Spontaneous expression of emotion while interacting with the AM (study 1).

## References

Andrés del Valle, A.C., Opalach, A., 2006. The Persuasive Mirror. Proceedings of International Conference on Persuasive Technology, Eindhoven, the Netherlands.

Bennett, M., Lengacher, C., 2008. Humor and Laughter May Influence Health: III. Laughter and Health Outcomes. Evidence-Based Complementary and Alternative Medicine, 5(1), 37-40.

Boehm, J.K., Lyubomirsky, S., 2009. The promise of sustainable happiness, in: Lopez, S.J. (Ed.), Handbook of positive psychology 2nd ed. Oxford University Press, Oxford.

Bryant, F.B., 1989. A Four-factor Model of Perceived Control: Avoiding, Coping, Obtaining, and Savoring. Journal of Personality and Social Psychology, 57(1), 773-797.

Calvo, R.A., D'Mello, S., 2010. Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications. IEEE Transactions on Affective Computing, 1(1), 18-37.

Chapman, A.J., 1973. Social facilitation of laughter in children. Journal of Experimental Social Psychology(9), 528-541.

Chen, J., 2007. Flow in games (and everything else). Communications of the ACM, 50(4).

Chesney, M.A., Darbes, L.A., Hoerster, K., Taylor, J.M., Chambers, D.B., Anderson, D.E., 2005. Positive emotions: Exploring the other hemisphere in behavioral medicine. International Journal of Behavioral Medicine, 12(2), 50-58.

Coan, J.A., Allen, J.J.B., 2007. Handbook of Emotion Elicitation and Assessment. Oxford University Press.

Csikszentmihalyi, M., 1990. Flow: The psychology of optimal experience.

Csikszentmihalyi, M., Csikszentmihalyi, I.S., 2006. A life worth living : contributions to positive psychology. Oxford University Press, Oxford; New York.

Darwin, C., 1972. The expression of the emotions in man and animals, 3rd ed. Oxford University Press.

Devereux, P.G., Ginsburg, G.P., 2001. Sociality effects on the production of laughter. The Journal of General Psychology, 128(2), 227-240.

Devereux, P.G., Heffner, K., 2007. Psychophysiological Approaches to the Study of Laughter: Towards an Integration with Positive Psychology, in: Ong, A.D., Van Dulmen, M. (Eds.), Handbook of Methods in Positive Psychology. Oxford Publisher, pp. 223-249.

Emmons, R.A., McCullough, M.E., 2003. Counting blessings versus burdens: Experimental studies of gratitude and subjective well-being in daily life. Journal of Personality and Social Psychology, 84(1), 377-389.

FaceReader, 2008. http://www.noldus.com/human-behavior-research/products/facereader.

Fernandez, R., Picard, R.W., 1997. Signal Processing for Recognition of Human Frustration. IEEE ICASSP, Seattle, WA. .

Fredrickson, B.L., 1998. What Good Are Positive Emotions? . Review of General Psychology 2(3), 300-319.

Fredrickson, B.L., 2001. The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. American Psychologist, 52(2), 218-226.

Fredrickson, B.L., Cohn, M.A., 2008. Positive Emotions, in: Lewis, M., Barrett, L.F. (Eds.), Handbook of Emotions. Guilford Press, New York.

Fredrickson, B.L., Joiner, T., 2002. Positive Emotions Trigger Upward Spirals Toward Emotional Well-Being. Psychological Science, 13(2), 172-175.

Fuller, R.G., 1974. Effects of group laughter on responses to humorous material: A replication and extension. Psychological Reports.

Gajadhar, B.J., Kort, Y.A.W.d., IJsselsteijn, W.A., 2008. Influence of Social Setting on Player Experience of Digital Games, CHI 2008 Proceedings - Works In Progress, Florence, Italy, 3099-3104.

Glenn, P., 2003. Laughter in Interaction. Cambridge University Press.

Gross, J.J., 1999. Emotion regulation: Past, present, future. Cognition and Emotion 13(5), 551-573.

Gross, J.J., Richards, J.M., John, O.P., 2006. Emotion regulation in everyday life, in: Snyder, D.K., Simpson, J.A., Hughes, J.N. (Eds.), Emotion Regulation in Families: Pathways to Dysfunction and Health. American Psychological Association, Washington DC.

Haakana, M., 2010. Laughter and smiling: Notes on co-occurrences Journal of Pragmatics, 42(1), 1499-1512.

IJsselsteijn, W.A., Kort, Y.A.W.d., Poels, K., 2007. The Game Experience Questionnaire: Development of a self-report measure to assess the psychological impact of digital games. TU Eindhoven.

Ikegami, T., 2002. The role of state self-esteem in positive mood effects on person impression: When does a positive mood lead to a favorable view of others? Japanese Psychological Research, 44, 20-33.

Isbister, K., Hook, K., 2005. Evaluating Affective Interfaces: Innovative Approaches, A workshop in the CHI 2005...

Janin, A., Baron, D., Edwards, J., Ellis, D., Gelbart, D., Morgan, N., Peskin, B., Pfau, T., Shriberg, E., Stolcke, A., Wooters, C., 2003. The ICSI meeting corpus. Proceedings of IEEE International conference on acoustics, speech and signal processing, 364-367.

Jefferson, G., 1979. A technique for inviting laughter and its subsequent acceptance declination, in: Psathas, G. (Ed.), Everyday Language Studies in Ethnomethodology. Irvington, New York, pp. 79-96.

KeyWorx, 2011. http://www.keyworx.org/.

Kori, S., 1989. Perceptual dimensions of laughter and their acoustic correlates. Intern. Confer. of the Phonetic Sciences Tallinn 255-258.

La Pointe, L.L., Mowrer, D.M., Case, J.L., 1990. Comparative Acoustic Analysis of the Laugh Responses of 20 Year Old and 70 Year Old Males. International Journal of Aging and Human Development, 31(1), 1-9.

Loewwnstein, G., Lerner, J.S., 2003. The Role of Affect in Decision Making, in: Davidson, R.J., Scherer, K.R., Goldsmith, H.H. (Eds.), Handbook of affective science. Oxford University Press, New York, pp. 619 - 642

Lyubomirsky, S., 2007. The how of happiness: A new approach to getting the life you want. Penguin Books, New York.

Lyubomirsky, S., King, L., Diener, E., 2005. The Benefits of Frequent Positive Affect: Does Happiness Lead to Success? . Psychological Bulletin, 131(6).

Mackie, D., Worth, L., 1989. Processing deficits and the mediation of positive affect in persuasion. Journal of Personality and Social Psychology, 57(1), 27-40.

Manstead, A.S.R., 2005. The social dimension of emotions. The Psychologist, 18(8).

Martin, M., 1990. On the induction of mood. In the Clinical Psychology Review, 10(6).

Melder, W.A., Truong, K.P., Uyl, M.D., Leeuwen, D.A.V., Neerincx, M.A., Loos, L.R., Stock Plum, B., 2007. Affective multimodal mirror: sensing and eliciting laughter, In the International workshop on Human-centered Multimedia. ACM Press, 31-40.

Mowrer, D.E., Lapointe, L.L., Case, J., 1987. Analysis of five acoustic correlates of laughter. Journal of Nonverbal Behavior, 11(3), 191-199.

Nichols, M.P., Zax, M., 1977. Catharsis in psychotherapy. Gardner Press, New York.

NoldusTechnologies, 2011. http://www.noldus.com.

Ochs, M., Sabouret, N., Corruble, V., 2008. Modeling the Dynamics of Non-Player Characters' Social Relations in Video Games. AAAI conference on Artificial Intelligence and Interactive Digital Entertainement, AAAI Press, 150-158.

Olivier, M., Benoit, M., Irène, K., Arman, S., Jordi, A., 2005. Multimodal Caricatural Mirror. 1st Summer Workshop on Multimodal Interfaces - ENTERFACE, 13-20.

Pantic, M., Rothkrantz, L.J.M., 2003. Towards an affect-sensitive multimodal human-computer interaction. Proceedings of the IEEE, 91(9), 1370 - 1390

Picard, R., 1997. Affective Computing. The MIT Press, Cambridge, MA.

Picard, R.W., 1999. Affective Computing for HCI. HCI International (the 8th International Conference on Human-Computer Interaction) 829-833.

Preuschoft, S., 1997. The Social Function of 'Smile' and 'Laughter': Variations Across Primate Species and Societies, in: Segerstrale, Molnar (Eds.), Nonverbal Communication: Where Nature Meets Culture.

Provine, R.R., 2000. Laughter. A Scientific Investigation. Faber & Faber, London.

QConferencing, 2011. http://www.qconferencing.eu/.

Read, J.C., 2008. Validating the Fun Toolkit: an instrument for measuring children's opinions of technology. In the Journal of Cognition, Technology & Work, 10(2), 119-128.

Ruch, W., Ekman, P., 2001. The expressive pattern of laughter, in: Kaszniak, A. (Ed.), Emotion, qualia and consciousness, pp. 426-443.

Seligman, M.E.P., Csikszentmihalyi, M., 2000. Positive psychology. American Psychologist 55(1), 5-14.

Seligman, M.E.P., Steen, T.A., Park, N., Peterson, C., 2005. Positive Psychology Progress: Empirical Validation of Interventions. American Psychologist, 60, 410-421.

Shahid, S., Krahmer, E., Swerts, M., 2008. Alone or Together: Exploring the Effect of Physical Copresence on the Emotional Expressions of Game Playing Children Across Cultures. In the International Conference of Fun and Games, Eindhoven, the Netherlands, Springer.

Shapiro, S.L., Astin, J.A., Bishop, S.R., 2005. Mindfulness-based stress reduction for health care professionals: Results from a randomized trial. International Journal of Stress Management, 12, 164-176.

Sheldon, K.M., King, L., 2001. Why positive psychology is necessary. American Psychologist 56(3), 216-217.

Sin, N.L., Lyubomirsky, S., 2009. Enhancing well-being and alleviating depressive symptoms with positive psychology interventions: A practice-friendly meta-analysis. Journal of Clinical Psychology, 65, 467–487.

Soussignan, R., 2002. Duchenne Smile, Emotional Experience, and Autonomic Reactivity: A Test of the Facial Feedback Hypothesis. Emotion, 2(1), 52-74.

Sweetser, P., Wyeth, P., 2005. GameFlow: A model for evaluating player enjoyment in games. Computers in Entertainment, 3(3).

Tkach, C., Lyubomirsky, S., 2006. How do people pursue happiness?: Relating personality, happiness-increasing strategies, and well-being. Journal of Happiness Studies, 7(2), 183-225.

Trouvain, J., 2001. Phonetic aspects of speech-laughs. Orality & Gestuality (Orage), Aix-en-Provence, 634-639.

Truong, K.P., Raaijmakers, S., 2007. Automatic discrimination between laughter and speech Speech Communication, 49(2), 144-158

Truong, K.P., van Leeuwen, D.A., Neerincx, M.A., 2007. Unobtrusive Multimodal Emotion Detection in Adaptive Interfaces: Speech and Facial Expressions, in: Schmorrow, D.D., Reeves, L.M. (Eds.), Foundations of Augmented Cognition, LNAI 4565 proceedings

Vaid, J., Ramachandran, V.S., 2001. Laughter and humor, in: Blakemore, C., Jennett, S. (Eds.), The Oxford Companion to the Body. Oxford University Press, Oxford.

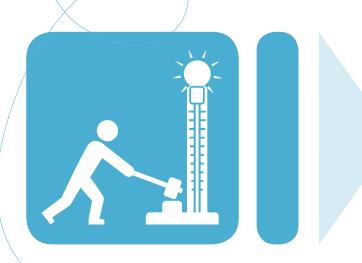
Vosburg, S.K., 1998. Mood and the quantity and quality of ideas. Creativity Research Journal 11(4), 315-324.

Wagner, H., Lee, V., 1999. Facial behavior alone and in the presence of others. In P. Philippott et al. (Eds.), The social context of nonverbal behavior, New York: Cambridge University Press.

Wagner, H.L., Smith, J., 1991. Facial expressions in the presence of friends and stranger. In the Journal of nonverbal Behaviour, 15(4).

Zhou, J., George, J.M., 2003. Awakening employee creativity: The role of leader emotional intelligence. The Leadership Quarterly, 14(4-5), 545-568.





# CHAPTER 6

General Discussion & Conclusion

In this dissertation, we explored to what extent non-verbal emotional expressions of people engaging in playful interactions depend on aspects of the social context and how variation in social settings, including differences in cultural background of players and physical co-presence of a gamepartner, influences players' emotional expressions, game experience and social interaction. We conducted four empirical studies, in each study systematically using a combination of subjective and objective measures. Each of these studies highlights one way in which the social context shapes nonverbal expressions of emotion and social interaction among players.

# Chapter 1: The effect of co-presence and culture on the facial expressions

In our first study we investigated how physical co-presence of a friend as a game partner influences the non-verbal emotional expressions of game playing children in a cross-cultural setting. We focused on 8 and 12 year old children, either from Pakistan or The Netherlands, and studied how they signaled joy (a positive emotion) and disappointment (a negative emotion) after winning or losing a specifically designed card game, while playing alone or together with their friends. This allowed us to focus on the spontaneous expression of emotions in a natural and playful settings, eliciting emotions across cultures in a consistent way.

Our hypothesis was that if the children's emotional expressions have a considerable social and cultural aspect then we would expect to find differences in expressivity not only between individual players and pairs of players but also between Pakistani players and Dutch players. The results of our study confirmed this hypothesis. The presence of a friend as a game partner clearly influenced the emotional expression of children in that they were more expressive while playing with their friends than while playing alone. Furthermore, we found that Pakistani children were overall more expressive than their Dutch counterparts confirming that the cultural background influenced the way children experience and display emotions.

The findings of our study are potentially relevant for Human-Computer Interaction in general, and Affective Computing in particular. Our results show that games offer an interactive, natural and highly efficient way to collect affective data without any artificial posing. The card game turns out to be particularly useful, because it appears to generalize to multiple settings: it works quite well in both cultures, across different age groups and in different social contexts. In the future, we plan to expand the GAME ("Game As a Method of eliciting Emotions") paradigm and we will use games to study other basic emotions, such as surprise, and more complex emotions such as shame and pride in cross-cultural settings.

Our findings also suggest that it would be good to include facial expression analyses as an evaluation metric of (computer) games. Such analyses seem to be especially useful when game-players are children, whose verbal skills are not as good as those of adults, which may be a disadvantage when one wants to evaluate a product (such as a game). Finally the affective data, collected in our study, can be used for training future affective systems, because we know from earlier research that emotion recognition systems perform best when the training data matches the real, application data.

# **Chapter 2: Child-robot Interaction**

In the second study, we investigated how children experience the presence of a social robot (the iCat) as a game partner and how they develop a social and emotional bond with this robot while playing the same collaborative game as used in Study 1. This allowed us to propose a new method to evaluate children's appreciation of social robots, by asking whether playing games with such robots is like playing alone (an unsocial experience) or more like playing with a friend (an ideal and truly social experience). In this study, we controlled for differences between children of different cultures (Pakistani and Dutch) and age groups (8 and 12 years old), since these could potentially influence the appreciation of the child-robot interaction.

Our results show that for children of both cultures 'playing with a social robot' is a social activity and is more similar to 'playing with a friend' than to 'playing alone' even though, in general, a friend is more appreciated than the iCat. The results do suggest that, despite the limitations of the iCat, many children (especially the younger ones in both cultures) associate social attributes to the iCat, imagined that the iCat is socially competent, and interacted with the iCat in a socially appropriate manner. In other words, in line with the CASA theory (Computers As Social Actors) (Reeves and Nass, 1996), children indeed treated the iCat as a social actor.

The results of our Study 2 also reinforce the importance of cultural background and age differences while designing interactive systems, confirming the findings of Study 1. Cultural background not only influenced how children behaved during child-child interactions, but also played a key role in shaping children's responses and attitudes towards the iCat. Pakistani children, for example, were generally more engaged and showed a stronger interest in playing with the iCat than Dutch children. Furthermore, Pakistani children were not only more expressive but also seemed to have a richer social interaction than their Dutch counterparts, touching and hugging iCat more often, sitting closer to the iCat and paying more attention to its feedback. Furthermore, younger children (8 years old) were found to be more expressive and considered playing with the iCat more engaging than older children (12 years old).

Although, in general, playing with a robot was clearly considered more fun than playing alone, it is fair to say that playing with a friend was even more fun. However, there is ample room for improving the child-robot interaction, and based on our results, we suggest that one way of designing a better interaction between children and robots would be to take individual differences (such as cultural background and age) explicitly into account during the design process.

It is important to note that, certainly when compared to a friend, the robot was still a "stranger" for the children. It is possible that children's first interaction with a robot is more comparable to interacting with a randomly selected, unfamiliar partner to them than interacting with self-selected partners (friends). In a follow-up experiment, we therefore plan to compare child-robot interaction with a 'playing with strangers' condition. Furthermore, in our current study, the children's interaction with the iCat was restricted to a few minutes. Previous research has shown that repeated exposure to a robot may change the children's attitude towards this robot in both a negative and a positive way. In future studies, we plan to investigate repeated interactions between robot and child, where we would like to explore how the longitudinal interaction effects the overall evaluation. 'Repeated exposures' is a theme to which we return in Study 4.

Taken together, our first two studies, are among the first to show systematic differences in expression of emotion between Dutch and Pakistani children. It would be important to replicate these findings, not only using games as a tool, but also using other emotion eliciting methods to see how whether such cultural differences are also found during other, possibly non-playful, interactions.

# Chapter 3: Mutual gaze in video-mediated gameplay

In the third study, we explored the effect of variations in gaze on the perceived social presence, emotional expressions, social interaction and game playing behavior of children during videomediated gameplay. We also investigated to what extent this video-mediated gameplay experience resembled that of a natural co-present experience (that is a situation in which children are physically co-present in pairs and interact in a face-to-face manner). Mutual gaze is known to play a key role in regulating face-to-face communication, and our hypothesis therefore was that mutual gaze would influence the players' game experience and perceived sense of presence during video-mediated communication.

Our results indeed suggest that children seem to enjoy playing a card game least in a no-gaze condition, while their appreciations for the mutual gaze and co-present condition are comparable (often with a preference for the former over the latter). Having the possibilities for mutual gaze had a direct effect on the perceived intimacy and closeness of children. Children reported that they felt the presence of their partner in the same physical space more strongly in the mutual gaze condition than in the no gaze condition. Children also had more fun, and were more engaged in the mutual gaze than in the no gaze condition, which is generally in line with Media Richness theories (Daft and Lengel, 1986; Short et al., 1976). The perception test and behavioral analysis confirmed that children were more expressive and showed richer social and affective cues while playing games when mutual gaze was possible, while having no possibilities for mutual gaze had a negative effect on expressiveness. In the physically co-present condition, children were also more engaged in game play and were rated as more expressive than children in the no gaze condition, although they were rated as most expressive in the (non-physically co-present) mutual gaze condition.

This study yielded interesting insights into video-mediated gameplay, which are potentially relevant for designing video-mediated communication systems, particularly for young children and playful settings. The children we focused on here seem particularly relevant because during this age (7-8 years) they strongly exploit mutual gaze for supporting and regulating communication. Our results indeed show that the possibility of eye-gaze is a crucial ingredient for the perceived sense of presence, enjoyment and positive game experience. Interestingly, our results also show that in the right context and settings, video-mediated communication experiences can not only match but even supersede physically co-present experiences, which is often considered an ideal form of humanhuman communication. The advantage of video-mediated communication in our experiment is possibly due to a "novelty-effect"; for the vast majority of the children this was the first time they played a game in such a video-mediated way.

In this study, we only focused on 8 year old Dutch children. In line with our first two studies it would be interesting to see how variations in level of gaze during video-mediated communication would influence the perceived social presence, emotional expressions and game playing behavior of older children and ones from other cultures, such as Pakistani. We would expect that, compared to Dutch children, the novelty effect could be stronger for Pakistani children, because of their lesser exposure to such communication systems, thereby possibly eliciting stronger emotional and social responses from the children.

Another thing to note is that, although our mutual gaze condition enabled a reasonable level of natural, mutual gaze, it was still not an ideal form of eye-gaze, where participants can look directly into each other's eyes, as is possible in face-to-face communication. Therefore, in another follow-up study, we would like to replicate Study 3 using so-called Eye Catchers. An Eye Catcher is a video-phone with a high-resolution camera installed behind the screen that is invisible to the participants sitting in front of it because of its half-silvered mirror. The camera is located right in the center of the screen, which offers true eye-gaze during video conferencing. Such a setup would be interesting for our purposes, because if children who communicate via an Eye Catcher are more expressive than those communicating via the original "mutual gaze" condition, then we can conclude that true eye-gaze is most important for mediated communication and the lack of it during our "mutual gaze" condition was perhaps partly compensated for by the novelty effect of the communication setup. However, if children's expressiveness while playing via Eye Catchers would be comparable to the "mutual gaze" condition, then we can conclude that a less ideal form of mutual gaze appears to be enough for specific tasks and contexts, and having a true, direct mutual gaze facilities does not add much to the overall communication experience.

# **Chapter 4: Affective Positive Interactions**

In the fourth and final study, we investigated if and how an affective interface (an adaptive laughing mirror) induces positive affect and laughter in individual participants and in people who participate with the physical co-presence of a friend. In addition, we also investigated whether the induced effect lasts for a longer period and whether it is repeatable after a fixed interval. The application we developed is called the Affective Mirror (AM), and what it does is distorting a participant's face in a manner which might be perceived as "funny" and which adapts itself according to the perceived user state. It was developed as a tool for inducing positive emotions and laughter in participants.

We found that, in general, the AM interface worked well for inducing positive emotions in that participants indeed reported feeling more positive after their AM session. Our results revealed that this effect is repeatable, in the sense that the AM could induce positive emotions in participants for a second time after a fixed interval, although the expressiveness of participants did decrease somewhat with repeated exposure. Furthermore, the positive effect of the AM was found to be stronger for pairs of participants than for individual one; when people participate in pairs, they report feeling happier and enjoying the interaction more, and independent judges rate them as more expressive as well.

Overall the results show that affective mirror interface can be used as a reliable tool for inducing positive emotions in a natural manner and the audio-visual data can be used for training affective systems. The results also show the importance of two key factors which can influence the performance of automatic affect recognition systems: the way human behavior changes after a repeated exposure to an affective system and the way human participants change their behavior in the

presence of others. The data collected in this study is already being used for training a commercial emotion recognition system, and it would be interesting to study how the new version of the affect recognition system works in similar setups.

In this study, similar to the study 3, we only focused on Dutch participants. Again, it would be interesting to use the Affective Mirror with Pakistani participants and see how they express positive emotions during repeated exposure and in the presence of friends. In our studies, we found that the Pakistani children were more expressive in showing positive emotions than Dutch ones. It is quite possible that we would see a similar effect for adults. In another study where we induced positive and negative emotions in Pakistani and Dutch adults (Shahid et al., 2008), we found that Pakistani adults, compared to the Dutch ones, were indeed more expressive while showing positive emotions. The Affective Mirror interface might help in replicating these results. Finally it would be interesting to run a similar study with children instead of adult participants, because it is likely that Affective Mirror, if tested with children, will induce stronger emotional responses both in the repeated exposure and physical co-presence conditions.

Taken together, these four studies clearly showed that the GAME paradigm offers a useful way to study the non-verbal expression of emotions. Each of these studies, in its own way, contributed to our understanding of factors influencing the expression of emotion, where especially the physical copresence of another game participant (such as a friend or a social robot) and the cultural background of users (e.g., Dutch or Pakistani) were shown to be important in shaping the non-verbal expression of emotion during playful interactions. So, when you visit the Tilburg Kermis (or any other fun fair) next year, hopefully you will notice that not only the attractions are worth exploring, but also the facial expressions of the visitors - real Fun is on the Face.

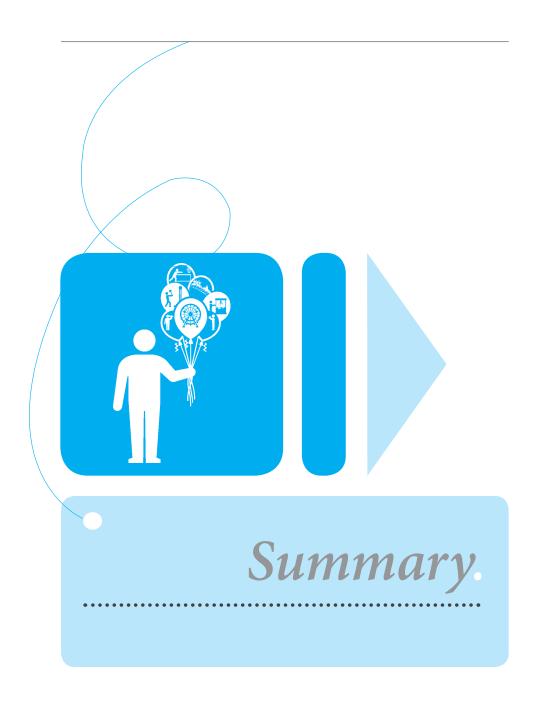
## **References:**

Daft, R., Lengel, R., 1986. Organizational Information Requirements, Media richness and Structural Design. Management Science, 32(5), 554-571.

Reeves, B., Nass, C., 1996. The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places. Cambridge University Press, New York.

Shahid, S., Krahmer, E., Swerts, M., 2008. Real vs. acted emotional speech: Comparing Caucasian and South-Asian speakers and observers. Proceedings of the Speech Prosody, Brazil, 669-672.

Short, J., Williams, E., Christie, B., 1976. The social psychology of telecommunications. Contemporary Sociology.



In this dissertation, we explored the role of social context in shaping the non-verbal expression of emotions during playful interactions. To do this, we proposed a new method for eliciting emotions in a natural manner for different social settings, which we call the GAME (Game As a Method of eliciting Emotions) paradigm. We reported on four studies, each one using a game developed under the GAME paradigm. Each of these four studies explores one way in which the social context may shape non-verbal expressions of emotion and social interaction among players. Moreover, we systematically explored different methods for measuring emotional responses and social interactions among players, ranging from self-reports and interviews to objective measures and behavioral analyses.

In our first study, described in chapter 2, we investigated the influence of physical co-presence of a friend as a game partner on the emotional expressions of game playing children across cultures. We explored how children of different age groups (8 and 12 years old) and with different cultural backgrounds (Dutch and Pakistani) signal positive and negative emotions while playing a game either alone or together. A simple number guessing game was developed and used as a tool for inducing emotions, which was played by Pakistani and Dutch individual and pairs of children. The data collected was used in a series of cross-cultural perception studies, in which Dutch and Pakistani observers classified the emotional expressions of the Dutch and Pakistani children. If emotional expressions primarily serve to reflect the individual and are not influenced by social context (culture and co-presence), then the emotional responses of children will be highly comparable for both cultures while playing alone or with a friend. However, if the expressions also have a considerable social component and are rooted in culture, we should find differences in expressiveness across cultures and between individual players and pairs of players.

The results of study 1 showed that computer games offer a natural and highly efficient way to collect affective data. The card game we developed was a straightforward implementation of an extremely simple game, certainly when compared to games currently on the market. Still, our results showed that children responded strongly, and with a variety of emotional behaviours, after winning or losing this game. The results of our study showed that the presence of a friend as a game partner clearly influenced the emotional expression of children in that they were more expressive while playing with their friends than while playing alone. We also found that Pakistani children were overall more expressive than their Dutch counterparts confirming that the cultural background influenced the way children display emotions. Finally, we found that younger (8 year old) children were more expressive than older ones (12 year old). These findings are relevant for Human-Computer Interaction in general, and Affective Computing in particular. We know from earlier research that emotion recognition systems perform best when the training data matches the real, application data, so it seems important to take factors such as age, culture and co-presence into account when designing affective applications.

In the second study, described in chapter 3, we studied the influence of the presence of a social robot (the iCat) as a game partner on the social and emotional response of children across cultures. We investigated how children of different age groups (8 and 12 year olds) and belonging to different cultural backgrounds (Pakistani vs. Dutch) experience interacting with a social robot (iCat) during collaborative game play. We employ the same collaborative card game and general set-up as used in Study 1 for eliciting emotions and social interaction. This allows us to effectively propose a new

method to evaluate children's appreciation of social robots, by asking whether playing a game with a social robot is more similar to playing this game alone or with a friend. Ideally, playing with a social robot will be experienced more like playing with a friend (an ideal and truly social experience) than playing alone (an unsocial experience). Our reasoning was that if children find playing with a social robot comparable to playing with a friend, then we would expect similar socio-emotional responses during child-child and child-robot interactions. Similar to the first study, if cultural background shapes the way children emotionally interact with a social robot then we would also expect to find differences in expressiveness and social interaction across cultures.

A combination of self-report scores, perception test results and behavioral analyses indicated that child-robot interaction in game playing situations is highly appreciated by children, although more by Pakistani and younger children than by Dutch and older children. Results also suggest that children had more fun playing with the robot than playing alone, but had more fun still when playing with a friend. Similarly, children were more expressive when playing with the robot than they were when playing alone, but less expressive than when playing with a friend.

The second study highlighted some important, but previously understudied aspects of childrobot interaction in an authentic social setting. Our results not only stress the importance of using new benchmarks for evaluating child-robot interaction but also reinforce the importance of cultural background and age differences while designing interactive systems, confirming the findings of Study 1. Cultural background not only influenced how children behaved during child-child interactions, but also played a key role in shaping children's responses and attitudes towards the iCat. Based on these results, we propose that one way of designing a better interaction between children and robots is to take individual differences such as cultural background and age explicitly into account during the design process. Furthermore, we argue that cross-cultural robot design should go beyond mere translations and high-level localizations. The underlying social norms of a particular culture should be part of the robot design process.

In our third study, described in chapter 3, we extended the concept of study 1 and studied how pairs of children interact socially and express their emotions while playing games in a video mediated communication environment. We not only investigated the influence of the presence of a remote friend as a game partner on the game experience children but also studied how variations in gaze (mutual gaze condition vs. no gaze condition) during video-mediated gameplay influence the perceived social presence, emotional expressions, social interaction and game playing behavior of children. We employed the same collaborative card game and general set-up as used in Study 1 for eliciting emotions and social interactions. This helped us to study to what extent the video-mediated gameplay experience resembles that of the natural co-present experience (as discussed in chapter 1).

Mutual gaze plays a key role in regulating human-human communication, suggesting that the absence of mutual gaze during mediated communication may influence players' game experience and perceived sense of presence in a technologically mediated environment. Furthermore, if the copresent (face-to-face) game playing experience is considered an ideal form of communication then we would expect differences in the way people express emotions and exchange social cues during video-mediated gameplay and co-present gameplay.

We used a combination of self-reports, objective facial expression analysis and behavioral analyses for measuring the perceived social presence, emotional expressions and game experience. Our results showed that eye-gaze is a crucial ingredient for the perceived sense of presence, enjoyment and positive game experience; conversely, the absence of mutual eye-gaze dramatically influences the quality of interaction in the video-mediated environment. Our results indeed suggest that children seem to enjoy playing a card game least in a no-gaze condition, while their appreciations for the mutual gaze and co-present condition are comparable (often with a preference for the former over the latter). Having the possibilities for mutual gaze had a direct effect on the perceived intimacy and closeness of children. Children reported that they felt the presence of their partner in the same physical space more strongly in the mutual gaze condition than in the no gaze condition.

The results of this study stress the importance of mutual gaze, and we therefore argue that it should become an integral component of future video mediated communication systems, particularly in those designed for playful settings and children. Interestingly, our results also show that in a playful context and with the right settings, video-mediated communication experiences can not only match, but even supersede physically co-present experiences, which are often considered the ideal form of human-human communication.

In our fourth and final study, described in chapter 5, we presented a new affective interface (a digital laughing mirror), which distorts a user's face in a manner which might be perceived as funny, and which adapts itself according to the perceived user's state, and we evaluated if and how this affective interface can be used as a tool for inducing positive emotions. In this study, consistent with the first three studies, we explored the influence of physical co-presence of a friend on the induction and expression of positive emotions and laughter. If emotion induction and positive expressions are influenced by the social context, we would again expect a stronger response of participants in the presence of their friends. In this study, we also focused on one additional factor, namely "repeated exposure" and investigate whether the induced positive affect is repeatable after a fixed interval. If the laughing mirror is strong enough in inducing laughter repeatedly then we would expect comparable emotional responses (both felt and expressed) after a first and a second session.

Consistent with study 2 and 3, we used a combination of self-reports, objective facial expression analysis and behavioral analyses for measuring the emotional expressions, game experience and mood of participants. Results show that the AM interface worked well for inducing positive emotions and participants systematically felt more positive after an affective mirror session. Our results revealed that this effect is repeatable, in the sense that the AM could induce positive emotions in participants for a second time after a fixed interval, although the expressiveness of participants did decrease somewhat with repeated exposure. Moreover, the positive effect of the AM was found to be stronger for pairs of participants than for individual ones; when people participate in pairs, they report feeling happier and enjoying the interaction more, and independent judges rate them as more expressive as well.

Overall the results show that the affective mirror interface can be used as a reliable tool for inducing positive emotions in a natural manner and the audio-visual data collected in this way can be used for training affective systems. The results also show the importance of two key factors which can influence the performance of automatic affect recognition systems: the way human behavior

changes after a repeated exposure to an affective system and the way human participants change their behavior in the presence of others.

To conclude, the main objective of this dissertation was to study the effectives of the GAME paradigm as a new method for eliciting emotions in different social contexts and to explore how systematic variations in the social context shape non-verbal expressions of emotion and social interactions among players. The four studies described in this thesis, showed that the GAME paradigm offers a useful way to study the non-verbal expression of emotions. Each of these studies, in its own way, contributed to our understanding of factors influencing the expression of emotion, where especially the physical presence of another game participant (such as a friend or a social robot) and the cultural background of users (e.g., Dutch or Pakistani) were shown to be important in shaping the non-verbal expression of emotion during playful interactions.



## Journal papers/book chapters:

Shahid, S., Krahmer, E., Swerts, M. (In Press). Video-mediated and Co-present Gameplay: Effects of Mutual Gaze on Game Experience, Expressiveness and Perceived Social Presence. Interacting with Computers (Elsevier).

Al Mahmud, A., Mubin, O., Shahid, S. (In Press). Designing with and for Older Adults: Experience from Game Design. Book chapter in Human-Computer Interaction: The Agency Perspective, Zacarias, Marielba; Oliveira, José Valente de (Eds.) (Springer).

Ahmad, M., I., Hassan, T., Saeed, S., Shahid, S., Krahmer, E. (2011). Guess who? an interactive and entertaining game-like platform for investigating human emotions. In the Proceedings of the 14th International Conference on Human Computer Interaction (HCI 2011), Springer-Verlag (LNCS series), Orlando, Florida.

Mattheij, R., Szilvasi, L., de Beer, L., Rakiman, K., Shahid, S. (2011). GooGreen: A Personalizing Green Advice. In the Proceedings of the 14th International Conference on Human Computer Interaction (HCI 2011), Springer-Verlag (LNCS series), Orlando, Florida.

Al Mahmud, A., Mubin, O., Shahid, S., & Martens, J.B. (2010). Designing Social Games for Children and Older Adults: Two Related Case Studies. The Journal of Entertainment Computing, 2(3), Elsevier.

Shahid, S., Krahmer, E., Swerts, M. (2008). Alone or Together: Exploring the effect of physical copresence on the emotional expressions of game playing children across cultures. In the Springer Proceedings of the International Conference on Fun and Games (FnG 2008), Springer-Verlag (LNCS series), Eindhoven, the Netherlands. - Best Paper Award

## Working journal papers:

Shahid, S., Krahmer, E., Neerincx, M., Swerts, M. Affective Positive Interactions: Role of Repeated Exposure and Social Presence. Submitted.

Shahid, S., Krahmer, E., Swerts, M. The Effect of Co-presence and Culture on the Emotional Expressions of Game Playing Children. Submitted.

Shahid, S., Krahmer, E., Swerts, M. Child-robot Interaction: How does playing a game with a social robot compare to playing a game alone or with a friend? Submitted.

# Conference/workshop proceedings (peer reviewed):

Shahid, S., Krahmer, E., Swerts, M. (2011). Child-robot Interaction: Playing Alone or Together? In the Proceedings of the ACM Conference on Human Factors in Computing Systems, (CHI EA) Atlanta, USA.

Shahid, S., Krahmer, E., Swerts, M. (2011). Who is More Expressive During Child-Robot Interaction: Pakistani or Dutch Children? In the IEEE/ACM Proceedings of the International Conference on Human Robot Interaction (HRI 2011), Lausanne, Switzerland.

De Kruyff, P., Steentjes, A, Shahid, S. (2011). The Alkaline Arcade: A Child-Friendly Fun Machine for Battery Recycling. In the ACM Proceedings of the International Conference on Advances in Computer Entertainment Technology (ACE 2011), Lisbon, Portugal.

Shahid, S., Krahmer, E.J., Swerts, M.J.G., and Mubin, O. (2010). Child-Robot Interaction during Collaborative Game Play: Effects of Age and Gender on Emotion and Experience. In the ACM Proceedings of the OzCHI 2010 Conference, Brisbane, Australia.

Shahid, S., Krahmer, E., Swerts, M. (2010). Playing with iCat: Investigating Children's Appreciation of Game Plays with a Social Robot. In the ACM Proceedings of the International Conference on Advances in Computer Entertainment Technology (ACE 2010), Taipei, Taiwan.

Shahid, S., Krahmer, E., Swerts, M. (2010). GamE Paradigm: Affective Gaming for Affect Elicitation. Creative Showcase Demo - In the ACM Proceedings of the International Conference on Advances in Computer Entertainment Technology (ACE 2010), Taipei, Taiwan.

Van de Wiel, E., Meijering, M., Shahid, S. (2010). CookKing: A King of Healthy, Fun and Social Cooking. In the ACM Proceedings of the International Conference on Advances in Computer Entertainment Technology (ACE 2010), Taipei, Taiwan.

Shahid, S., Krahmer, E.J., Swerts, M.J.G. (2010). GamE in action: Using the GamE paradigm as a tool for investigating human emotions. In the proceedings of the workshop 'Video Games as Research Tools' in CHI 2010. Atlanta, USA.

Hong, A., van Heugten, S., Kooken, T., Vinke, N., Vromans, M., and Shahid, S. (2010). The Augenda: Structuring the Lives of Autistic Teenagers. In the ACM Proceedings o the 12<sup>th</sup> International ACM SIGACCESS conference on Computers and Accesibility (ASSETS 2010), Orlando, Florida.

Mubin, O., Shahid, S., van de Sande, E., Krahmer, E.J., Swerts, M.J.G., Bartneck, C. and Feijs, L.M.G (2010). Using Child-Robot Interaction to Investigate the User Acceptance of Constrained and Artificial Languages. In the Proceedings of the 19th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN 2010, Viareggio, Italy.

De Beer, R., Hol, S., Keijers, R., Shahid, S., Al Mahmud, A., Mubin, O. (2010). PMD: Designing a Portable Medicine Dispenser for Persons Suffering from Alzheimer 's Disease. In the 12th International Conference on Computers Helping People with Special Needs (ICCHP 2010), Vienna, Austria, Springer-Verlag (LNCS series).

Blommaert, A., Philippart, P., Rassaert, C., Theunissen, E., Widdershoven, S., and Shahid, S. (2010). Day Pad: A Daily life Assistant for Autistic Children. Future Innovation - In the ACM Proceedings of the 12th Annual International ACM Conference on HCI with Mobile Devices and Services (Mobile HCI 2010), Lisbon, Portugal.

Shahid, S., Krahmer, E.J., Swerts, M.J.G., Melder, W., & Neerincx, M. (2009). You Make Me Happy: Using an Adaptive Affective Interface to Investigate the Effect of Social Presence on Positive Emotion

Induction. In the Proceedings of the IEEE 3<sup>rd</sup> International Conference on Affective Computing and Intelligent Interaction, ACII'09. Amsterdam, the Netherlands

Shahid, S., Krahmer, E.J., Swerts, M.J.G., Melder, W., & Neerincx, M. (2009). Exploring the Social and Temporal dimension of Emotion Induction using an adaptive Affective Mirror. In the Extended Abstracts of the CHI '09 Conference on Human Factors in Computing Systems. (ACM, New York, NY). Boston, USA.

Mubin, O., Shahid, S., Bartneck, C., Krahmer, E.J., Swerts, M.J.G., & Feijs, L.M.G. (2009). Using Language Tests and Emotional Expressions to Determine the Learnability of Artificial Languages. Work in Progress Paper To Appear in the Extended Abstracts of the CHI '09 Conference on Human Factors in Computing Systems. (ACM, New York, NY). Boston, USA.

Shahid, S., Krahmer, E. (2009). Human-centred Design in the Developing World: Towards designing didactic games for children. In the proceedings of the 'Human-centered Computing in International Development' workshop. In the CHI '09 Conference on Human Factors in Computing Systems. Boston, USA.

Shahid, S., (2009). Exploring the Emergency Response Workers' Interaction with Large Displays during Interventions. In the 'Crowd-Computer Interaction' workshop. In the CHI '09 Conference on Human Factors in Computing Systems. Boston, USA.

Shahid, S., Mubin, O., & Al Mahmud, A. (2009). RACE: Towards Exploring the Design Dimensions of a Route Assisting and Communication System for the Elderly. In the Proceedings of the HCI 2009 International Conference (Series: Springer LNCS). San Diego, USA

Al Mahmud, Mubin, O., & Shahid,S. (2009). User Experience with in-car GPS Navigation Systems: Comparing the Young and Elderly Drivers. In the ACM Proceedings of the 11th Annual International ACM Conference on HCI with Mobile Devices and Services (Mobile HCI), Bonn, Germany

Al Mahmud, A., Mubin, O., Shahid, S., & Martens, J.B. (2008). Designing and Evaluating the Tabletop Game Experience for Senior Citizens. In the ACM Proceedings of the NordiCHI 2008 conference, Lund, Sweden.

Shahid, S., Mubin, O., & Al Mahmud, A. (2008). Walk 2 Win: Towards Designing a Mobile Game for Elderly's Social Engagement. In the ACM Proceedings of the 22nd Annual British HCI Conference (British HCI 2008), Liverpool, England

Al Mahmud, A., Mubin, O., Shahid, S., Juola, J., & de-Ruyter, B. (2008). EZ Phone: Persuading Mobile Users to Conserve Energy. In the ACM Proceedings of the 22nd Annual British HCI Conference (British HCI 2008), Liverpool, England

Shahid, S., Krahmer, E., Swerts, M. (2008). GamE: Games as a Method for Eliciting Emotions. In the Proceedings of 6th International Conference on Methods and Techniques in Behavioral Research (Measuring Behavior 2008), Maastricht, the Netherlands.

Shahid, S., Mubin, O., & Al Mahmud, A. (2008). Friend Forever - A Healderly Partner. To Appear in the Proceedings of the 10th Annual International ACM Conference on HCI with Mobile Devices and Services (Mobile HCI), Amsterdam, Netherlands.

Shahid, S., Krahmer, E., Swerts, M. (2008). Real vs. Acted Emotional Speech: Comparing Caucasian and South-Asian Speakers and Observers. In the Proceedings of the 4th International Conference on Speech Prosody (SP 2008), Campinas, Brazil

Shahid, S., Krahmer, E., Swerts, M. (2007). Audiovisual Emotional Speech of 8 and 12 Years Old Children: A Cross-cultural Comparison. In the Proceedings of Eighth Conference in the Annual Series of Interspeech (Interspeech 2007) Antwerp, Belgium.

Shahid, S., Moran, O., Li, Y., Jia, D. (2007). Photo Shaker: A novel interaction technique for browsing photos on handheld devices. A section in the book 'Information Visualization: Design for Interaction' (2nd Edition) by Robert Spence, Prentice Hall, ISBN: 978-0132065504.

Al Mahmud, A., Mubin, O., Octavia, J.R., Shahid, S., Yeo, L.C., Markopoulos, P., Martens, J.B. & Aliakseyeu, D. (2007). Affective Tabletop Game: A New Gaming Experience for Children. Proceedings of the 2nd Annual IEEE Workshop on Horizontal Interactive Human-Computer Systems (Tabletop 2007), Newport, Rhode Island, USA.

Al Mahmud, A., Mubin, O., Octavia, J.R., Shahid, S., Yeo, L.C., Markopoulos, P., & Martens, J.B. (2007). aMAZEd: Designing an Affective Social Game for Children. Proceedings of the 6th Annual ACM Conference on Interaction Design and Children (IDC 2007), Aalborg, Denmark.

Al Mahmud, A., Dadlani, P., Mubin, O., Shahid, S., & Midden, C. (2007). iParrot: Towards Designing a Persuasive Agent for Energy Conservation. Proceedings of the 2nd Annual Conference on Persuasive Technology (PERSUASIVE 2007), Stanford, USA (Series: Springer - LNCS).

Hashmi N., Al-Malaise, A., Shahid, S., Afzal, M., Z. (2006). Managing Workflows in Data Environments using Se-mantic Technologies. In the IEEE International Conference on Intelligent Computing, YunNan, China.

## Abstracts - Oral & poster presentations (peer reviewed):

Shahid, S., Krahmer, E., Swerts, M. (2011). Using Facial Expressions and Social Presence Self-reports for Investigating the Role of Eye-gaze in Computer-mediated Gameplay. Oral Presentation - In the proceedings of the Conference of the International Society for Research on Emotions, ISRE'11. Kyoto, Japan.

Shahid, S., Krahmer, E., Swerts, M. (2010). Robot is also a Friend? Investigating how differences in social presence effect the emotional expressions of children across cultures. Paper presented as a 'Hot Topic Presentation' in the Third European Conference on Emotion (CERE 2010), Lille, France

Elshout, R. and Shahid, S. (2010). Floorbook: An Interactive Floor Plan for Enhancing Social Interaction among Office Workers. In the Proceedings of the 14th CHI Nederland Conference (The Web and Beyond), Amsterdam, the Netherlands.

Shahid, S., Sande, E. van de, Mubin, O., Krahmer, E., Swerts, M. (2009). Child-Robot Interaction: Investigating the Emotional Expressions and Social Interaction of Children in a Collaborative Game. In the proceedings of the 2<sup>nd</sup> International Conference on Human-Robot Personal Relationship, HRP'09. Leiden, the Netherlands.

Shahid, S., Krahmer, E., Swerts, M. (2009). Comparing facial expressions and emotion self-reports of game- playing children. Oral Presentation - In the proceedings of the Conference of the International Society for Research on Emotions, ISRE'09. Leuven, Belgium

Shahid, S., Krahmer, E., Swerts, M. (2009). Using subjective and objective emotional response for measuring social presence and game experience of children. In the HCI 2009 International Conference. San Diego, USA

Mubin, O., Al Mahmud, A., & Shahid, S. (2009). Experiencing Location Aware Systems: a Case Study. In the HCI 2009 International Conference. San Diego, USA

Valkenburg, L., de Pinth, P., Vollering, T., Shahid, S. (2009). Buddy Bo: An Educational Tool to Enlarge the Vocabulary of Young Children. In the Proceedings of the 13th CHI Nederland Conference, Leiden, the Netherlands.

Shahid, S., Krahmer, E. (2008). Cross-Cultural Participatory Design: A Case Of Designing an Educational Game For Illiterate Children. In the 'The Challenges for Participatory Design in the Developing World' Workshop. In the Participatory Design (PD2008) conference, Bloomington, Indiana, USA.

Shahid, S., Krahmer, E., Swerts, M. (2008). Using games as a tool for exploring cross-culture emotional differences in children. Extended Abstract in the 2<sup>nd</sup> International Symposium in Cross-Cultural Psychology, Tilburg, the Netherlands.

Shahid, S., Mubin, O., & Al Mahmud, A. (2008). Exploring the Design Dimensions of a Mobile Guidance System for Senior Citizens. In the Proceedings of the 12th CHI Nederland Conference (The Web and Beyond), Amsterdam, the Netherlands.

Shahid, S., Krahmer, E., Swerts, M. (2007). Comparing Acted and Real Expressions of Emotion: An Experimental Investigation. Appeared as an Extended Abstract in the Emotions 2007 Conference, Tilburg, the Netherlands, 2007.

Klann, M., Shahid, S. (2005). Playing with Fire. In the Proceedings of ACM ECSCW Workshop on Computer Games and CSCW, Paris, France.



- 1. Pashiera Barkhuysen. Audiovisual Prosody in Interaction. Promotores: M.G.J. Swerts, E.J. Krahmer. Tilburg, 3 October 2008.
- Ben Torben-Nielsen. Dendritic morphology: function shapes structure. Promotores: H.J. van den Herik, E.O. Postma. Co-promotor: K.P. Tuyls. Tilburg, 3 December 2008.
- 3. Hans Stol. A framework for evidence-based policy making using IT. Promotor: H.J. van den Herik. Tilburg, 21 January 2009.
- Jeroen Geertzen. Dialogue act recognition and prediction. Promotor: H. Bunt. Co-promotor: 4. J.M.B. Terken. Tilburg, 11 February 2009.
- Sander Canisius. Structured prediction for natural language processing. Promotores: A.P.J. 5. van den Bosch, W. Daelemans. Tilburg, 13 February 2009.
- Fritz Reul. New Architectures in Computer Chess. Promotor: H.J. van den Herik. Co-6. promotor: J.W.H.M. Uiterwijk. Tilburg, 17 June 2009.
- 7. Laurens van der Maaten. Feature Extraction from Visual Data. Promotores: E.O. Postma, H.J. van den Herik. Co-promotor: A.G. Lange. Tilburg, 23 June 2009 (cum laude).
- 8. Stephan Raaijmakers. Multinomial Language Learning. Promotores: W. Daelemans, A.P.J. van den Bosch. Tilburg, 1 December 2009.
- Igor Berezhnov. Digital Analysis of Paintings. Promotores: E.O. Postma, H.J. van den Herik. 9. Tilburg, 7 December 2009.
- 10. Toine Bogers. Recommender Systems for Social Bookmarking. Promotor: A.P.J. van den Bosch. Tilburg, 8 December 2009.
- Sander Bakkes. Rapid Adaptation of Video Game AI. Promotor: H.J. van den Herik. Copromotor: P. Spronck. Tilburg, 3 March 2010.
- 12. Maria Mos. Complex Lexical Items. Promotor: A.P.J. van den Bosch. Co-promotores: A. Vermeer, A. Backus. Tilburg, 12 May 2010 (in collaboration with the Department of Language and Culture Studies).
- Marieke van Erp: Accessing Natural History. Discoveries in data cleaning, structuring, and retrieval. Promotor: A.P.J. van den Bosch. Tilburg, 30 June 2010.

- 14. Edwin Commandeur: Implicit Causality and Implicit Consequentiality in Language Comprehension. Promotores: L.G.M. Noordman, W. Vonk. Co-promotor: R. Cozijn. Tilburg, 30 June 2010.
- Bart Bogaert: Cloud Content Contention. Promotores: H.J. van den Herik, E.O. Postma. Tilburg, 30 March 2011.
- 16. Xiaoyu Mao: Airport under Control. Promotores: H.J. van den Herik, E.O. Postma. Copromotores: N. Roos, A. Salden. Tilburg, 25 May 2011.
- Olga Petukhova: Multidimensional Dialogue Modelling. Promotor: H. Bunt. Tilburg, 1 September 2011.
- Lisette Mol: Language in the Hands. Promotores: F. Maes, E.J. Krahmer, M.G.J. Swerts. Tilburg, 7 November 2011 (cum laude).
- 19. Herman Stehouwer: Statistical Language Models for Alternative Sequence Selection. Promotores: A.P.J. van den Bosch, H.J. van den Herik. Co-Promotor: M.M. van Zaanen. Tilburg, 7 December 2011.
- Terry Kakeeto-Aelen: Relationship Marketing for SMEs in Uganda. Promotores: J. Chr. van Dalen, H.J. van den Herik. Co-promotor: Dr. B. van der Walle. Tilburg, 1 February 2012.
- Suleman Shahid: Fun & Face: Exploring non-verbal expressions of emotion during playful interactions. Promotores: E.J. Krahmer, M.G.J. Swerts. Tilburg, 25 May 2012.



- 1. Johan van den Akker (CWI<sup>2</sup>) DEGAS An Active, Temporal Database of Autonomous Objects
- 2. Floris Wiesman (UM) Information Retrieval by Graphically Browsing Meta-Information
- 3. Ans Steuten (TUD) A Contribution to the Linguistic Analysis of Business Conversations within the Language/Action
- 4. Dennis Breuker (UM) Memory versus Search in Games
- 5. Eduard W. Oskamp (RUL) Computerondersteuning bij Straftoemeting

#### 1999

- 1. Mark Sloof (VU) Physiology of Quality Change Modelling; Automated modelling of Quality Change of Agricultural Products
- 2. Rob Potharst (EUR) Classification using decision trees and neural nets
- 3. Don Beal (UM) The Nature of Minimax Search
- 4. Jacques Penders (UM) The Practical Art of Moving Physical Objects
- Aldo de Moor (KUB) Empowering Communities: A Method for the Legitimate User-Driven 5. Specification of Network Information Systems
- 6. Niek J.E. Wijngaards (VU) Re-design of Compositional Systems
- 7. David Spelt (UT) Verification Support for Object Database Design
- 8. Jacques H.J. Lenting (UM) Informed Gambling: Conception and Analysis of a Multi-Agent Mechanism for Discrete Reallocation

- 1. Frank Niessink (VU) Perspectives on Improving Software Maintenance
- 2. Koen Holtman (TU/e) Prototyping of CMS Storage Management
- 3. Carolien M.T. Metselaar (UVA) Sociaal-organisatorische gevolgen van kennistechnologie; een procesbenadering en actorperspectief
- 4. Geert de Haan (VU) ETAG, A Formal Model of Competence Knowledge for User Interface Design
- 5. Ruud van der Pol (UM) Knowledge-based Query Formulation in Information Retrieval
- 6. Rogier van Eijk (UU) Programming Languages for Agent Communication
- 7. Niels Peek (UU) Decision-theoretic Planning of Clinical Patient Management
- 8. Veerle Coupé (EUR) Sensitivity Analyis of Decision-Theoretic Networks
- Florian Waas (CWI) Principles of Probabilistic Query Optimization
- 10. Niels Nes (CWI) Image Database Management System Design Considerations, Algorithms and Architecture
- 11. Jonas Karlsson (CWI) Scalable Distributed Data Structures for Database Management

<sup>&</sup>lt;sup>2</sup> Abbreviations: SIKS - Dutch Research School for Information and Knowledge Systems; CWI -Centrum voor Wiskunde en Informatica, Amsterdam; EUR - Erasmus Universiteit, Rotterdam; KUB - Katholieke Universiteit Brabant, Tilburg; KUN - Katholieke Universiteit Nijmegen; OU - Open Universiteit; RUL - Rijksuniversiteit Leiden; RUN - Radbourd Universiteit Nijmegen; TUD -Technische Universiteit Delft; TU/e - Technische Universiteit Eindhoven; UL - Universiteit Leiden; UM - Universiteit Maastricht; UT - Universiteit Twente, Enschede; UU - Universiteit Utrecht; UvA - Universiteit van Amsterdam; UvT - Universiteit van Tilburg; VU - Vrije Universiteit, Amsterdam.

- 1. Silja Renooij (UU) Qualitative Approaches to Quantifying Probabilistic Networks
- 2. Koen Hindriks (UU) Agent Programming Languages: Programming with Mental Models
- 3. Maarten van Someren (UvA) Learning as problem solving
- 4. Evgueni Smirnov (UM) Conjunctive and Disjunctive Version Spaces with Instance-Based Boundary Sets
- 5. Jacco van Ossenbruggen (VU) Processing Structured Hypermedia: A Matter of Style
- 6. Martijn van Welie (VU) Task-based User Interface Design
- 7. Bastiaan Schonhage (VU) Diva: Architectural Perspectives on Information Visualization
- 8. Pascal van Eck (VU) A Compositional Semantic Structure for Multi-Agent Systems Dynamics
- Pieter Jan 't Hoen (RUL) Towards Distributed Development of Large Object-Oriented Models, Views of Packages as Classes
- Maarten Sierhuis (UvA) Modeling and Simulating Work Practice BRAHMS: a multiagent modeling and simulation language for work practice analysis and design
- 11. Tom M. van Engers (VUA) Knowledge Management: The Role of Mental Models in Business Systems Design

- 1. Nico Lassing (VU) Architecture-Level Modifiability Analysis
- 2. Roelof van Zwol (UT) Modelling and searching web-based document collections
- 3. Henk Ernst Blok (UT) Database Optimization Aspects for Information Retrieval
- 4. Juan Roberto Castelo Valdueza (UU) The Discrete Acyclic Digraph Markov Model in Data Mining
- Radu Serban (VU) The Private Cyberspace Modeling Electronic Environments inhabited by Privacyconcerned Agents
- Laurens Mommers (UL) Applied legal epistemology; Building a knowledge-based ontology of the legal domain
- 7. Peter Boncz (CWI) Monet: A Next-Generation DBMS Kernel For Query-Intensive Applications
- 8. Jaap Gordijn (VU) Value Based Requirements Engineering: Exploring Innovative E-Commerce Ideas
- Willem-Jan van den Heuvel (KUB) Integrating Modern Business Applications with Objectified Legacy Systems
- 10. Brian Sheppard (UM) Towards Perfect Play of Scrabble
- 11. Wouter C.A. Wijngaards (VU) Agent Based Modelling of Dynamics: Biological and Organisational Applications
- 12. Albrecht Schmidt (Uva) Processing XML in Database Systems
- 13. Hongjing Wu (TU/e) A Reference Architecture for Adaptive Hypermedia Applications
- Wieke de Vries (UU) Agent Interaction: Abstract Approaches to Modelling, Programming and Verifying Multi-Agent Systems
- 15. Rik Eshuis (UT) Semantics and Verification of UML Activity Diagrams for Workflow Modelling
- 16. Pieter van Langen (VU) The Anatomy of Design: Foundations, Models and Applications
- 17. Stefan Manegold (UVA) Understanding, Modeling, and Improving Main-Memory Database Performance

- 1. Heiner Stuckenschmidt (VU) Ontology-Based Information Sharing in Weakly Structured Environments
- 2. Jan Broersen (VU) Modal Action Logics for Reasoning About Reactive Systems
- Martijn Schuemie (TUD) Human-Computer Interaction and Presence in Virtual Reality Exposure 3. Therapy
- 4. Milan Petkovic (UT) Content-Based Video Retrieval Supported by Database Technology
- 5. Jos Lehmann (UVA) Causation in Artificial Intelligence and Law - A modelling approach
- 6. Boris van Schooten (UT) Development and specification of virtual environments
- 7. Machiel Jansen (UvA) Formal Explorations of Knowledge Intensive Tasks
- 8. Yongping Ran (UM) Repair Based Scheduling
- 9. Rens Kortmann (UM) The resolution of visually guided behaviour
- 10. Andreas Lincke (UvT) Electronic Business Negotiation: Some experimental studies on the interaction between medium, innovation context and culture
- 11. Simon Keizer (UT) Reasoning under Uncertainty in Natural Language Dialogue using Bayesian Networks
- 12. Roeland Ordelman (UT) Dutch speech recognition in multimedia information retrieval
- 13. Jeroen Donkers (UM) Nosce Hostem Searching with Opponent Models
- 14. Stijn Hoppenbrouwers (KUN) Freezing Language: Conceptualisation Processes across ICT-Supported Organisations
- 15. Mathijs de Weerdt (TUD) Plan Merging in Multi-Agent Systems
- 16. Menzo Windhouwer (CWI) Feature Grammar Systems Incremental Maintenance of Indexes to Digital Media Warehouses
- 17. David Jansen (UT) Extensions of Statecharts with Probability, Time, and Stochastic Timing
- 18. Levente Kocsis (UM) Learning Search Decisions

- 1. Virginia Dignum (UU) A Model for Organizational Interaction: Based on Agents, Founded in Logic
- 2. Lai Xu (UvT) Monitoring Multi-party Contracts for E-business
- 3. Perry Groot (VU) A Theoretical and Empirical Analysis of Approximation in Symbolic Problem Solving
- 4. Chris van Aart (UVA) Organizational Principles for Multi-Agent Architectures
- 5. Viara Popova (EUR) Knowledge discovery and monotonicity
- 6. Bart-Jan Hommes (TUD) The Evaluation of Business Process Modeling Techniques
- 7. Elise Boltjes (UM) Voorbeeldig onderwijs; voorbeeldgestuurd onderwijs, een opstap naar abstract denken, vooral voor meisjes
- Joop Verbeek (UM) Politie en de Nieuwe Internationale Informatiemarkt, Grensregionale politiële 8. gegevensuitwisseling en digitale expertise
- Martin Caminada (VU) For the Sake of the Argument; explorations into argument-based reasoning
- 10. Suzanne Kabel (UVA) Knowledgerich indexing of learning-objects
- 11. Michel Klein (VU) Change Management for Distributed Ontologies
- 12. The Duy Bui (UT) Creating emotions and facial expressions for embodied agents
- 13. Wojciech Jamroga (UT) Using Multiple Models of Reality: On Agents who Know how to Play
- 14. Paul Harrenstein (UU) Logic in Conflict. Logical Explorations in Strategic Equilibrium
- 15. Arno Knobbe (UU) Multi-Relational Data Mining

- 16. Federico Divina (VU) Hybrid Genetic Relational Search for Inductive Learning
- 17. Mark Winands (UM) Informed Search in Complex Games
- 18. Vania Bessa Machado (UvA) Supporting the Construction of Qualitative Knowledge Models
- 19. Thijs Westerveld (UT) Using generative probabilistic models for multimedia retrieval
- 20. Madelon Evers (Nyenrode) Learning from Design: facilitating multidisciplinary design teams

- 1. Floor Verdenius (UVA) Methodological Aspects of Designing Induction Based Applications
- 2. Erik van der Werf (UM) AI techniques for the game of Go
- 3. Franc Grootjen (RUN) A Pragmatic Approach to the Conceptualisation of Language
- 4. Nirvana Meratnia (UT) Towards Database Support for Moving Object data
- 5. Gabriel Infante-Lopez (UVA) Two-Level Probabilistic Grammars for Natural Language Parsing
- 6. Pieter Spronck (UM) Adaptive Game AI
- 7. Flavius Frasincar (TU/e) Hypermedia Presentation Generation for Semantic Web Information Systems
- Richard Vdovjak (TU/e) A Model-driven Approach for Building Distributed Ontology-based Web Applications
- 9. Jeen Broekstra (VU) Storage, Querying and Inferencing for Semantic Web Languages
- Anders Bouwer (UVA) Explaining Behaviour: Using Qualitative Simulation in Interactive Learning Environments
- 11. Elth Ogston (VU) Agent Based Matchmaking and Clustering A Decentralized Approach to Search
- 12. Csaba Boer (EUR) Distributed Simulation in Industry
- 13. Fred Hamburg (UL) Een Computermodel voor het Ondersteunen van Euthanasiebeslissingen
- Borys Omelayenko (VU) Web-Service configuration on the Semantic Web; Exploring how semantics meets pragmatics
- 15. Tibor Bosse (VU) Analysis of the Dynamics of Cognitive Processes
- 16. Joris Graaumans (UU) Usability of XML Query Languages
- 17. Boris Shishkov (TUD) Software Specification Based on Re-usable Business Components
- 18. Danielle Sent (UU) Test-selection strategies for probabilistic networks
- 19. Michel van Dartel (UM) Situated Representation
- 20. Cristina Coteanu (UL) Cyber Consumer Law, State of the Art and Perspectives
- 21. Wijnand Derks (UT) Improving Concurrency and Recovery in Database Systems by Exploiting Application Semantics

- 1. Samuil Angelov (TU/e) Foundations of B2B Electronic Contracting
- Cristina Chisalita (VU) Contextual issues in the design and use of information technology in organizations
- 3. Noor Christoph (UVA) The role of metacognitive skills in learning to solve problems
- 4. Marta Sabou (VU) Building Web Service Ontologies
- 5. Cees Pierik (UU) Validation Techniques for Object-Oriented Proof Outlines
- Ziv Baida (VU) Software-aided Service Bundling Intelligent Methods & Tools for Graphical Service Modeling

- Marko Smiljanic (UT) XML schema matching balancing efficiency and effectiveness by means of clustering
- Eelco Herder (UT) Forward, Back and Home Again Analyzing User Behavior on the Web 8.
- Mohamed Wahdan (UM) Automatic Formulation of the Auditor's Opinion
- 10. Ronny Siebes (VU) Semantic Routing in Peer-to-Peer Systems
- 11. Joeri van Ruth (UT) Flattening Queries over Nested Data Types
- 12. Bert Bongers (VU) Interactivation Towards an e-cology of people, our technological environment, and the arts
- 13. Henk-Jan Lebbink (UU) Dialogue and Decision Games for Information Exchanging Agents
- 14. Johan Hoorn (VU) Software Requirements: Update, Upgrade, Redesign towards a Theory of Requirements Change
- 15. Rainer Malik (UU) CONAN: Text Mining in the Biomedical Domain
- 16. Carsten Riggelsen (UU) Approximation Methods for Efficient Learning of Bayesian Networks
- 17. Stacey Nagata (UU) User Assistance for Multitasking with Interruptions on a Mobile Device
- 18. Valentin Zhizhkun (UVA) Graph transformation for Natural Language Processing
- 19. Birna van Riemsdijk (UU) Cognitive Agent Programming: A Semantic Approach
- 20. Marina Velikova (UvT) Monotone models for prediction in data mining
- 21. Bas van Gils (RUN) Aptness on the Web
- 22. Paul de Vrieze (RUN) Fundaments of Adaptive Personalisation
- 23. Ion Juvina (UU) Development of Cognitive Model for Navigating on the Web
- 24. Laura Hollink (VU) Semantic Annotation for Retrieval of Visual Resources
- Madalina Drugan (UU) Conditional log-likelihood MDL and Evolutionary MCMC
- 26. Vojkan Mihajlovic (UT) Score Region Algebra: A Flexible Framework for Structured Information Retrieval
- 27. Stefano Bocconi (CWI) Vox Populi: generating video documentaries from semantically annotated media repositories
- 28. Borkur Sigurbjornsson (UVA) Focused Information Access using XML Element Retrieval

- Kees Leune (UvT) Access Control and Service-Oriented Architectures 1.
- 2. Wouter Teepe (RUG) Reconciling Information Exchange and Confidentiality: A Formal Approach
- Peter Mika (VU) Social Networks and the Semantic Web
- 4. Jurriaan van Diggelen (UU) Achieving Semantic Interoperability in Multi-agent Systems: a dialoguebased approach
- 5. Bart Schermer (UL) Software Agents, Surveillance, and the Right to Privacy: a Legislative Framework for Agent-enabled Surveillance
- Gilad Mishne (UVA) Applied Text Analytics for Blogs 6.
- 7. Natasa Jovanovic' (UT) To Whom It May Concern - Addressee Identification in Face-to-Face Meetings
- Mark Hoogendoorn (VU) Modeling of Change in Multi-Agent Organizations
- 9. David Mobach (VU) Agent-Based Mediated Service Negotiation
- 10. Huib Aldewereld (UU) Autonomy vs. Conformity: an Institutional Perspective on Norms and Protocols
- 11. Natalia Stash (TU/e) Incorporating Cognitive/Learning Styles in a General-Purpose Adaptive Hypermedia System

- 12. Marcel van Gerven (RUN) Bayesian Networks for Clinical Decision Support: A Rational Approach to Dynamic Decision-Making under Uncertainty
- 13. Rutger Rienks (UT) Meetings in Smart Environments; Implications of Progressing Technology
- 14. Niek Bergboer (UM) Context-Based Image Analysis
- 15. Joyca Lacroix (UM) NIM: a Situated Computational Memory Model
- Davide Grossi (UU) Designing Invisible Handcuffs. Formal investigations in Institutions and Organizations for Multi-agent Systems
- 17. Theodore Charitos (UU) Reasoning with Dynamic Networks in Practice
- 18. Bart Orriens (UvT) On the development an management of adaptive business collaborations
- 19. David Levy (UM) Intimate relationships with artificial partners
- 20. Slinger Jansen (UU) Customer Configuration Updating in a Software Supply Network
- Karianne Vermaas (UU) Fast diffusion and broadening use: A research on residential adoption and usage of broadband internet in the Netherlands between 2001 and 2005
- 22. Zlatko Zlatev (UT) Goal-oriented design of value and process models from patterns
- 23. Peter Barna (TU/e) Specification of Application Logic in Web Information Systems
- 24. Georgina Ramírez Camps (CWI) Structural Features in XML Retrieval
- 25. Joost Schalken (VU) Empirical Investigations in Software Process Improvement

- Katalin Boer-Sorbán (EUR) Agent-Based Simulation of Financial Markets: A modular, continuous-time approach
- 2. Alexei Sharpanskykh (VU) On Computer-Aided Methods for Modeling and Analysis of Organizations
- 3. Vera Hollink (UVA) Optimizing hierarchical menus: a usage-based approach
- 4. Ander de Keijzer (UT) Management of Uncertain Data towards unattended integration
- Bela Mutschler (UT) Modeling and simulating causal dependencies on process-aware information systems from a cost perspective
- 6. Arjen Hommersom (RUN) On the Application of Formal Methods to Clinical Guidelines, an Artificial Intelligence Perspective
- 7. Peter van Rosmalen (OU) Supporting the tutor in the design and support of adaptive elearning
- 8. Janneke Bolt (UU) Bayesian Networks: Aspects of Approximate Inference
- 9. Christof van Nimwegen (UU) The paradox of the guided user: assistance can be counter-effective
- 10. Wouter Bosma (UT) Discourse oriented summarization
- 11. Vera Kartseva (VU) Designing Controls for Network Organizations: A Value-Based Approach
- 12. Jozsef Farkas (RUN) A Semiotically Oriented Cognitive Model of Knowledge Representation
- 13. Caterina Carraciolo (UVA) Topic Driven Access to Scientific Handbooks
- 14. Arthur van Bunningen (UT) Context-Aware Querying; Better Answers with Less Effort
- 15. Martijn van Otterlo (UT) The Logic of Adaptive Behavior: Knowledge Representation and Algorithms for the Markov Decision Process Framework in First-Order Domains
- 16. Henriette van Vugt (VU) Embodied agents from a user's perspective
- 17. Martin Op 't Land (TUD) Applying Architecture and Ontology to the Splitting and Allying of Enterprises
- 18. Guido de Croon (UM) Adaptive Active Vision

- 19. Henning Rode (UT) From Document to Entity Retrieval: Improving Precision and Performance of Focused Text Search
- 20. Rex Arendsen (UVA) Geen bericht, goed bericht. Een onderzoek naar de effecten van de introductie van elektronisch berichtenverkeer met de overheid op de administratieve lasten van bedrijven
- 21. Krisztian Balog (UVA) People Search in the Enterprise
- 22. Henk Koning (UU) Communication of IT-Architecture
- 23. Stefan Visscher (UU) Bayesian network models for the management of ventilator-associated pneumonia
- 24. Zharko Aleksovski (VU) Using background knowledge in ontology matching
- 25. Geert Jonker (UU) Efficient and Equitable Exchange in Air Traffic Management Plan Repair using Spender-signed Currency
- 26. Marijn Huijbregts (UT) Segmentation, Diarization and Speech Transcription: Surprise Data Unraveled
- 27. Hubert Vogten (OU) Design and Implementation Strategies for IMS Learning Design
- 28. Ildiko Flesch (RUN) On the Use of Independence Relations in Bayesian Networks
- 29. Dennis Reidsma (UT) Annotations and Subjective Machines Of Annotators, Embodied Agents, Users, and Other Humans
- 30. Wouter van Atteveldt (VU) Semantic Network Analysis: Techniques for Extracting, Representing and Querying Media Content
- 31. Loes Braun (UM) Pro-Active Medical Information Retrieval
- 32. Trung H. Bui (UT) Toward Affective Dialogue Management using Partially Observable Markov Decision Processes
- 33. Frank Terpstra (UVA) Scientific Workflow Design; theoretical and practical issues
- 34. Jeroen de Knijf (UU) Studies in Frequent Tree Mining
- 35. Benjamin Torben-Nielsen (UvT) Dendritic morphologies: function shapes structure

- Rasa Jurgelenaite (RUN) Symmetric Causal Independence Models 1.
- 2. Willem Robert van Hage (VU) Evaluating Ontology-Alignment Techniques
- 3. Hans Stol (UvT) A Framework for Evidence-based Policy Making Using IT
- 4. Josephine Nabukenya (RUN) Improving the Quality of Organisational Policy Making using Collaboration Engineering
- 5. Sietse Overbeek (RUN) Bridging Supply and Demand for Knowledge Intensive Tasks - Based on Knowledge, Cognition, and Quality
- Muhammad Subianto (UU) Understanding Classification
- 7. Ronald Poppe (UT) Discriminative Vision-Based Recovery and Recognition of Human Motion
- 8. Volker Nannen (VU) Evolutionary Agent-Based Policy Analysis in Dynamic Environments
- 9. Benjamin Kanagwa (RUN) Design, Discovery and Construction of Service-oriented Systems
- 10. Jan Wielemaker (UVA) Logic programming for knowledge-intensive interactive applications
- 11. Alexander Boer (UVA) Legal Theory, Sources of Law & the Semantic Web
- 12. Peter Massuthe (TU/e, Humboldt-Universitaet zu Berlin) Operating Guidelines for Services
- 13. Steven de Jong (UM) Fairness in Multi-Agent Systems
- 14. Maksym Korotkiy (VU) From ontology-enabled services to service-enabled ontologies (making ontologies work in e-science with ONTO-SOA)
- 15. Rinke Hoekstra (UVA) Ontology Representation Design Patterns and Ontologies that Make Sense

- 16. Fritz Reul (UvT) New Architectures in Computer Chess
- 17. Laurens van der Maaten (UvT) Feature Extraction from Visual Data
- 18. Fabian Groffen (CWI) Armada, An Evolving Database System
- Valentin Robu (CWI) Modeling Preferences, Strategic Reasoning and Collaboration in Agent-Mediated Electronic Markets
- 20. Bob van der Vecht (UU) Adjustable Autonomy: Controling Influences on Decision Making
- 21. Stijn Vanderlooy(UM) Ranking and Reliable Classification
- 22. Pavel Serdyukov (UT) Search For Expertise: Going beyond direct evidence
- 23. Peter Hofgesang (VU) Modelling Web Usage in a Changing Environment
- 24. Annerieke Heuvelink (VU) Cognitive Models for Training Simulations
- 25. Alex van Ballegooij (CWI) "RAM: Array Database Management through Relational Mapping"
- 26. Fernando Koch (UU) An Agent-Based Model for the Development of Intelligent Mobile Services
- 27. Christian Glahn (OU) Contextual Support of social Engagement and Reflection on the Web
- 28. Sander Evers (UT) Sensor Data Management with Probabilistic Models
- 29. Stanislav Pokraev (UT) Model-Driven Semantic Integration of Service-Oriented Applications
- 30. Marcin Zukowski (CWI) Balancing vectorized query execution with bandwidth-optimized storage
- 31. Sofiya Katrenko (UVA) A Closer Look at Learning Relations from Text
- 32. Rik Farenhorst (VU) and Remco de Boer (VU) Architectural Knowledge Management: Supporting Architects and Auditors
- 33. Khiet Truong (UT) How Does Real Affect Affect Recognition In Speech?
- Inge van de Weerd (UU) Advancing in Software Product Management: An Incremental Method Engineering Approach
- 35. Wouter Koelewijn (UL) Privacy en Politiegegevens; Over geautomatiseerde normatieve informatieuitwisseling
- 36. Marco Kalz (OU) Placement Support for Learners in Learning Networks
- 37. Hendrik Drachsler (OU) Navigation Support for Learners in Informal Learning Networks
- Riina Vuorikari (OU) Tags and self-organisation: a metadata ecology for learning resources in a multilingual context
- Christian Stahl (TU/e, Humboldt-Universitaet zu Berlin) Service Substitution ~ A Behavioral Approach Based on Petri Nets
- 40. Stephan Raaijmakers (UvT) Multinomial Language Learning: Investigations into the Geometry of Language
- 41. Igor Berezhnyy (UvT) Digital Analysis of Paintings
- 42. Toine Bogers (UvT) Recommender Systems for Social Bookmarking
- Virginia Nunes Leal Franqueira (UT) Finding Multi-step Attacks in Computer Networks using Heuristic Search and Mobile Ambients
- 44. Roberto Santana Tapia (UT) Assessing Business-IT Alignment in Networked Organizations
- 45. Jilles Vreeken (UU) Making Pattern Mining Useful
- 46. Loredana Afanasiev (UvA) Querying XML: Benchmarks and Recursion

- 1. Matthijs van Leeuwen (UU) Patterns that Matter
- 2. Ingo Wassink (UT) Work flows in Life Science
- Joost Geurts (CWI) A Document Engineering Model and Processing Framework for Multimedia 3. documents
- 4. Olga Kulyk (UT) Do You Know What I Know? Situational Awareness of Co-located Teams in Multidisplay Environments
- 5. Claudia Hauff (UT) Predicting the Effectiveness of Queries and Retrieval Systems
- Sander Bakkes (UvT) Rapid Adaptation of Video Game AI
- 7. Wim Fikkert (UT) A Gesture interaction at a Distance
- 8. Krzysztof Siewicz (UL) Towards an Improved Regulatory Framework of Free Software. Protecting user freedoms in a world of software communities and eGovernments
- Hugo Kielman (UL) Politiële gegevensverwerking en Privacy, Naar een effectieve waarborging
- 10. Rebecca Ong (UL) Mobile Communication and Protection of Children
- 11. Adriaan Ter Mors (TUD) The world according to MARP: Multi-Agent Route Planning
- 12. Susan van den Braak (UU) Sensemaking software for crime analysis
- 13. Gianluigi Folino (RUN) High Performance Data Mining using Bio-inspired techniques
- 14. Sander van Splunter (VU) Automated Web Service Reconfiguration
- 15. Lianne Bodenstaff (UT) Managing Dependency Relations in Inter-Organizational Models
- 16. Sicco Verwer (TUD) Efficient Identification of Timed Automata, theory and practice
- 17. Spyros Kotoulas (VU) Scalable Discovery of Networked Resources: Algorithms, Infrastructure, **Applications**
- 18. Charlotte Gerritsen (VU) Caught in the Act: Investigating Crime by Agent-Based Simulation
- 19. Henriette Cramer (UvA) People's Responses to Autonomous and Adaptive Systems
- 20. Ivo Swartjes (UT) Whose Story Is It Anyway? How Improv Informs Agency and Authorship in Emergent Narrative
- 21. Harold van Heerde (UT) Privacy-aware data management by means of data degradation
- 22. Michiel Hildebrand (CWI) End-user Support for Access to \\ Heterogeneous Linked Data
- 23. Bas Steunebrink (UU) The Logical Structure of Emotions
- 24. Dmytro Tykhonov (TUD) Designing Generic and Efficient Negotiation Strategies
- 25. Zulfigar Ali Memon (VU) Modelling Human-Awareness for Ambient Agents: A Human Mindreading Perspective
- 26. Ying Zhang (CWI) XRPC: Efficient Distributed Query Processing on Heterogeneous XQuery Engines
- 27. Marten Voulon (UL) Automatisch contracteren
- 28. Arne Koopman (UU) Characteristic Relational Patterns
- 29. Stratos Idreos (CWI) Database Cracking: Towards Auto-tuning Database Kernels
- 30. Marieke van Erp (UvT) Accessing Natural History Discoveries in data cleaning, structuring, and retrieval
- 31. Victor de Boer (UVA) Ontology Enrichment from Heterogeneous Sources on the Web
- 32. Marcel Hiel (UvT) An Adaptive Service Oriented Architecture: Automatically solving Interoperability Problems
- 33. Robin Aly (UT) Modelling Representation Uncertainty in Concept-Based Multimedia Retrieval
- 34. Teduh Dirgahayu (UT) Interaction Design in Service Compositions

- 35. Dolf Trieschnigg (UT) Proof of Concept: Concept-based Biomedical Information Retrieval
- 36. Jose Janssen (OU) Paving the Way for Lifelong Learning; Facilitating competence development through a learning path specification
- 37. Niels Lohmann (TU/e) Correctness of services and their composition
- 38. Dirk Fahland (TU/e) From Scenarios to components
- 39. Ghazanfar Farooq Siddiqui (VU) Integrative modelling of emotions in virtual agents
- 40. Mark van Assem (VU) Converting and Integrating Vocabularies for the Semantic Web
- 41. Guillaume Chaslot (UM) Monte-Carlo Tree Search
- 42. Sybren de Kinderen (VU) Needs-driven service bundling in a multi-supplier setting the computational e3-service approach
- Peter van Kranenburg (UU) A Computational Approach to Content-Based Retrieval of Folk Song Melodies
- 44. Pieter Bellekens (TU/e) An Approach towards Context-sensitive and User-adapted Access to Heterogeneous Data Sources, Illustrated in the Television Domain
- 45. Vasilios Andrikopoulos (UvT) A theory and model for the evolution of software services
- 46. Vincent Pijpers (VU) e3alignment: Exploring Inter-Organizational Business-ICT Alignment
- 47. Chen Li (UT) Mining Process Model Variants: Challenges, Techniques, Examples
- 48. Milan Lovric (EUR) Behavioral Finance and Agent-Based Artificial Markets
- 49. Jahn-Takeshi Saito (UM) Solving difficult game positions
- 50. Bouke Huurnink (UVA) Search in Audiovisual Broadcast Archives
- 51. Alia Khairia Amin (CWI) Understanding and supporting information seeking tasks in multiple sources
- 52. Peter-Paul van Maanen (VU) Adaptive Support for Human-Computer Teams: Exploring the Use of Cognitive Models of Trust and Attention
- 53. Edgar Meij (UVA) Combining Concepts and Language Models for Information Access

- 1. Botond Cseke (RUN) Variational Algorithms for Bayesian Inference in Latent Gaussian Models
- Nick Tinnemeier (UU) Organizing Agent Organizations. Syntax and Operational Semantics of an Organization-Oriented Programming Language
- Jan Martijn van der Werf (TU/e) Compositional Design and Verification of Component-Based Information Systems
- Hado van Hasselt (UU) Insights in Reinforcement Learning Formal analysis and empirical evaluation of temporal-difference learning algorithms
- Base van der Raadt (VU) Enterprise Architecture Coming of Age Increasing the Performance of an Emerging Discipline
- Yiwen Wang (TU/e) Semantically-Enhanced Recommendations in Cultural Heritage
- 7. Yujia Cao (UT) Multimodal Information Presentation for High Load Human Computer Interaction
- 8. Nieske Vergunst (UU) BDI-based Generation of Robust Task-Oriented Dialogues
- 9. Tim de Jong (OU) Contextualised Mobile Media for Learning
- 10. Bart Bogaert (TU) Cloud Content Contention
- 11. Dhaval Vyas (UT) Designing for Awareness: An Experience-focused HCI Perspective
- 12. Carmen Bratosin (TU/e) Grid Architecture for Distributed Process Mining

- 13. Xiaoyu Mao (UvT) Airport under Control. Multiagent Scheduling for Airport Ground Handling
- Milan Lovric (EUR) Behavioral Finance and Agent-Based Artificial Markets
- Marijn Koolen (UVA) The Meaning of Structure: the Value of Link Evidence for Information Retrieval
- 16. Maarten Schadd (UM) Selective Search in Games of Different Complexity
- 17. Jiyin He (UVA) Exploring Topic Structure: Coherence, Diversity and Relatedness
- Mark Ponsen (UM) Strategic Decision-Making in Complex Games
- Ellen Rusman (OU) The Mind's Eye on Personal Profiles
- Qing Gu (VU) Guiding service-oriented software engineering A view-based approach
- Linda Terlouw (TUD) Modularization and Specification of Service-Oriented Systems
- 22. Junte Zhang (UVA) System Evaluation of Archival Description and Access
- 23. Wouter Weerkamp (UVA) Finding People and their Utterances in Social Media
- 24. Herwin van Welbergen (UT) Behavior Generation for Interpersonal Coordination with Virtual Humans On Specifying, Scheduling and Realizing Multimodal Virtual Human Behavior
- 25. Syed Wagar ul Qounain Jaffry (VU) Analysis and Validation of Models for Trust Dynamics
- 26. Matthijs Aart Pontier (VU) Virtual Agents for Human Communication Emotion Regulation and Involvement-Distance Trade-Offs in Embodied Conversational Agents and Robots
- 27. Aniel Bhulai (VU) Dynamic website optimization through autonomous management of design patterns
- 28. Rianne Kaptein (UVA) Effective Focused Retrieval by Exploiting Query Context and Document Structure
- 29. Faisal Kamiran (TUE) Discrimination-aware Classification
- 30. Egon van den Broek (UT) Affective Signal Processing (ASP): Unraveling the mystery of emotions
- Ludo Waltman (EUR) Computational and Game-Theoretic Approaches for Modeling Bounded Rationality
- 32. Nees-Jan van Eck (EUR) Methodological Advances in Bibliometric Mapping of Science
- Tom van der Weide (UU) Arguing to Motivate Decisions
- 34. Paolo Turrini (UU) Strategic Reasoning in Interdependence: Logical and Game-theoretical Investigations
- 35. Maaike Harbers (UU) Explaining Agent Behavior in Virtual Training
- Erik van der Spek (UU) Experiments in serious game design: a cognitive approach
- 37. Adriana Burlutiu (RUN) Machine Learning for Pairwise Data, Applications for Preference Learning and Supervised Network Inference
- 38. Nyree Lemmens (UM) Bee-inspired Distributed Optimization
- Joost Westra (UU) Organizing Adaptation using Agents in Serious Games
- Viktor Clerc (VU) Architectural Knowledge Management in Global Software Development
- 41. Luan Ibraimi (UT) Cryptographically Enforced Distributed Data Access Control
- Michal Sindlar (UU) Explaining Behavior through Mental State Attribution
- Henk van der Schuur (UU) Process Improvement through Software Operation Knowledge
- Boris Reuderink (UT) Robust Brain-Computer Interfaces
- 45. Herman Stehouwer (UvT) Statistical Language Models for Alternative Sequence Selection
- Beibei Hu (TUD) Towards Contextualized Information Delivery: A Rule-based Architecture for the Domain of Mobile Police Work
- 47. Azizi Bin Ab Aziz (VU) Exploring Computational Models for Intelligent Support of Persons with Depression

- 1. Najja Terry Kakeeto (UvT) Relationship Marketing for SMEs in Uganda
- 2. Muhammad Umair (VU) Adaptivity, emotion, and Rationality in Human and Ambient Agent Models
- 3. Adam Vanya (VU) Supporting Architecture Evolution by Mining Software Repositories
- 4. Jurriaan Souer (UU) Development of Content Management System-based Web Applications
- 5. Marijn Plomp (UU) Maturing Interorganisational Information Systems
- 6. Wolfgang Reinhardt (OU) Awareness Support for Knowledge Workers in Research Networks
- Rianne van Lambalgen (VU) When the Going Gets Tough: Exploring Agent-based Models of Human Performance under Demanding Conditions
- 8. Gerben de Vries (UVA) Kernel Methods for Vessel Traject
- 9. Ricardo Neisse (UT) Trust and Privacy Management Support for Context-Aware Service Platforms
- 10. David Smits (TUE) Towards a Generic Distributed Adaptive Hypermedia Environment
- 11. J.C.B. Rantham Prabhakara (TUE) Process Mining in the Large: Preprocessing, Discovery, and Diagnostics
- Kees van der Sluijs (TUE) Model Driven Design and Data Integration in Semantic Web Information Systems
- 13. Suleman Shahid (UvT) Fun and Face: Exploring non-verbal expressions of emotion during playful interactions

